The Facilitation of Trust in Automation: A Qualitative Study of Behaviour and Attitudes Towards Emerging Technology in Military Culture

by

Megan Field

A thesis submitted to

The University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Engineering
Department of Electronic, Electrical and Systems Engineering
The University of Birmingham
March 2020

ABSTRACT

High speciality and criticality domains categorise the most researched areas in the field of Trust in Automation. Minimal studies have explored the nuances of the psycho-social environment and organisational culture in the development of appropriate mental models on dispositional trust. To aid integration of human operators with emergent specialised systems, there is ambition to introduce Human-Human/Human-System analogies with AI Avatars and 3D representations of environments (Ministry of Defence, 2018). Due to the criticisms in the literature of Human-Human and Human-System teaming analogues this research has explored personal narratives of civilians and military personnel about technology, adaptability and how to facilitate beneficial attitudes and behaviours in appropriate trust, reliance and misuse. A subdivision of the research explores the socio-cultural idiosyncrasies within the different echelons of the military as variances in authority and kinship provide insight on informing training targeted to unique domains. The thesis proposes that there are core hindrances to tacit trust facilitation with automation as cognitive rigidity towards individual and group identities impact socially constructed social responses and internal mental models. Furthermore, as automation broaches category boundaries there may be resistance and discomfort as a result of unpredictable social contracts whereby transactional and relational trust-related power dynamics are unknown or unpredictable.

DEDICATION

To my sickly body – you have been my greatest challenge, my albatross, and my greatest teacher.

You are the determination in every page.

ACKNOWLEDGEMENTS

Thank you to Prof. Chris Baber and Prof. Bob Stone for their encouragement and supervision in the completion of this thesis. I appreciate the support received by Colin Etting for their cooperation and assistance in facilitating the data collection for this thesis. I gratefully acknowledge the funding received towards my PhD from EPSRC and BAE for this industrial CASE studentship.

A very special thank you to Sharon for their invaluable advice, feedback and kind words. From formatting, proof-reading to helping keep my thoughts in order, you have been the one constant through my doctoral journey. Thank you for being my study support and my friend.

My deepest appreciation to the HIT Team, Rob and Andy; Drone Soc pals and the office 311 comrades – thank you for all the conversations, laughs and companionship over the years. Especially to Jo and our rekindled friendship.

I am indebted to my friends and family who have been by my side, for their support, their encouragement and unwavering cheerleading. Without them I would not have had the courage, resilience or fortitude to have embarked on the challenges the journey of the last five years has thrown at me. To my little acorns especially, Fortitudine Crescamus – May we grow in strength, always together.

To my mum, dad, my beloved sister Laura, and Sophie; a heartfelt thank you for believing in me and the steadfast pride you have always had in me.

And finally, to Martin. Your tolerance, patience and indomitable adoration makes me think you are a product of my imagination. Thank you for everything you are.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACC	Adaptive Cruise Control
AI	Artificial Intelligence
AMVE	Augmented, Mixed and Virtual Environment
ANT	Actor Network Theory
ATCO	Air Traffic Control Officer
AWMS	Automated Warehouse Management System
AWO	Assistant Warfare Officer
BAE	BAE Systems Ltd
CAD	Computer Aided Design
CAQDAS	Computer Assisted Qualitative Data Analysis Software
C2	Command and Control (Centre)
C3I	Command, Control, Communication and Intelligence
CMS	Combat Management System
CoTS	Commercial of the Shelf
CPO	Chief Petty Officer
DiE	Design Induced Errors
DRI	Detection, Recognition and Identification (Task)
EASC	Embodied Active Situated Cognition
EPSRC	Engineering and Physical Sciences Research Council
EU	Europe
HAI	Human Automation Interaction
HCA	Hierarchical Content Analysis
HCI/HSI	Human Computer Interaction/Human System Interaction
HCT	Human Computer Trust
H-H	Human to Human (teammate)

HK	Hong Kong
HM*	Her Majesty's Armed Forces
HMM	Human Mental Models
H-S/H-T	Human to System/Human to Technology (teammate)
HSI	Human System Interaction
IPA	Interpretive Phenomenological Analysis
LOA	Level of Automation
MMDS	Mental Models in Dynamic Systems
MOOSE	Meta-analysis of Observational Studies in Epidemiology
MoD	Ministry of Defence
MoTS	Modified of the Shelf
NASA-TLX	NASA Task Load Index
OCEAN	Openness to Experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
Ps	Participants
PO	Petty Officer
RAF	Royal Air Force
RQ #	Research Question
SA	Situational Awareness
SART	Situational Awareness Rating Technique
SDT	Self-Determinism Theory
S-E	Socio-emotional
SME	Subject Matter Expert / System Matter Expert
SMM	Shared Mental Models
STEC	Spatial, Temporal and Environmental Cues
STEM	Science, Technology, Engineering and Mathematics
STS	Socio-Technical System/s

TAM	Trust Acceptance Model
TIA	Trust in Automation
UAV	Unmanned Air Vehicle
UGV	Unmanned Ground Vehicle
UK	United Kingdom
VLE	Virtual Locomotion Environment
WO	Warfare Officer

LIST OF APPENDICES

JOURNAL PUBLICATIONS & CONFERENCE PAPERS

[1] *"Review of Mental Models as a Method of Trust Facilitation for Human System Interaction," Technical Paper Presented at 2018 Defence and Security Doctoral Symposium (DSDS18), Swindon, England. 13-15 October 2018.*

[2] *"Building risk matrices from interview transcripts utilising HCA and IPA" in the Proceedings of Contemporary Ergonomics and Human Factors 2019, Stratford-Upon-Avon, England, 29-1 April May. https://publications.ergonomics.org.uk/uploads/Building-risk-matrices-from-interview-transcripts-utilising-HCA-and-IPA.pdf*

[3] *"Trust in Automation: How this is shaped by the human operator and the underwater domain" paper presented at UDT 2019: Undersea Defence Technology Annual Conference, Stockholm, Sweden, 13-15 May 2019.*

Page Left Intentionally Blank

1 INTRODUCTION

1.1 INTRODUCTION

The research proposes to explore the trust narratives towards automation and emerging technology and how this might affect appropriate Human System Interaction (HSI) for future development and facilitation of forthcoming Command and Control (C2) systems. The predictions of human behaviour can provide both insight and design recommendations to align with future trends and technological development.

The implementation and development of future command and control environments pose new user impacts and design recommendation gaps. Current and future trends seek to use automation to reduce workload and increase efficiency, in collaboration with potential novel user interfaces that may utilise emerging technology. However, the consequences of these emergent specialised systems on the human operator are under-researched within the literature.

The aim of this thesis is to explore the trust factors which may influence appropriate facilitation with artificial actants. Furthermore, to explore the concept of trust and how the unique psycho-social technical system dynamics and intra-organisational inter-personal relationships therein of HM Armed Forces shape operator Human-Automation Interaction (HAI).

The reignition of interest in trust research has increased markedly in the last decade. The growth has become a *"transdisciplinary topic with many interdisciplinary opportunities [with] multidisciplinary outputs"* (Möllering, 2017, p. 107). Therefore, to strategically facilitate trust in automation, the role and concept of trust must be defined. For example, McEvily and Tortoriello (2011) identified 38 distinct trust-relevant constructs in an analysis of organisational literature. The factors, determinants and dimensions of trust are fragmented within the literature (PytlikZillig, et al., 2016). Therefore, a key concern of this thesis is to elaborate on the crucial trust elements and expressions within the Military sociotechnical system. Furthermore, interpret and elaborate on the trust narratives which may shape and

influence the development of inter-personal, intra-organisational relationships with human and system teammates. Therefore, the first research question must be:

Research Question 1 (RQ1): How do the narratives that people tell about automation relate to their trust in technology?

This thesis seeks to investigate the factors influencing trust and their role in shaping the mental models of operators. This will be investigated through exploration of the worldview held by operators, and how they conceptualise trust expressions. Narrative research explores the lived experiences and stories of participants to examine the underlying phenomena.

To understand the narrative, it is important to provide structure for sensemaking of the ambiguity and complexity of the phenomena. The following section seeks to discuss and define what is military culture and the different defining characteristics of the holarchy. Also, to provide the framework into exploration of military narratives, the values and beliefs which may shape their interpretations, and background information into military social psychology.

1.2 BACKGROUND INFORMATION: MILITARY CULTURE

1.2.1 Chaoplexic Warfare and Narrative Framing

With rapid advancements in defence technologies, some research has raised concerns on the vulnerabilities of the utilisation of automation, system-teammates and digital native operators in the existing regimented battlespace (Crosston, 2017). An example of the complexities has been proposed by Bousquet explains (2008, p. 915):

> *"Ours is the age of the network. Whether framed in socio-economic, technological or ideological terms, our present times are increasingly seen as characterised by the rise and spread of fluid decentralised forms of social organisation in which information and communication devices play a key role…technology certainly contributes to defining the space of possible social formations and practical organisational*

arrangements, it remains nonetheless primarily a tool which is given meaning and put to specific uses within a broader socio-cultural setting"

The framework of Chaoplexic Warfare is beneficial to explore the past and future sociotechnical variations within military organisation and key technologies and scientific concepts within certain forms of warfare (refer to Table 1). Bousquet outlines each of the key technologies and their corresponding warfare aids to frame the different types of battlespace. For example, mechanistic warfare is framed as predictable, controlled, synchronised and utilises simple technologies, such as a clock.

Table 1 - The four regimes of the scientific way of warfare [Bousquet, 2008 (pg. 917)]

	Mechanism	**Thermodynamics**	**Cybernetics**	**Chaoplexity**
Key Technology	Clock	Engine	Computer	Network
Scientific Concepts	Force; matter in motion; linearity; geometry	Energy; entropy; probability	Information; negentropy; negative feedback; homeostasis	Information; non-linearity; positive feedback; self-organisation; emergence
Form of Warfare	Close order drill; rigid tactical deployments	Mass mobilisation; motorisation; industrialisation	Command and control; automation	Decentralisation; swarming

This thesis utilises Chaoplexic definitions as a cognitive tool to frame, and thus better describe larger and more complex concepts. The use of narrative frameworks within the military is a growing area of research as there are beneficial cognitive effects of communicating information effectively and efficiently using strategic communications like these (Finlayson & Corman, 2013). Using simple concepts as a carrier, more diverse and complex communications can be used to shape, guide and influence team interactions. Furthermore, as transmissions become more decentralised, technologically mediated or influenced by other asymmetries, simplifying complex information exchange is more imperative than ever for maintaining teammate cooperation.

1.2.2 Characteristics of Military Culture

The definition of culture illustrates different concepts across interdisciplinary fields and suffers from different connotations in different paradigms (Ajiferuke & Boddewyn, 1970, p. 154). Within this thesis, the fuzzy concept (e.g. due to psychogenic differences) of culture is defined as the learned and shared meanings, ideas and symbols that distinguish a group or category of people (Hofstede, 1991; Soeters, et al., 2006). The concept of culture is often used in the transnational sense, however the effects within organisational cultures are often underrepresented in the literature (Hoff & Bashir, 2015). One of the most distinct organisational microcosms cross-culturally is the military. Military culture is characterised by robust hierarchical order which enacts established rules, utilises high levels of discipline and compliance with regulations and acceptance of authority (Pak, et al., 2017, p. 202). Recent literature purports that the intense training and indoctrination of personnel might have a psychosocial effect on expectations and behaviour with system teammates (Kennedy, et al., 2015). However, it is debated whether the comradery may create inappropriate trust and the maladaptive responses cause over-reliance or under-reliance. For example, if blind trust is given to automated teammates, this may lead to immoderate or high expectations of automated teammates capability and impact situational awareness (Stanton, 2011).

As is noted in Chapter 3 (Systematic Review of Trust, Automation and Mental Models), many studies examining military culture often use college students for convenience (Wang, et al., 2009; Dzindolet, et al., 2001). A core principle in this thesis is to explore if generalising studies to a military population is appropriate, as the role of organisational culture on expressions of trust and reliance may influence attitudes towards automation.

Studies by Pak *et al* (2017) showed trust in automation was often highly dependent on the sociocultural domain, the maturity and reliability of automation and most importantly, the user group. They found students could calibrate trust ratings dependent on system reliability. However, older adults and military cadets could not. It was expected that older participants were unable to calibrate trust appropriately (Ho, et al., 2005), but the outcomes with cadets was not anticipated. They propose that cultural differences

could be the attributable variable and that the sociocultural impact of development within a military organisation is a source of influence. Furthermore, they documented a noteworthy effect of stages of automation on trust, whereby cadets had significantly lowered trust for decision automation across all the domains where this was absent in civilian students.

As this is an emerging pedagogy, specifics of the differences are unknown between military and civilian populations. The following sections attempt to cover some of the prior research and literature on military social psychology.

1.2.3 Social Cohesion and Military Psychology

> *"Any social unit that has…shared history will have evolved a culture. The strength of that culture depends on the length of time, and the emotional intensity of the actual historical experiences they have shared…there may be major turnover in the leaders or members, the mission or primary task may change, the underlying technology in which the group is built may evolve. (Schein, 2010, pp. 17-18)*

Organisational culture, particularly those with distinct and commanding leadership, are built upon social patterning, external adaption and internal integration. There are further distinctions in national cultures that include complexities such *as "[historical] knowledge, belief, art, morals, law, [and] customs"* (Tylor, 1871; Peoples & Bailey, 2011, p. 22). Despite this however*, "cultural, religious, and ethnic diversity within the military, the military is a culture in its own right"* (Fenell, 2008). The emphasis on the distinctions outlined above, is due to the somewhat limited selection of psycho-social research utilising the British Armed Forces, compared to the vast majority of studies utilising the US Armed Forces (Kirke, 2010). The unique idiosyncrasies of the national cultures and the within-group social attitudes, beliefs, behaviour and values, all shape the worldview and schema of the personnel and organisation (Spencer-Oatey, 2004, p. 4).

A recent interest in the literature has sought to explore the transition experiences of British Military Veterans and how separation from these robust social units has an effect on social identity (Binks & Cambridge, 2018; Cooper, et al., 2018). For example, readapting obstinate or persistent worldviews and

reintroduction to civilian culture (Hogg & Reid, 2006; Oakes, 2011; Akerlof, 1997). Furthermore, there is a separation effect and residual impact of ingroup cooperation and unit trust (Brewer, 2008). There are key barriers, breakdowns in support and entire organisations devoted to helping leavers reidentify self (Heal, et al., 2019).

Regarding trust in military units, there is an imperative for team units to rely on each other, to accurately share information and to cooperate with teammates. To form strong, effective teams where mutually assured success is the ideal, organisations can build trust through group cohesion. In the military this has be defined as "*the bonding together of members of an organisation/unit in such a way as to sustain their will and commitment to each other, their unit, and the mission"* (Johns, et al., 1984, p. 4). This can be enabled through formally sanctioned military activities, or through unofficial subculture behaviour (Hockey, 2016). Examples of formal commands include executed orders from higher ranking teammates which are embedded in the organisational elements through discipline enforcement and official responsibility. Informal commands are those behaviours entrenched in off-duty contexts, social activities which are unrestricted by authoritative restrictions (Kirke, 2010).

There are also loyalty structures within Military cohesion, partly in response to the in-group 'belonging' on all levels of the organisational hierarchy. Houppert (2007, p. 84) posits that early recruitment replaces adolescent dependence from family unit-based kinship towards dependence on the military unit kin; *"the soldier must learn that [they] can trust no one but [their] buddies".* Military identity and kinship have been explored in the British Armed Forces through situated narrative accounts and found that soldier identity was inextricably linked to collective, group endeavours. That these kinship relationships are often relayed in dynamics stronger or closer than immediate family (Woodward & Jenkings, 2011).

There are additional levels of cohesion at the personal, organisational and institutional level. Similarly, to sociotechnical systems models, there are interactions at the individual (peer), management (Company leader), organisational (Battalion leader) and institutional (Regiment leader) level. Figure 1 illustrates some of the psychosocial factors which play a role in group cohesion and the command structure.

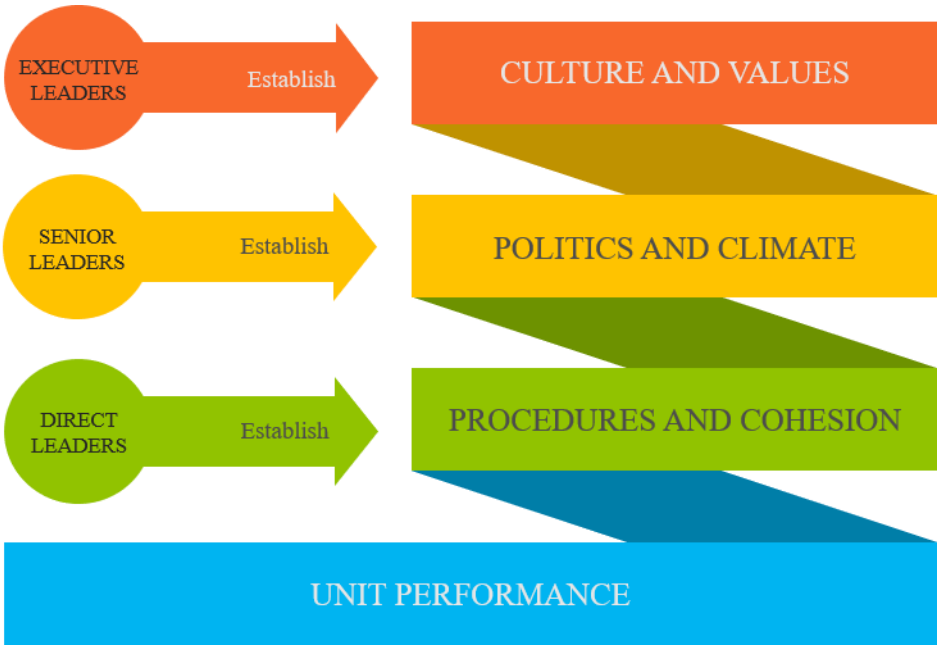

Figure 1 - Factors and how they affect hierarchy on unit performance (Jacobs, 1991)

Furthermore, it is documented in formal management research, that in developing group cohesion that specific actions or commands by leaders can alter the attitudes or behaviours to increase conformity, compliance and unit effectiveness. Table 2 outlines the Stage Model for Unit Development (Bartone & Kirkland, 1991, p. 397) which explores the characteristics required to influence social change and obedience of team members.

Table 2 – Stage Model for Excellent Unit Development (Adapted from Bartone & Kirkland [1991; p397])

	Stage 1: **Neonatal Unit**	***Stage 2:*** **Adolescent Unit**	***Stage 3:*** **Maturing Unit**	***Stage 4:*** **Excellent Unit**
Key Development Task	Peer bonding; Establish horizontal cohesion and trust	Leader- led bonding; Establish vertical trust and cohesion	Team proficiency; Consolidation and integration	Generativity; Maintain excellence
Dominant Process of Social Influence	Conformity/compliance with peer reference group	Identification with unit leaders	Internalisation of organisational values	Continued striving for growth/challenge
Critical Leader Characteristics	*Key Leader Behaviour for Each Stage*			
Competence	Master basic skills and tasks	Demonstrate ability to manage activities	Demonstrate proficiency in tactics/manoeuvres	Organise creative and challenging tasks
Caring	Actively show concern for health and wellbeing	Teach self-sufficiency	Organise systems to care for dependents	Teach caring leadership and skills
Respect	Treat members with dignity	Provide predictable schedules/hours	Decentralise power/authority	Treat unit as colleagues
Commitment	Share member's sense of mission	Give training priority over staged shows	Use errors and failures as learning tools	Keep the focus on skill retention
Chief Sources of Feedback for Leaders	Unit members are mutually supportive	Members apply to unit norms themselves	Members seek out new challenges	Members feel special; use unique methods

Team management and trusted leadership has an indirect positive effect on team cooperation, cohesiveness and subordinate trust (Chiu & Chiang, 2019). Furthermore, team cohesiveness continuously shows better, more efficient performance outcomes than those with less cohesive teams (Goodwin, et al., 2018). This is, in some measure, as a consequence of shared understanding and shared mental models.

Mental models are affected by feedback from external stimulus in the environment and the shared culture system in which the individual is embedded. Exploratory mental models allow individuals to cognitively calculate expected outcomes which are shaped through prior experiences, social heuristics or emotional responses. Culture, through the lens of human mental models, is often seen as intergenerational transfer of past experiences. The use, therefore of shared mental models, can reduce the cognitive load in complex information processing (see Figure 2 (Ostrom, 2005)). There are limitations in working memory capacity, especially within technology mediated teams and human-system interaction (Sparkes & Huf, 2003). The influx of high cognitive load information in the form of data communication has an impact on naturalistic decision making (Klein, 2008) in both individuals and teams. Furthermore, in these technology mediated spaces in the military, there is a required increase of speed of command decisions. Human-system teams adds complexity to cognitive load through building mental models with a 'black box' system teammate (Castelvecchi, 2016) – of unknown perceived and actual capacity and social compatibility issues (Flemisch, et al., 2008). Recent literature suggests that a 'grey box' or sufficiently transparent system can provide users with the underlying processes for adequately learning information to establish trust (Christensen & Lyons, 2017).

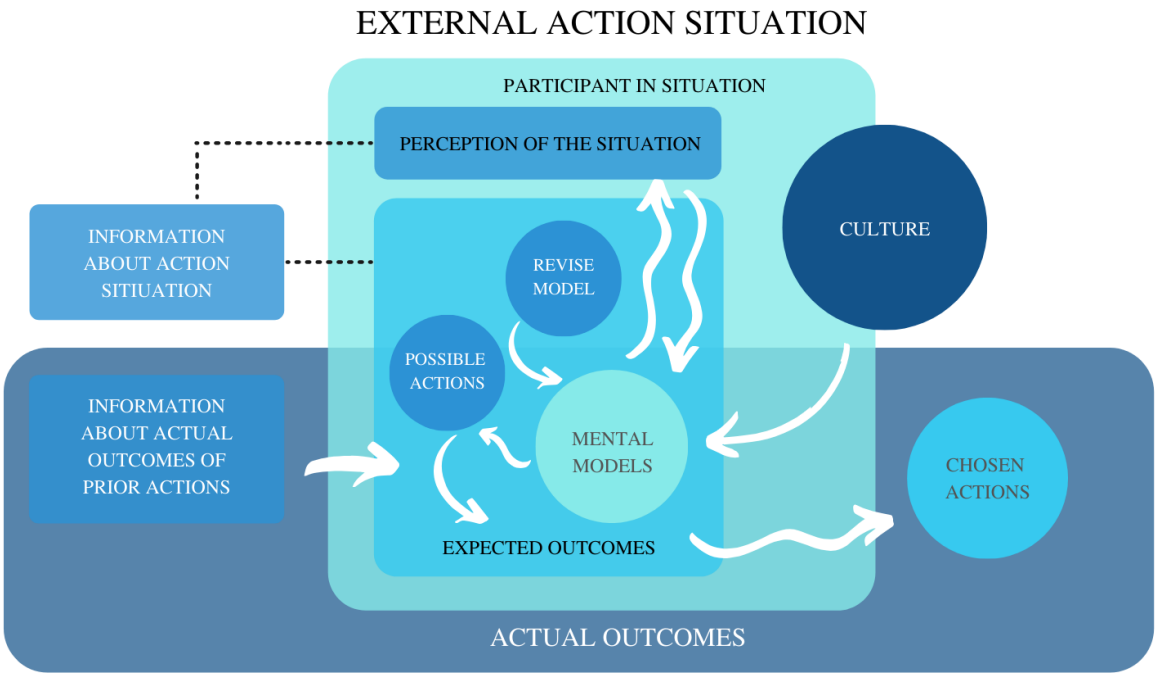

Figure 2 - Relationship between information, actions and internal mental models (redrawn from Ostrom [2005, p. 105])

There are a number of human factors issues regarding the optimisation of trust with autonomous teammates. Firstly, there are problems with the intensified amount of sensor and data processing, and how this is communicated to human teammates. Trust calibration is linked to, as above, the lack of transparency with non-human actants. Transparency is difficult at these increased levels of machine complexity. Secondly, the accommodation of errors and the impact on trust regarding system reliability (e.g. whether the error is low-level failures or high-level deficiencies) can cause tension and distrust between human-system teammates. Lastly, these mixed actor-actant teams may cause novel cognitive challenges which can be disruptive to teaming, decision making and situational awareness in military operations (Stokes, 2017).

1.2.4 Summary

> *"These visions of the future go well beyond the traditional brittle automation paradigms with limited flexibility or confidence for human operators. As such, a novel, multimodal, interdisciplinary and system-driven approaches are required…Greater*

attention will need to be given to the human side of the equation. As collaboration and teaming are social endeavours that require socio-emotional (S-E) skills for optimal functioning, understanding leveraging and training for S-E will be critical for successful human-machine collaboration." (Matthews, et al., 2019)

Some of the key issues raised in this subsection are concerned with system cohesion in human-automation teaming. Whether this is mediating the operator's internal sense of control and competence (Stokes, 2017), or the unique socio-cultural impact which pervades in military personnel long after service, and other mechanisms in which trust is formed. Therefore, a principle research question to explore, is as follows:

Research Question 2 (RQ2): How does trust in automation differ between military personnel and civilians?.

From the information provided in this section, there is a significant impact culture, social reality and cognitive processes have on human-system trust facilitation and calibration. Therefore, the factors influencing these, and how they relate to trust dynamics and mechanisms are imperative to explore in this thesis.

1.3 RESEARCH QUESTIONS

The aim of the research is to explore the underlying narratives impacting on trust facilitation in HM Armed Forces and how the sociotechnical culture has an impression on their attitude and behaviour in comparison to civilian personnel. There is a specific focus on Sub-Surface operators due to the idiosyncrasies of the echelon from a sociotechnical cultural perspective within the military.

The thesis proposes the following research questions to investigate the facilitation of trust with technological adaption and the nature Military involvement has on mental model schemata:

Research Question 1 (RQ1): Facilitating Trust

1) How do the narratives that people tell about automation relate to their trust in technology?

Research Question 2 (RQ2): Military Culture:

2) How does trust in automation differ between military personnel and civilians?

These questions fundamentally underscore this thesis, and the literature reviewed therein. The next section summarises the overall thesis structure.

1.4 THESIS STRUCTURE

This section provides an overview of the structure of the thesis and summarises the chapter contents. These are as follows: Chapter 1 (this chapter) covers the scope, context and objectives of the thesis. It also provides additional background information concerning military social psychology which provides narrative framing. This is followed by Chapter 2 which explores the literature concerning trust, how it can be defined and the key concepts which provide a solid knowledge foundation for later interpretation. Chapter 3 provides a systematic literature review into trust,. automation and human mental models.

Chapter 4 details the methodological choices and theoretical framework to support the data analysis. Subsequently followed by Chapter 5 which uses the techniques and approaches of the previous chapter. This chapter analyses a documentary case study, which assists in grounding the narrative lens of the submariner participant subset; and the exploratory study which aids in bounding the vast concept boundaries.

The two subsequent chapters comprise the analysis of the two distinct participant cohorts. Chapter 6 covers the individual and collective issues of Civilians attitudes towards trust and emerging technology. The concepts and expressions within these verbal protocols were subsequently used to explore responses in the Military cohort in the following chapter. Chapter 7, therefore, covers the individual and collective issues of Military personnel towards trust and emerging technology.

Chapter 8 summarises the output of the analysis and details the narrative personas extracted from the data. In the final chapter, Chapter 9, the thesis is concluded. This chapter elaborates on the implications of the research outcomes, addresses the research questions and provides recommendations for improvements, limitations incurred and suggestions for future work.

In addition to the chapter outlines, the thesis structure additionally includes the primary appendices. Appendix A covers the weighted thematic outcomes for the core studies during the project. This information can be cross-referenced in the main body of the text. Appendix B presents security and ethical procedure information used when collecting qualitative user data. Finally, Appendix C presents the detailed sources from the Systematic Literature Review in a tabulated format.

2 TRUST

Parts of this chapter have been published in:

[1] "Trust in Automation: How this is shaped by the human operator and the underwater domain" paper presented at UDT 2019: Undersea Defence Technology Annual Conference, Stockholm, Sweden, 13-15 May 2019.

The sections 'Related Work' on Page 1-3 of [1] have been used to prepare this chapter.

This chapter presents a literature review on trust as a concept and its theoretical underpinnings. Trust definitions vary and are a multi-disciplinary concept and this chapter seeks to outline the background and pedagogy going forward in this thesis.

Trust research has increased in the literature across domains (Lyon, 2015), especially in the last decade. Trust is a fundamentally human-centric social phenomenon with cross-cultural determined complexities, and thus this section outlines the foundation knowledge and the broad array which trust research scopes. Facilitation is a prominent focus in high-criticality organisations, such as the military increasing levels of automation and use of emerging technology (Lyons, et al., 2018). Therefore, attitudes towards Human-Automation Interaction (HAI) are also touched upon in this chapter.

2.1 TRUST AS A MULTIDISCIPLINARY CONSTRUCT

Trust literature spans a diverse breadth across many domains, from social sciences to business management, ethics, neuroscience and more (Rousseau, et al., 1998; Robbins, 2016; Drnec, et al., 2016). The conceptualisations of trust differ with diverse methodological approaches, elements and characteristics which span the domains and decades. How trust is viewed as a construct takes many forms, whether it is a behavioural intention, a social cognition or an internalised value (Colquitt, et al., 2007). Colquitt *et al* (2007) discusses the nuances of trust as its synonymous characteristics of trustworthiness and trust propensity. This, as well as dependency, control and power between social actors is expanded upon in the following section (2.2). However, the trust discourse in the social

sciences is diverse as there are many interlacing theories in sociology, social psychology and psychology. Section 2.3 examines some of the trust discourse around personality traits, childhood development and our social worldviews. The rise of behavioural and cognitive psychology in the mid-20th century provoked a significant rise in experimental design, statistical analysis and theoretical modelling to the field of trust research. Section 2.4 provides a brief introduction to some of the more prominent frameworks and psychometrics. This section explores some of the highly criticised areas of the literature regarding defining trust as a construct; as some researchers consider trust synonymous with reliability, risk behaviour, cooperation and decision making.

The final section (2.5) in this chapter explores the impact of technology – whether as mediator or participatory actant – and the rise of trust perspectives in sociotechnical systems. Sociotechnical system ergonomics is important in exploring trust as there is significant crossover of objective psychometrics and social theories. Triangulation of these can be used to look at the micro and macro lens of social psychology when exploring the interplay between and within groups. In section 2.5, the human factors within sociotechnical system ergonomics are discussed and their specific relation towards trust in automation and human-machine interface arguments.

2.2 TRUST, TRUSTWORTHINESS AND TRUST PROPENSITY

As mentioned in the previous section, the attempt to define trust is complex and often ill-defined. To properly address the terminology of trust, it is important to distinguish some of the most common expressions. The Integrative Model of trust (Mayer, et al., 1995; Mayer, et al., 2007) attempted to define trust by delineating the actions between trustor and trustee; and the anticipation of positive expectations which is depicted in many other concepts of trust (Barber, 1983). The former seeks to understand how parties process information regarding perception of others and the inherent risk behaviours therein. This can also influence how vulnerable individuals are willing to be (Rousseau, et al., 1998). In the last 20 years, research into trust has examined the role of emotion, attachment and integrity (furthered explore in the following sections: 2.3 and 2.4).

A meta-analysis of trust antecedents by Colquitt *et al* (2007) defines the terms as follows: 1) Trust is being vulnerable to someone despite untrustworthy actions or communication; 2) Trustworthiness is defined by characteristics and behaviours of one person that inspires positive expectations in another; 3) Trust propensity is based on actual or perceived reliability on others. Though this meta-analysis explores H-H dynamics, it has since been used and applied to HCI (Kulms & Kopp, 2018). Exploring human-system interaction from a social cognition aspect is a key theme of inquest in this thesis.

To explore this further, it is fundamental to understand that trust is transformational. For example, where limited interaction or knowledge is obtained about a trustor, initial trust is oft based on the trustee's trust propensity. However, following interaction, a trustee will form knowledge-based trust and trustworthiness is gauged through this schema adaption. Although trust is human-centric, trustworthiness can also be applied to systems and non-human actants. Hall and McQuay (2010, p. 19) state that:

> *"Trustworthiness is assurance that a system will perform as expected despite environmental disruptions, human and operator error, hostile attacks, and design and implementation errors"*

They also define trust as a cognitive mediator in response to environmental stimuli and that observable behaviour can be used as a predictive variable. However, as discussed at length in Chapter 3 (Section 3.3.3 Human System Interactions); reliability as a measure of trust can be inappropriate. As Lee and See (2004, p. 53) suggest, *"In the context of trust and reliance, trust is an attitude and reliance is a behaviour".* The distinction between trust and reliance has implications in trust research and beyond as individuals' oft show a proclivity to distrust reliable automation (Dzindolet, et al., 2003).

The nuances and variances of the definition of trust and its components are diverse and cannot all be outlined in one thesis. However, those with social and cultural implications are of specific interest in this research. An instance of this would be strategic or moralistic trust. This is interrelated to cognitive decision making as it is contingent on personal experience of others trustworthiness, information and incentives (Hardin, 2002). The formation of mental models and social schemata to form trust heuristics

and relationship forming between actants is of interest in exploring macroergonomics. Furthermore, how distrust and their associated behavioural responses of social withdrawal and revenge can have analogous implications in human-system teaming (Lewicki, et al., 2006).

A comprehensive exploration into the existing variety of trust definitions across disciplines can be found in Robbins (2014, pp. 18-20). They simplify the range of delineations of trust as the following; 1) *how* – the psycho-social microfoundations of external (strategic) and internal (moralistic) trust*;* 2) *whom* – is the trustee specific (internal or within group) or general (external or outside of group)*;* 3) *what* – simplex (concerns a specific task or matter) or social trust (concerns many tasks or all matters).

As seen in this section, the psychological and sociological scope of trust is varied and complex. The following two sections aim to explore these domains further.

2.3 INTRODUCTION TO TRUST AND SOCIAL SCIENCES

Trust research in the social sciences gained traction in the mid-20th century with Neo-Freudian, Erik Erikson (Erikson, 1950) and psycho-social development. He theorised that personality is cultivated through experiences, especially those in childhood development. Trust acquisition, it was suggested, occurred during the ages of 0-2 years, when the individual receives consistent, predictable and reliable care and that these schemata are pervasive in building relationships going forward. Furthermore, that inconsistencies and unpredictability propagate mistrust, suspicion and anxieties. Ainsworth and Bell (1970) reinforced trust development in childhood experiences through their research into attachment styles.

Personality trait research followed and further attempted to conceptualise intrinsic indicators of trusting behaviour in DeNeve and Cooper's (1998) meta-analyses of personality traits and subjective wellbeing. Their analysis of literature explored the Five Factor Model (McCrae & John, 1992), a personality psychopathology model, which posits that five major traits explain differences in personality and behaviour. These traits are Openness to Experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism (OCEAN). They suggest that, although trust evolves over time with experience, that

intrinsic traits influence trust formation (such as agreeableness and conscientiousness). Higher dispositions of agreeableness have shown an interaction with cognitive ability to predict trust (Lyons, et al., 2018; Lyons & Stokes, 2012). These methodologies have also been used to predict trust inclinations human-computer and human- system interaction (Dillon & Watson, 1996). In addition to internal factors, researchers also acknowledge that cultural dimensions interact with trust attitudes and trust adjacent traits. For example, Chien *et al.* (2016) examined OCEAN qualities and trust attitudes with different cultures against technology. Stability of social hierarchy was a factor, in that an individual from Honour culture (e.g. Turkey) in general had low initial trust to others, whereas a member of Dignity culture (U.S.A) showed higher general trust effects. Social influence and the notion of trust are being increasingly adopted into predicting behaviour by others. This is especially important when it comes to government, institutions or machines regarding authority, accountability and agency effect (or agentic shift) (Milgram, 1963).

Sociologists have transformed the work of Milgram towards dispositional or situational factors influencing trust and obedience behaviours, such that social influence plays a major role in conformity and personal agency. Rotter (1966), although later criticised for reductionist and deterministic evaluation, explored the individual's locus of control on trust. For example, an individual's locus of control refers to how much control an individual feels they have over their own behaviour (internal loci believe control is due to their own actions and behaviours). Research has broadened to social influence factors and their effect on compliance or resistance. In their paper on interpersonal trust, they explored social organisation, and the willingness to trust others in a social group or unit (Rotter, 1967, p. 664; Rotter, 1980). They define interpersonal trust as *"…the generalised expectation that the verbal statements of others can be relied upon".* The construction of the Interpersonal Trust Scale explored trust inclination between a variety of social objects, and how in- or out group actants impact compliance. Kramer and Tyler (1996) further explored trust as a calculative social orientation influenced by risk and how social exchange and power differentials can impact trust behaviours. Social Identity Theory perspectives further investigates how interpersonal position influences trusting behaviour, whereby categorisation of identity and self is inter-related with group membership and stereotypes (Tajfel, 1979),

and that trust can be manipulated through in-group favouritism and group distinctiveness (Tanis & Postmes, 2005). Tanis and Postmes (2005) report that individuating information reduces perceived ambiguity, and that physical presence cues provide intimacy and immediacy. These findings are important when discussing the rise of technology-mediated communication as well as human-system teaming interfaces. The facilitation of trust through group membership where interfaces or interaction barriers reinforce otherness, may inhibit formation of a shared social identity and thus trusting behaviour.

Shared social identity is a key theme in trust research in the social sciences. Barbara Misztal, in their book Trust in Modern Society, set out to combine notions of trust together into a unified typology (Misztal, 2013). Through extensive discussion, they explore social relationships and their interactions to discern trust as related to social order through stable, cohesive and collaborative order. These are roughly defined as trust as habitus (habit and collective memory); trust as passion (values based on interpersonal relationship interactions); and trust as policy (coping with othering and personal autonomy).

In the early conceptions of trust, Luhmann's systems theory sought to use phenomenological meaning making of reducing complexity in social interactions (Luhmann, 1979). Through this, he theorised that trust was based in the reduction of uncertainty of future events via lived experiences, social heuristics and schemata. Furthermore, trust is inter-related with power dynamics through the adaption of complex social environments and interactions to increase opportunities. Morgner (2018) updates Luhmann's research with the shift of structures in society and subsequent role of trust by addressing gaps and other theoretical considerations for contemporary society.

The role of trust and systems theory is pertinent to Human-Automation Interaction research as *"Trust is only involved when the trusting expectation makes a difference to a decision"* (Luhmann, 1979, p. 24). If trust is based off knowledge of the actions of others, transparency of technology is vital when uncertainty is bred in the unknowable actions of the system. Trust facilitation therefore must include mechanisms for managing the uncertainty about autonomous entities and the information

communication pathways. Furthermore, trust can play a substantial role in decentralized systems (Tang, et al., 2012). The increased levels of distribution in sociotechnical systems, and the role of trust are discussed further in section 2.5. In addition, the role of distributed and decentralised human-system teaming in the military are included in section 1.2.1.

2.4 TRUST FRAMEWORKS AND PSYCHOMETRICS

As has been discussed in the previous sections, the conceptualisation of trust is often heavily associated with risk and reliability (Coleman, 1990; Molm, 2007; Yamagishi, 2001). This section seeks to identify their role in trust frameworks and psychometrics, and if they are appropriate antecedents when researching trust and related themes.

One of the key frameworks in the literature is that of Information Theory (Gerck, 2002). Gerck pursued the information communication of trust in heterogeneous environments, from human to machine, machine to machine, and machine to human. They observed control and power distributions in both open and closed loop systems and how trust is related to reliance on information. They posit that same channel information is unreliable and truthfulness in the communication cannot be appropriately obtained when transferred using that same channel. For example, they suggest that *"trust me"* is an empty affirmation as a self-affirmation cannot communicate trust. This situation applies to system-teaming and human-human teams. Whereby, information can be benign or malignant and, as Sztompka suggests*: "Trust is a bet about the future contingent actions of others"* (1999, p. 25).

Other frameworks explore how impressions towards social actants can rely on prior impressions influencing subsequent interactions. Social psychologist, Solomon Asch, proposed the Models of Social Impression Formation (1946) whereby inferences of the perceptions and attributes of others are built upon information available. For example, individuals may employ incomplete pieces of information and cognitive shortcuts using representative heuristics (Tversky & Kahneman, 1974). An extension of this is the Continuum Model of Impression Formation (Fisk & Neuberg, 1990). They theorise that impressions, and thus trusting behaviour in response, are formed upon social categorisation to existing

schemata based on stereotypes. These can then be adapted and adjusted with personal relevance and resource management to acquire more information to confirm or recategorize these mental models. Organisation research into trust has additionally explored the use of categories and schemas to create prototypes of people (such as stereotypes) in cooperative teams (Brewer, 2008) and to evaluate the intentions of unfamiliar others (Kramer, 1999; Messick & Kramer, 2001).

Other researchers have explored the intersectionality of sociological and psychological conceptualisations of trust by studying the behavioural expressions of psychometric techniques in the intersubjective and multidimensional social reality. Lewis and Weigert (1985) enquire into Barber and Luhmann's (Barber, 1983; Luhmann, 1979) seminal research and how cognitive, behavioural and situational manifestations of trust are not temporally or spatially bound. They argue that trust as a psychological state can be easily confused with other contributary factors such as faith, hope and risk behaviours. Furthermore, that the use of psychometrics is reductionist and deterministic of human behaviour and cognition. One of the main critiques of psychometrics is that survey items are recurrent in social science approaches. Measures of trust are often set as expectations or cooperativeness towards behavioural observations of perceived cognitive processes. However, these are often misinterpreted as closely allied elements such as risk behaviours (such as aversion), trustworthiness (as discussed in section 2.2) and cultural bias (Miller & Mitamura, 2003; Sturgis, 2010). Other criticisms lie in the measurement validity in trust literature. Conceptualising trust and observing internal cognitive processes has a difficult task of identifying causes and consequences. Many concepts rely on how a person trusts, how the association relates to them, and what the individual is trusting the other actant with. Relational trust, as discussed in the section 2.2, is specific to trustworthiness, the relationship between trustor and trustee and the motivations to facilitate trust.

Trust frameworks are generally built on the individual's assessment of the trustworthiness of the trustee and the vulnerability of the position this leaves the trustor. Trusting behaviour is motivated by the positive outcomes, such as the affect to trustee (emotional trust) or by perceived rational reasoning of the trustor that merits trust (cognitive trust), or a combination. Scanzoni (1979) describes trust as an *"...actor's willingness to arrange and repose [their] activities on Other because of confidence that*

Other will provide expected gratifications". But, as Luhmann states, trust is not simply predictive: *"Trust is not a means that can be chosen for particular ends, much less an end/means structure capable of being optimized"* (Luhmann, 1979, p. 88). Transactional based trust, whereby knowledge acquisition builds upon a trustor's predictions and assessments, creates positive trust facilitation through positive and affirming expectations and experiences (Lewicki & Bunker, 1996).

Other researchers of cognitive based trust explore trust decisions and behaviours as a result of processed information based on emotional and relational bonds (McAllister, 1995). As illustrated in the table below (Table 3), Lewis and Weigert (1985, p. 970) reflect that many choices of trust are based on *"good reasons"* or evidence of trustworthiness. Colquitt, *et al* (2007) in their paper on trust, trustworthiness and trust propensity; attempt to define the terminology of what constitutes *"good reasons"*. They vary from study to study under the guise of antecedents, factors, dimensions or conditions. It is recommended that the Ability, Benevolence and Integrity definitions of Mayer *et al* (1995, p. 723) reflect appropriately both cognitive and affect-based trust factors.

Table 3 - Rationality and Emotionality Bases, Types of Trust and Boundary States [Lewis and Weigert (1985, p. 973)]

		Emotionality		
		High	**Low**	**Virtually Absent**
Rationality	**High**	Ideological Trust	Cognitive Trust	Rational Prediction
	Low	Emotional Trust	Mundane, Routine Trust	Probably Anticipation
	Virtually Absent	Faith	Fate	Uncertainty, Panic

More recently, researchers have attempted to model personalised multi-faceted algorithms for trust recommendations specifically for distributed and decentralised mixed actant teams. Quinn *et al* (2009) sought to appropriately model concrete trust concepts such as competency, credibility, honesty,

reputation, and reliability, with the more abstract concepts such as belief, confidence, and faith. Therefore, developing a model of trust which is human-centric and can semantically evaluate the diversity and interlaced meanings of trust.

Others have explored trust in sociotechnical terms through technology acceptance and adoption models and theories. As outlined in the research statement of this thesis, facilitation of trust with system-teammates is imperative for appropriate use and future development. One of the primary goals of technology acceptance and adoption models is to understand the issues which influence user's refusal or reluctance to engage appropriately with systems. As has been discussed, decision making and the underlying cognitive and social processes which influence them, are closely associated with trust dynamics. Taherdoost (2018) compiled a thorough literature review on technology acceptance and adoption models. Refer to Figure 3 for the overview of models covered at length in their manuscript.

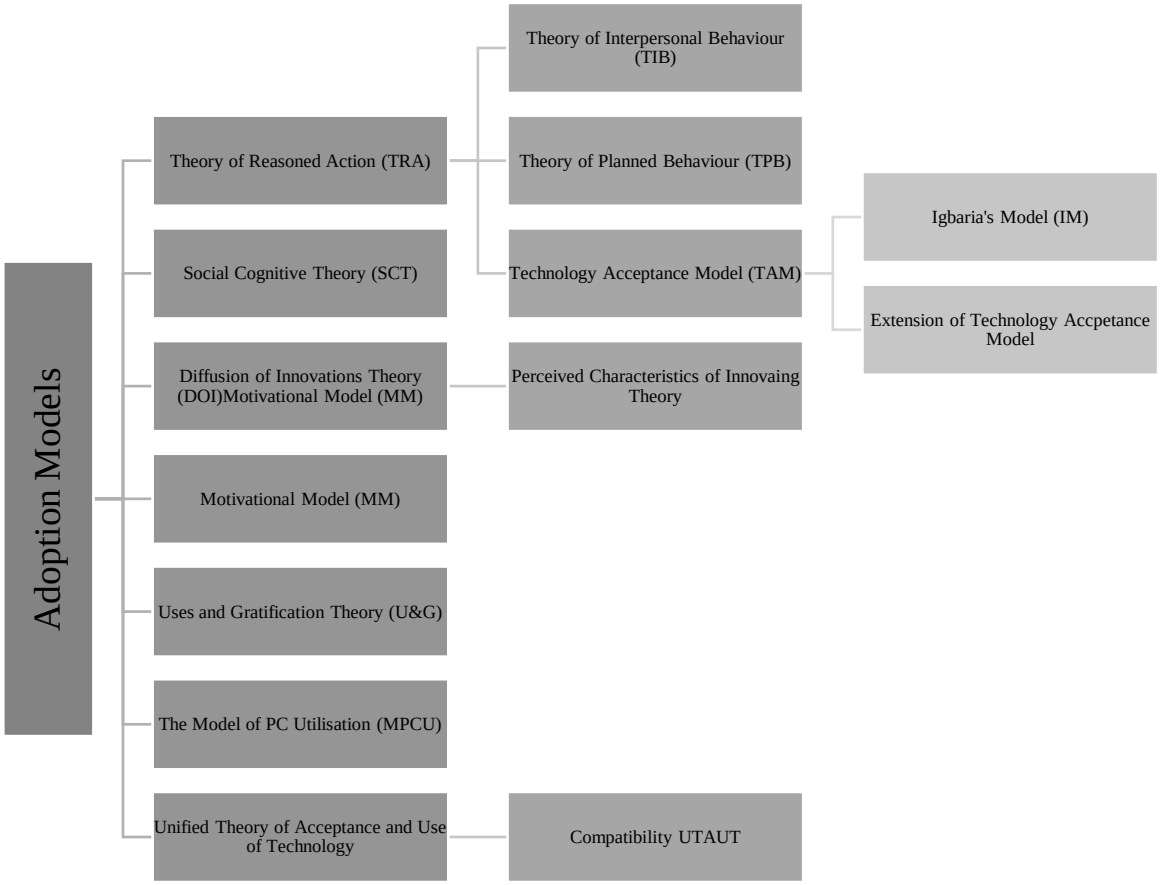

Figure 3 - Overview of Adoption and Acceptance Models [Taherdoost (2018, pg. 962)]

Discussed in the following chapter (Systematic Review of Trust, Automation and Mental Models) the Technology Acceptance Model (Davis, 1989) is used extensively in Trust in Automation research. This is primarily attributed to the factor of trust producing positive impacts on perceived usefulness and explaining user behaviour (Gefen, et al., 2003; Venkatesh & Morris, 2000; Wu, et al., 2011). A comprehensive meta-analysis assessing the effective factors of the Technology Acceptance Model, including trust variables in a diverse range of domains has been conducted by Doulani (2018). However, criticisms of how TAM defines trust, is that behavioural intentions and attitude more appropriately align with many of the conceptual descriptions of cognitive trust. Furthermore, one of the fundamental critiques of psychometrics and self-reporting in general is the accuracy of individual introspection of cognitive processes compared to more holistic accounts within verbal protocol (Nisbett & Wilson, 1977; Fiske & Taylor, 2013; Schoenherr & Hamstra, 2016).

The following section addresses some of the literature unique to sociotechnical systems which attempt to conceptualise and explain inter- and intra- personal relationships, formation of trust influences and other factors prompting cognitive dissonance.

2.5 TRUST IN SOCIOTECHNICAL SYSTEMS

Trust is a social construct but only one element to social reality. In discussing inter-personnel concepts or measures, items such as hope, confidence, risk and power are introduced, as has been mentioned in the above sections. Our social reality, however, is not just between individuals, but a myriad of different levels of relationship from the macro system ergonomics or sociological level; to intra-personal cognitive processes shaped by the world around them. Macroergonomics (or system ergonomics) emerged in the 1960's to aid in viewing the complex interrelated components of organisations, teams and technology (Katz & Kahn, 1966). These 'socio-technical' systems are the basis of many domains within human factors and ergonomics, and thus integral to examining the intersection of human-system teaming. Sociotechnical systems (STS) (Rasmussen, 1997; Carayon, 2006) is an approach to complex organisational structures that recognise interaction between human-job, human-machine, and human-

software interfaces (Hendrick & Kleiner, 2001). The term also refers to the interaction between society's complex infrastructures and human behaviour.

Sociotechnical Systems models are the junction of organisational design, human interaction and technologies (Waterson, et al., 2015). Although trust is not controllable in teams, elements of trust can be mediated somewhat by reliability elements increasing confidence in a system. In technological domains, trust and control are closely associated with reliability, and discussion that confidence is the result of trust and control, which echoes power distribution theories such as Luhmann and Baier (Baier, 1986; Luhmann, 1979). Cofta (2007, p. 12) suggests that trust in technology and distributed systems can be facilitated through security and confidence elements in computing systems. For example, reliability in positive expectations and transparency in the system could foster confidence in perceived and actual usability. Furthermore, Cofta proposed that there is an expectation that trust in a technical structure reflects social relationships. One of the main elements of sociotechnical system interactions, and subsequently crucial discourse for this thesis, is the debate on how social relationships are formed with non-human actants. Trust is attributed to interactions and relationships between social actors – whether between individuals or groups in social systems. The following theories discuss actants or non-human systems as social actors.

2.5.1 Actor Network Theory (ANT)

Actor Network Theory (Latour, 2005) is a social constructivism-based theory which attempts to describe factors and interactions between actors and actants in the fluid and fluctuating relationships between humans, objects, ideas and processes. He relays the importance of social relations with non-human actants when discussing technological infrastructure and power hierarchy (Latour, 1991, p. 110):

> *"We know how to describe human relations, we know how to describe mechanisms, we often try to alternate between context and context to talk about the influence of technology or vice-versa, but we are not yet expert at weaving together the two resources into an integrated whole."*

Latour (2013) further suggests that technology is "*... neither a human nor a non-human phenomenon, but a hybrid of both".* Technology is an integral part of our social reality, and we are fundamentally dependant on technology across domains and architecture. Despite technology designed, constructed and shaped by humans, technology consequently shapes human interactions and are indivisibly connected in a heterogenous complex system.

Actor Network Theory is important for exploring trust dynamics between human actors and artificial agents. Trust, as proposed in this chapter, is socially constructed or at least embedded in cognitive responses. By engaging with automated actants in a holistic systems approach, humanities research can be utilised to explore human-automation interaction where appropriate. Furthermore, human-interaction is rooted in human social responses and mechanisms, therefore, exploring trust facilitation must include social psychology.

Within ANT ontologies, Nimmo (2011) suggests there is no domain that exclusively consists of interactions between human actors, but these relations are mediated, transformed or enabled by non-human actors. Examples of these can be physical objects, technologies, animals or ecosystems. ANT is commonly associated with STS models as it attempts to map material and semiotic concepts with their associated relations. Figure 4 provides a visual representation to illustrate how actants and concepts are interrelated with each other when designing a graphical user interface. This material-semiotic method, or more specifically a theory, has similarities to STS modelling such as those used in macroergonomics. However, ANT is more paradigmatically linked with sociology as it lacks empirical tools or techniques, and employs inquiry approaches such as symbolic interactionalism (Star, 1992) and grounded theory (Myers, 1997). Regarding trust, and relationship forming, understanding the underlying cognitive processes in social psychology bond formation is crucial for interpreting social responses to communication technologies. The next section explores this theme further.

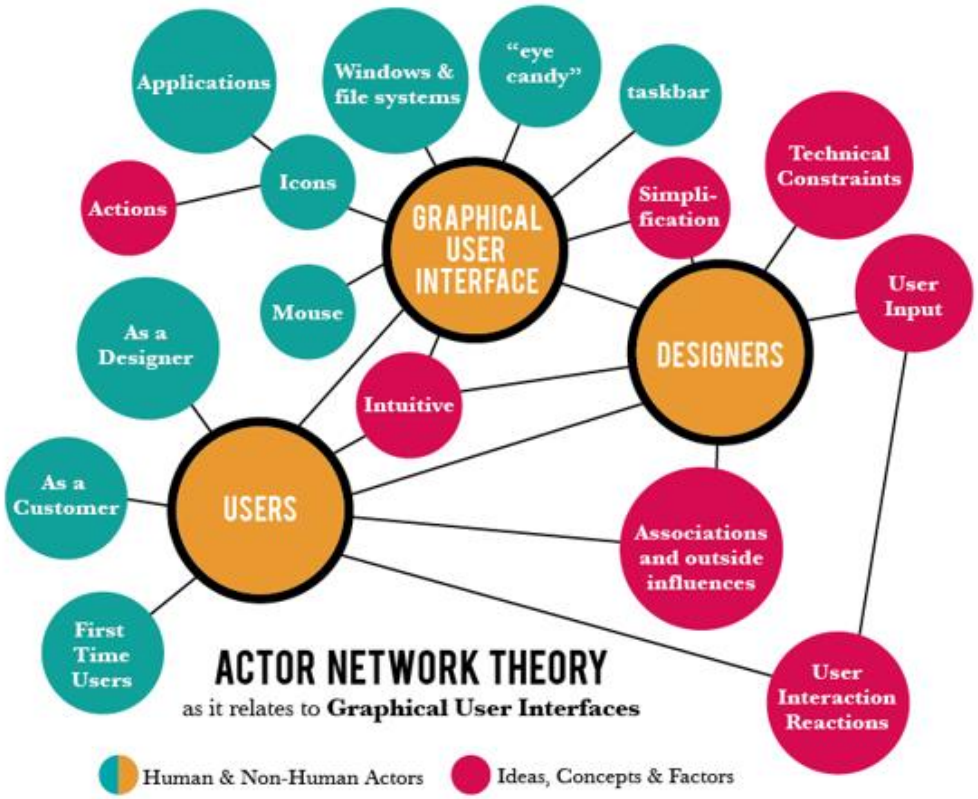

Figure 4 - Example Diagram of Actor Network Theory as it relates to Graphical User Interfaces (Webb, 2013)[Creative Commons]

2.5.2 The Media Equation

The Media Equation explores how people fundamentally react with computers, emerging technology and system actants in sociotechnical systems and everyday life (Reeves & Nass, 1996). They propose that human actors have expectations of non-human social actors to obey socially constructed rules borne of interpersonal interactions. Empirical studies observing the parasocial relationship towards interactions with technology have observed non-human actants perceived as social actors regardless of conscious or unconscious expectations of trust (Krämer, et al., 2012). Similar to the previous section, the use of this theory relates to trust research through exploration of trust themes from a socially constructed lens. If individuals utilise socially constructed mechanisms to trust artificial agents, then human-human dynamics are key for appropriately facilitating human-system teaming.

An expanding area of research which specifically examines the communicative interactions towards artificial entities is that of social robotics. Through using elements of human-human communication,

and identifying valuable criterions, meaningful and facilitative interactions can be made. For example, assumptions of knowledge, information communication and category membership are all concepts related to trust facilitation that are proposed for human-agent interaction. Technology can already elicit social cues, such as human-like language or fulfilment of social roles. Humanlike responses like these can activate existing social scripts, thus can positively effect social cues and the subsequent experiences can appear more humanlike (Szczuka, et al., 2019; Royakkers & van Est, 2015; Savela, et al., 2018).

Conversely, research such as Broadbent *et al* (2011) and others have observed visceral, unconscious unease at social robotics. In their experiment, they observed blood pressure increase in patients with pill-dispensing robots which had an image of a human face attached to it. In relation to the Media Equation, Nass and Moon (2000) submit that artificiality cues may interfere with social cues and subsequent social treatment may be negatively affected by these interactions. A crucial issue with the creation of human-like agents, which mimic humans, is the Moravec Paradox:

> *"It's comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult to impossible to give them the skills of a one year old when it comes to perception and mobility" (Moravec, 1988, p. 15)*

The difficulty of impersonating human behaviour and actions is most keenly observed in the study of anthropomorphism. As indicated in Chapter 1 (Introduction), the ambition to implement anthropomorphic system-team agents (MoD U.K, 2014) can be problematic. There are numerous underlying issues with human-like machines and technology in many domains and at different psycho-social levels (Bar-Cohen & Hanson, 2009). However, one of the more confounding issues with anthropomorphism, is with highly humanlike robotics. Strait *et al* (2015) employed novel experimental paradigms of traditional subjective ratings, behavioural measures and physiological responses. Their results identified highly humanlike robots as uncanny and exhibited greater avoidance than encounters with less humanlike and human agents.

The concerns stemming from the feeling of 'uncanny' are displayed keenly in social robotics as they are electrotechnical but share human qualities, whether through physical design or artificial intelligence.

Freud (1919) explored the uncanny in his psychoanalytic essays and theorised that the feeling of discomfort comes from familiar but repressed emotions associated with the social actor. More recently, and in the field of robotics, Masahiro Mori's (2012) influential essay, "The Uncanny Valley", explored the hypothesis that there is a point in which humanness in a non-human agent significantly affects negative emotional responses. They theorise the unnerved reactions to robotics are bound to expressions of self, and agents which exist close to the category boundary between human and non-human can be seen as a threat to human identity. Category boundaries describe agents, actors and actants that exist in the boundaries of perceived binary spectra. Furthermore, cognitive conflict has been empirically universally observed across morph spectra (Wiese & Weis, 2020). Research has shown this is clearest with human faces, perceived impression of humanness is at its most fragile or sensitive with facial recognition associated category boundaries (Looser & Wheatley, 2010). Social cues and responding social scripts are elicited intensely by facial aesthetics (Kappas, 1997), and therefore undergo the most scrutiny (Watt, et al., 2017). The central role of facial ambiguity in specific fears (e.g. mannequins) are explored in the essays of camera portraiture by Sobieszek (1999, p. 86). The concept of the 'ghost in the shell' explores human identity in portraiture and the eerie response to facial ambiguity:

> *"[1] Treating the face either as a blank somatic surface expressive of absolutely nothing yet infinitely suggestive; [2] Or is a fluid matrix of constantly shifting identities at once true and false, assumed and genuine, feigned and imagined."*

Ultimately, facial aesthetics are a crucial element of scrutiny in social interactions and ambiguity can create aversive responses in non-human actants (Tinwell, et al., 2013). When facilitating trust in systems interaction, negative stimuli associated with the feeling of uncanny (such as threat avoidance) should be mediated as appropriately as possible.

2.6 SUMMARY

Trust is an emergent property of social reality, embedded in and personally experienced by people in their everyday life (Robbins, 2014). This chapter provides a foundation of the depth and breadth of the

complex concept that is trust. Herein also identifies the interdisciplinary diversity of knowledge on this subject, as well as the theoretical and epistemological avenues they are examined from.

Facilitating trust in automation and emerging technology, this segment emphasises social presence as an influencing factor and that 'trust needs touch' (Handy, 1995, p. 46), physically or metaphysically. By that, technology mediated environments (Hess, et al., 2009) benefit from building trust through social presence, as identifiability increases vulnerability which can generate trusting beliefs or behaviours by lowering risk or perceived harm avoidance. There is significant literature on human agency, anonymity and antisocial behaviour in response to trust and power distribution (Zimbardo, 1969). This chapter also stresses the importance of social cues and scripts, both from a human mental model standpoint and that of artificial human actants in system teaming. Furthermore, as social robotics and anthropomorphism become more pervasive; the themes of the uncanny, social responses in communication technology and system-trust confidence have potentially unidentified psychosocial consequences.

This chapter highlights the limited human factors research in autonomous systems actively repairing trust, human-automation trust calibration and incorporating or adapting social science findings to human machine teaming (de Visser, et al., 2018). That literature in organisational managements and macroergonomics can provide principles which may aid in administering trust repair and improve facilitation. Researchers have shown assumed performance improvements are often not realised due to inappropriate reliance on automation which could be attributed to basic understandings, or misunderstandings, of trust (Stokes, et al., 2010). For human-system teaming, there is a perception of reciprocity of trust through the media equation (Turkle, 2010, p. 3) . By addressing cognitive based trust, that is derived from reliability, dependability and competence (McAllister, 1995) there is an option to superficially manipulate perceived trust propensity and trustworthiness expectation in HAI.

This chapter provides a robust foundation to address Research Questions 1 (**RQ1**): *How do the narratives that people tell about automation relate to their trust in technology?*

The following diagram (Figure 5) provides a summary of the key concepts raised in this chapter and presents theories which will be key to conceptualising participant expressions of trust.

Figure 5 - Diagram of Key Trust Concepts in the Literature

The next chapter provides a systematic review of existing literature concerning human-automation interaction and the use of mental models to explore gaps in trust calibration and facilitation.

3 SYSTEMATIC REVIEW OF TRUST, AUTOMATION AND MENTAL MODELS

This chapter have been published in: (Field, 2018)

[2] "Review of Mental Models as a Method of Trust Facilitation for Human System Interaction," Technical Paper Presented at 2018 Defence and Security Doctoral Symposium (DSDS18), Swindon, England. 13-15 October 2018.

This chapter presents a systematic review of the research into trust, automation and mental models. The dimension of trust, as discussed in the previous chapter, has a crucial impact on developing intra- and interpersonal interactions. It is also a cultural dimension which can be shaped by the social sphere individuals inhabit. The idiosyncrasies of military culture are of specific interest in this review but are not a limiting factor, as a wider context is also appreciated for generalisability.

3.1 INTRODUCTION

Automation is a critical future trend military domain (MoD U.K, 2014) and civilian industries. Increased levels of automation can provide greater system capability and assistance in decision making, data selection or process management, and performing tasks that are otherwise "*dangerous, time-consuming, or outside of human desire or capability*" (Parasuraman, et al., 2000; Lee & See, 2004). As automated systems advance, the role of the operator will change in nature from active control (i.e. human-in-the-loop) to a supervisory management position (i.e. 'Human-on-the-loop') (Atkinson, et al., 2012). Understanding the difference between levels of automation in a predominantly human-oriented and safety- or mission-critical system is an important consideration as the utilisation of more advanced, sophisticated technologies becomes ubiquitous in both military and civilian domains. However, trust in automation (TIA) remains a diverse and broad research field as technology advances faster than literature can appropriately keep up with (Siemens, 2008).

There are three main views throughout the literature on human factors and human system interaction; these are human-in-the-loop, ‘on-the-loop’ and ‘out-of-the-loop’ (Wogalter, 2006). The latter is problematic for human-automation interaction (HAI) research. It suggests the system may act independently of a human operator, or system integration has failed and operators are unsure on what the status of the system actually is, thus divorced from the interaction. Integration currently strives to achieve in-the-loop interaction, with human and system integrated well, and with both ‘actors’ aware of each other’s influences. On-the-loop refers is the ideal for HAI system integration activities and development programmes, with human actors removed from the overall system functionality and allocated tasks that are predominantly focused on surveillance and maintenance of the autonomous actions of the system. Factors that influence trust are imperative for appropriate use and reliance on the system and subsequent decision making. In the context of this review Human Mental Models (HMM) are frameworks (schema/schemata) that individuals cognitively construct based on prior experience and knowledge to support their expectations and interpretations of their environment.

3.2 REVIEW METHODOLOGY

A systematic review methodology was adopted to explore new and emerging research in the field of human system automation and mental models. This protocol was used to focus on the gaps in the literature as opposed to answering a narrow research question which systematic review methodology traditionally utilises (Arksey & O'Malley, 2005). However, systematic review procedure was applied where appropriate, such as inclusion criteria, quality filtering and data extraction where appropriate.

3.2.1 Systematic Review Framework

The search criteria for the databases included (Google Scholar, IEEE Xplore, Science Direct and Wiley Online Library databases) include keywords reflective of the broad scope of HAI research. Related articles and bibliographic citations were screened for applicable sources. To ensure relevance, a date range restriction of articles since 2005 was imposed. Figure 6 illustrates the search strategy adopted.

All articles were qualitatively assessed for relevancy. Such as human mental models as the focus of discussion, and/or experimentation with automation, systems interaction or computer interaction as secondary variables. Journals within critical workplaces, such as the military or healthcare were prioritised (with relevant age groups (working age, 18-60)). Clinical studies were excluded, as they lack ecological validity/environmental generalisability. Other exclusion criteria were unpublished or non-English literature, or where literature was not available in full text form.

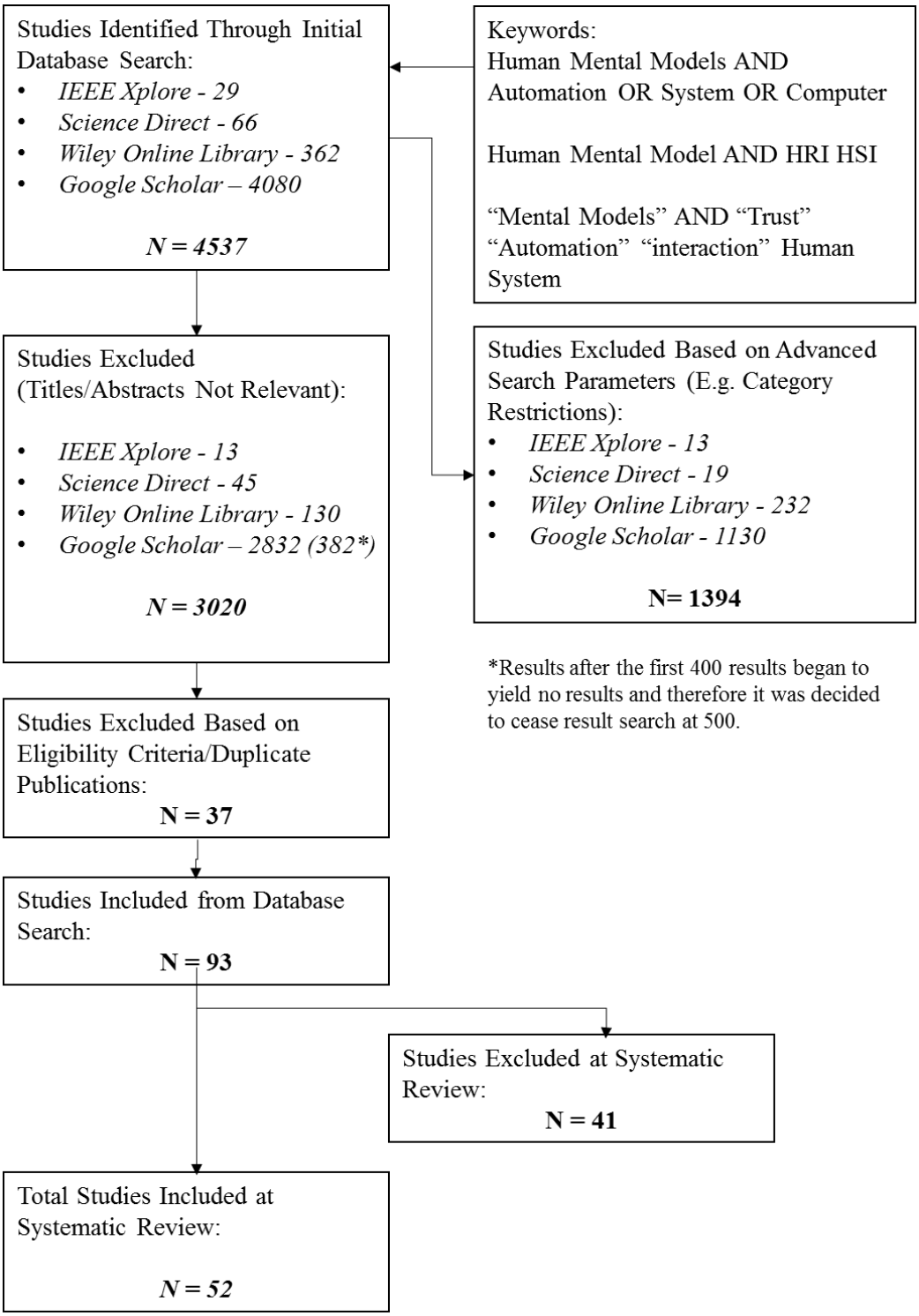

Figure 6 - Eligibility Screening Identification

Study design and methodological information were extracted where possible. Primarily qualitative research materials that provided scope and context include discussion pieces, reviews and meta-

analyses. Elements such as location, participant demographics, outcome measures and results were sought where possible to provide appropriate context and generalisability to the sources reviewed. Outcomes were clustered according to grouping criteria using a modified version of the Three Factor Human-Robot Trust Model outlined in Hancock *et al* (2011). Although this model focusses on Human-Robot[1] issues, these themes closely align to general system behaviour within sociotechnical domains

Quality assurance of the literature review adhered to existing systematic review protocol such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocols (Moher & Liberati, 2009) with regard to evaluation guidelines and inclusion criteria[2] and Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines (Stroup, et al., 2000).

The assessed literature are categorised into the relevant thematic categories. These are as follows: Human Related, subcategorised by Ability Based and Characteristics; System Related, subcategorised by Performance Based and Attribute Based; and Environmental, subcategorised by Team Collaboration and Tasking. Literature within these tables have been refined by author, alphabetically.

[1] 'robot' is under the umbrella of system in this review

[2] The risk of bias could not be conclusively assessed across all included sources in the recommended methodology (Higgins & Green, 2008). High risk sources were excluded, and medium risk literature was moderated.

3.3 RESULTS

3.3.1 Search Results

The eligibility screening process in Figure 6 provided keyword and terminology parameters for each database search. The resulting data pool was relatively large for the field of research (*N=4537*) (Hoff & Bashir, 2015). A portion of sources were excluded due to irrelevant titles or abstracts and a third excluded through search parameters (e.g. category restrictions available on electronic databases (e.g. full text versions/English only)). Several duplicates and reproductions of identical studies to multiple sources (conference proceedings and subsequent journal article omitted). Of the remaining 93 eligible sources, a further 41 were excluded through review of content, leaving a total 52 sources reviewed in this document.

Table 4 - Summary of Systematic Literature Results

Study: Author(s), Country	Subdomain	Exp[3]	Obv	Lit/D	Actors
Human Related					
Ability Based: 12					
Aydoğan, Sharpanskykh, & Lo (2014) *Netherlands*	Aviation	–	✓	✓	–
Birkmeier, Korn, & Flemisch (2011, October) *Germany*	Aviation	–	✓	✓	ATCo[4] (n=8)
Chua, Storey, & Chiang (2012) *North America*	Civilian	–	✓	–	High Skilled Engineer (n=14)
Fallon, Murphy, Zimmerman, & Mueller (2010, May) *North America*	Civilian/University	–	–	✓	–
Hawkley, Mares, & Giammanco (2005) *North America*	Military	–	–	✓	–
Hoffman & Woods (2011) *North America*	Civilian	–	–	✓	–
Lim, Dey, & Avrahami (2009, April) *North America*	Civilian	–	✓	–	51% female; 49% male; Mean Age 29.8 (n=55)
Neerincx, et al (2008, January) *Netherlands*	Naval	✓	–	–	Warfare Officers/Assistants (n=8)
Piccinini, Rodrigues, Leitão, & Simões (2015) *Portugal*	Automotive	✓	–	–	13 ACC[5] users (age 42.2 (SD=9.9), 13 non-ACC users (age 26.7 (SD=9.9). (n=26)
RTO/NATO (2007) *Multinational*	Military	–	–	✓	–

[3] Exp = Experimental study design / Obv = Observational study design / Lit/D = Literature review, discussion piece and other review-based sources
[4] ATCo – Air Traffic Controller
[5] ACC – Automated Cruise Control

Study: Author(s), Country	Subdomain	Exp[3]	Obv	Lit/D	Actors
Wilkison, Fisk, & Rogers (2007, October) *North America*	Civilian/University	–	✓	–	Undergraduate participants (58% F;42% M; Ages18-30)(n=12)
Zhang, Kaber, & Hsiang (2010) *North America*	Automotive	✓	–	–	Male Participants, Mean Age: 25 (SD=2.4) (n=12)
Characteristics: 19					
Arkin, Ulam, & Wagner (2012) *North America*	Military	–	–	✓	–
Beggiato & Krems (2013) *Germany*	Automotive	✓	–	–	51% Female;49% Male; Mean Age: 24 (n=51)
Beggiato, Pereira, Petzoldt, & Krems (2015) *Germany*	Automotive	✓	–	–	47%: Female; 53% Male; Mean Age: 28 (SD=1.82); (n=15)
Bruemmer, Gertman, & Nielsen (2007) *North America*	Automotive (USAR)	–	✓	–	(n=153)
Bunt, Lount, & Lauzon (2012, February) *Canada*	Civilian/University	–	✓	–	Exp1 (n=21); Exp2: (n=14)
Dehais, Causse, Vachon, & Tremblay (2012) *France*	Military/Automotive	✓	–	–	Mean age: 27.84 (SD = 6.53) (n=13)
Groom & Nass (2007) *North America*	Military/Civilian	–	–	✓	–
Hancock, et al. (2011) *North America*	Military/Civilian	–	–	✓	–
Hoff & Bashir (2015) *North America*	Military/Civilian	–	–	✓	–
Hoffman, Johnson, Bradshaw, & Underbrink (2013) *North America*	Civilian	–	–	✓	–
Lee, Lau, Kiesler, & Chiu (2005, April) *North America; Hong Kong*	Civilian	✓	–	–	Exp 1: n=60); Exp 2: (n=48)
Nachtwei (2011) *Germany*	Civilian	–	–	✓	–

Study: Author(s), Country	Subdomain	Exp[3]	Obv	Lit/D	Actors
Nothdurft, Lang, Klepsch, & Minker (2013, April) *Germany*	Civilian	✓	–	–	(n=48)
Oleson, Billings, Kocsis, Chen, & Hancock (2011, February) *North America*	Military/Civilian	–	–	✓	–
Olson, Fisk, & Rogers (2009, October) *North America*	Civilian	✓	–	–	Older population (ages between 60-80) (n=19)
Sanders, Oleson, Billings, Chen, & Hancock (2011, September) *North America*	Military	–	–	✓	–
Schaefer, et al. (2014) *North America*	Military	–	–	✓	–
Talone, Phillips, Ososky, & Jentsch (2015, September) *North America*	Military	✓	–	–	(n=100)
System Related					
Performance Based: 6					
Barg-Walkow (2013) *North America*	Civilian	✓	–	–	N= 60 (38% Female; 62% Male; Mean Age: 19.8 (0.21 SD)
Cassidy (2009) *North America*	Military/Naval	✓	–	–	26% Female;74% Male (n=42)
Dawson, Crawford, Dillon, & Anderson (2015, May) *North America*	Military/Civilian	✓	–	–	(n=40)
Mosier, et al. (2013) *North America*	Aviation	–	–	✓	–
Schaefer, Evans III, & Hill (2015) *North America*	Military	–	–	✓	–
Westin, Borst, & Hilburn (2016) *Netherlands*	Aviation/Civilian	–	–	✓	–
Attribute Based: 4					
Andersson (2010) *Sweden*	Civilian	–	✓	–	–
Sheridan & Nadler (2006) *North America*	Aerospace	–	–	✓	–

Study: Author(s), Country	Subdomain	Exp[3]	Obv	Lit/D	Actors
Shin, Busby, Hibberd, & McMahon (2005) *UK*	Civilian	–	–	✓	–
Silva & Hansman (2015) *North America*	Aviation	–	–	✓	–
Environment Related:					
Team Collaboration: 10					
Chen & Barnes (2013) *North America*	Military	–	–	✓	–
Hawley, Mares, & Giammanco (2006) *North America*	Military	–	–	✓	–
Joe, O'Hara, Medema, & Oxstrand (2014, June) *North America*	Civilian/Military	–	–	✓	–
Morita & Burns (2014) *Canada*	Civilian/University	–	✓	–	(n=200)
Ososky (2013) *North America*	Civilian/Military/University	✓	–	–	Undergraduate psychology student population. Mean Age: 18.78 (SD 1.61) (n=120)
Phillips, Ososky, & Jentsch (2014, September) *North America*	Military	–	–	✓	–
Phillips, Ososky, Grove, & Jentsch (2011, September) *North America*	Military	–	–	✓	–
Sætrevik (2013) *Norway*	Civilian	–	✓	–	–
Schaffernicht & Groesser (2011) *Multinational (South America/Switzerland)*	Civilian	–	–	✓	–
Smith, Borgvall, & Lif (2007) *UK*	Military	–	–	✓	–
Tasking: 1					
Clancey, Linde, Seah, & Shafto (2013) *North America*	Aviation/Aerospace	–	–	✓	–
TOTAL COUNT: 52					

3.3.2 Human-Automation Interactions

A common trend in HAI literature is priming (or knowledge of the system and its capabilities) which can impact cognitive schema (Beggiato & Krems, 2013; Beggiato, et al., 2015; Lee, et al., 2005, April; Piccinini, et al., 2015). Beggiato and Krems (2013) and Beggiato *et al* (2015) conducted studies of automated cruise control (ACC) and found initial information had an enduring effect on trust and acceptance and trust facilitation observed the Power Law of Learning (Newell & Rosenbloom, 1981). Piccinini *et al.*(2015) found automation error during critical situations had negative effects on behaviour and negative correlations between mental model and ACC operation. Zhang *et al* (2010) also observed learnt mistrust behaviours from automation failure, as a result of over-estimation of system capability with SA and confidence ratings reflective of negative mental schema changes.

Conversely, Lo *et al.* (2015) observed that prior knowledge or experience does not necessarily impact trust or performance with automation, but noted that schema differ with experience associated with goal orientated decision making. Dehais *et al* (2012) observed improper perservation behaviour[6] as a result of psychosocial attitudes and socio-technical demands, decreased performance, proper use of automation and overall trust facilitation in the operators' mental models. Estimated knowledge or capability of a system and the effect on developing appropriate attitudes towards over- or under-reliance on automation also featured in Lee *et al.* (2005, April).

The literature indicates TIA observes a positive trend when systems exhibit transparency and reliability, as these significantly influence construction of appropriate mental models and improvements to SA (Aydoğan, et al., 2014; Birkmeier, et al., 2011, October; Bunt, et al., 2012, February; Lim, et al., 2009, April). Birkmeier *et al* (2011, October) suggest increased trust in automation appropriate human-system interaction (HSI) is key for decision making and over-reliance is connected to LOA (Levels of Automation) and out-of-the-loop system architecture.

[6] Perservation behaviour is the repetition or continuation of a particular response despite the absence or cessation of the initial stimulus. For example, an individual may solely focus on completion of a task to the exclusion of initial overall goal, time taken to complete, communication with team members or other critical tasks (such as DRI).

Lim *et al* (2009) (2009, April) found positive outcomes from transparency in experimental conditions with TIA, however they found that priming did not have a statistically significant effect on mental models or perceptions of the system. Wilkinson *et al* (2007, October) conversely reported primed mental models had improved task performance and TIA, stating participants with lower acquisition (weak schema of the system), compared to higher acquisition (robust mental models), demonstrated higher misuse and disuse with automation.

Three studies used qualitatively grounded psychological approaches (Chua, et al., 2012; Arkin, et al., 2012; Bruemmer, et al., 2007). Bruemmer *et al* (2007) utilised metaphors as primers to modify existing narratives to manipulate existing schema. Similarly, Arkin *et al* (2012) used human psychology, behaviour and attitudes to develop artificial intelligence to simulate HMM and schema regarding moral traits for high-level automation in future battlespaces.

The literature reviews and meta-analyses reviewed cover a diverse range and the following are grouped by the main outcome or theme explored in their respective sources. Education and training to promote appropriate HMM featured in five sources (Fallon, et al., 2010, May; Hoff & Bashir, 2015; Oleson, et al., 2011, February; RTO/NATO, 2007; Schaefer, et al., 2014) and indicated schema can influence the operator's SA and behaviours through dispositional, situational and learned trust which impact facilitation with automation. Augmented, Mixed and Virtual Environment (AMVE) technology suggest embedding training within operational equipment to maintain appropriate schema (RTO/NATO, 2007) as do critical environments within the military domain Oleson *et al* (2011, February), through implementation of appropriate training interventions. Fallon *et al* (2010, May) reports cognitive elements affect operator trust facilitation, such as sense-making to reframe schema after automation failure and error. Shaefer *et al* (2014) reported error ($\bar{g}$ = +.44) and communication feedback failure ($\bar{g}$ = +.45) had negative effects on trust development, whereas scenario training to support continued training has a large mediation effect ($\bar{g}$ = +.79).

Function appropriation is discussed in the research by Hawley *et al* (Hawkley, et al., (2005). .) observing external factors (such as task allocation) and Hoffman and Woods (2011) exploring internal factors

(such as macro-cognition trade-offs between LOA and operator capability). Key outcomes of inappropriate task allocation (Hawkley, et al., (2005). .) suggests fragmentation of workload (creating residuals) occurs between operator and system that cannot be properly framed by existing schema, may lead to interaction errors. Literature recommendations suggest clarification of task and suitable responses to automation error feedback and increase fidelity and transparency to improve trust in operators. Hancock *et al* (2011) meta-analysis found a positive correlation ($r^2 = +0.26$) between factors improving TIA in operators, were supported with literature analysis within the field (2011, September).

A primary problem in psychosocial narrative frameworks in the literature stem from viewing system-teammates as parallel to human-human interaction. Negative impact on appropriate trust development and reliance (Groom & Nass, 2007; Hoffman, et al., 2013) and narrative framework divergence between operators and system designers can also facilitate mistrust with automation (Nachtwei, 2011), and was observed in this review.

3.3.3 Human-System Interactions

Perceived reliability and actual reliability were not aligned in either Barg-Walkow (2013) nor Cassidy (2009), which implies despite reliability increases, perception of automation capability may be limited to the operator schemata. Dawson *et al* (2015, May) reported those with prior training or information had more stable perceptions of capability and were more inclined to trust automation. However, the study also indicated training did not significantly impact HMM in operators, and in some cases detrimental to developing TIA. SA played a role in system transparency and appropriate use, the authors reported spatial, temporal and environmental cues (STEC) are key for operators in developing appropriate understanding of autonomous agent behaviour.

Andersson (2010) reported that miscalibration exists between actual and perceived reality, when there is a dissonance between operator's schema and the technical capability of the automation. For example, when the mental model attributes larger functional capacity to the technical processes underpinning an automated system, there is an over-reliance on that system. The research implies that LOA may affect

operator's mental models through degradation from non-continuous use (skill-fade) and creating cognitive dissonance.

Sheridan and Nadler (2006) reported high occurrences of HAI error as a result of misuse of the system. The main findings found fidelity and transparency are key to maintaining appropriate schema, otherwise incorrect, inadequate or inaccurate feedback may negatively impact trust facilitation. Cognitive divergence and dissonance between human-automation capabilities is a common trend in accident analysis scenarios. Silva and Hansman (2015) reported cognitive divergence as a failure in HSI through lack of transparency and feedback, such as the system changing state without operator input.

Shin, Busby, Hibberd and McMahon (2005) report increased system complexity generates new mechanisms for design and human error. Their analysis supports cognitive dissonance errors in the mismatch between internal representations of system function and reality of operators' existing schemata with increased LOA. Shin *et al* (2005) suggests there is only a partial overlap of mental models of system between designers and users regarding the underlying rationality of the system.

3.3.4 Environment and Social Influences

The environmental based research supports the assertion that positive associations in operator interpretation of accuracy with congruent information, develop robust mental models which improve task performance and facilitate appropriate use of automated systems.

Morita and Burns (2014) explore intuitive trust and HMM and frames group trust to explore the socio-cultural impact of external influences. The study investigates shared mental models (SMM), the impact of interpersonal trust and the regulating factors involved in best facilitating TIA with human-human teams. Sætrevik and Eid (2013) similarly report SMM in human-human-automation teams facilitate appropriate team processes and performance. They observe shared information reflects higher degrees of SMM and appropriate SA. Furthermore, misinformed leadership (influenced by weak mental models and weak SMM with teammates) had a negative impact on team similarity indexes and reported lower accuracy and performance.

Clancey *et al* (2013) explores the use of Brahms GÜM model to verify and validate a theoretical new assessment method for human-system simulations. The model uses cognitive framework to include interactions of pilots and air traffic control operators (ATCOs; modelled to represent human ontologies for different actors) – the research found distributed actants operating without knowledge of the other's actions (e.g. low transparency or increased cognitive divergence) create unexpected behaviour that is difficult to control and simulate.

Hawley *et al* (2006) found better accretion with automation when appropriate mental model framework aids were utilised, and weak schema were accompanied with lower performance outcomes and error prone behaviour. In the Phillips *et al* (2014, September) review, they reported increased levels of misuse, disuse and abuse where operators had weaker mental models and lower TIA and produced an overview of the many antecedents to facilitating TIA. Operator cognitive capacity (such as memory, task allocation, etc.) and HMM is discussed in another review by Phillips *et al* (2011, September) in which transference of human attributes from H-H teams onto H-S teams is discussed. Joe *et al* (2014, June) reviews the mimicry of human behaviour in H-S teams and automation and suggests avoiding utilising human narratives and analogies as system parallels. Cognitive capacity is also reviewed by Chen (2013) into the appropriateness of ecological interface design to HAI and the appropriateness of human-system information exchange impacting on the performance of H-H and H-S teams.

Schaffernicht and Groesser (2011) and Smith *et al* (2007) comprehensively explore metrics utilised in HMM research with a focal point towards individual and shared mental models in team communication and for mission success or goal attainment.

3.4 LIMITATIONS OF THE REVIEW

The limitations are methodological as the scope of the review and quality assessments for critical literature review require stricter guidelines. The synthesis of findings had challenges as guiding principles are ambiguous and undefined (Arksey & O'Malley, 2005). Some sources may have been missed in terminology selection mis-capture, article restrictions and unpublished technical reports.

There are differences in statistical significance and effect size in the trends and results, due to small populations of participants. The literature has limitations with lack of randomisation in experimental study protocols, self-reporting error and androcentric subject populations. The issue of self-reporting is accepted in the studies which utilise validated questionnaires (e.g. Trust in Automation Questionnaire) or appropriate qualitative design protocol. However, the review supports comparable outcomes in the existing literature.

3.5 CONCLUSIONS

The sources discussed support the use of HMM for exploring issues and limitations in facilitating TIA. The outcomes could be utilised in interface design recommendations, operator training and socio-technical bottlenecks.

The experimental design studies overall suggest priming and prior training can positively influence operators TIA through reduction of mistrust or inappropriate behaviour with the system, and appropriate reliance and knowledge of limitations of the system capability. However, cognitive dissonance between perceived reliability and actual reliability, may be a factor requiring more exploration regardless of fidelity of automated systems. Olson, Fisk, and Rogers (2009, October), Ososky (2013) and Wilkinson, Fisk and Rogers (2007, October) suggest distrust and incorrect estimations of automation accuracy were still apparent at 100% precision. Transparency between human and system interaction is vital for providing operators with congruent information which provide accurate and appropriate interpretation of the system's capability and reliability, thereby robustly framing their schemata.

The observational studies support these findings through recommendations of improving transparency, which thus improves SA, especially in environments with high LOA (Aydoğan, et al., 2014; Birkmeier, et al., 2011, October; Morita & Burns, 2014; Sætrevik, 2013). However, priming operators through education and training has mixed reviews in the observational studies (Lim, et al., 2009, April). Nonetheless, the majority of studies reviewed throughout the literature summarised herein indicate that lower acquisition mental models can increase inappropriate behaviour with the system, whereby robust

mental models support improved interaction and facilitation. Cognitive dissonance between actual and perceived reality (such as the capability of the system, or SA) and cognitive overloads impacts trust facilitation, performance and subsequent degradation of mental models if behaviour is not accurately framed by the operator's mental models.

Mental model framework divergence is key in studies reviewed that explored accident analysis of human-system teams, as incidents featured human-system interaction misuse and disuse heavily. Weak schema are associated with lower performance outcomes and error behaviour – together with transparency education and training, may have prevented the incidents and more robust models would re-converge appropriate framework. Training and transparency are recommended to facilitate HAI/HSI in human-human and human-automation team collaboration (both individual operators and groups).

The current literature seeks to explore underlying psychosocial impacts affecting performance and interaction within complex socio-technical environments through exploring mental models and schema. System transparency and operator priming through education can aid in facilitating appropriate trust as levels of automation rise. There are limitations within HAI research, as metrics used are varied and not cohesive across the literature and task performance outcomes are not necessarily representative of internal cognition and attitudes. Qualitative research and inquiry, although less objective, may yield rich contextual data to influence future HSI research in emerging and novel automation interaction and interfaces

The overall scope of current literature is the utilisation of mental models as a theoretical framework for inquiry, is an expanding field of research in the identification of the shifting limitations of TIA research as technology emerges at an ever-expanding pace. Transparency of automation behaviour within the socio-technical system is key to the improvement of appropriate reliance and performance of both system and operator. In addition, training has shown positive effects on trust as it aids in creating appropriate schema, SA and the priming of mental models. The socio-cultural context and environment of human-system interaction is also significant in improving performance and task outcome through team communication, collaboration and leadership.

3.6 EXPLANATORY MENTAL MODEL SUMMARY

As mentioned at the start of this Chapter, it is common to refer to Human Mental Models in the context of Automation. In broad terms, Human Mental Models involve a collection of knowledge held by an individual to account for the behaviour of a system. Human Mental Models thus are frameworks (schema/schemata) that individuals cognitively construct based on prior experience and knowledge to support their expectations and interpretations of their environment. Mental Models are explored across multiple domains, and thus the theories that account for them changes with each paradigm. Although widely accepted that they refer to internal cognitive representations (Al-Diban, 2012), the outcomes or constructed differs across philosophy, cognitive psychology and cognitive sciences. Moreover, there is an implication that a Mental Model has a structure that can be used to make predictions about a system (whether or not these predictions are proved correct, the Mental Model should contain sufficient knowledge to enable reasoning about the system). But this implies a commitment to reasoning and a concept of the structure of knowledge that feels much stricter than the theories explored in this thesis. Indeed, Mental Models tend to operate in the traditions of the cognitive sciences and in terms of individual's interaction with systems. As this thesis is concerned with broader questions of attitude to technology, it makes more sense to refer to 'explanatory models', i.e., rather looser frames which people apply to interpret social and technical systems (Heise, 1979; Harré, 1961; Brown & Clement, 1989). This allows focus on reasoning through a constructivist (rather than reductionistic and cognitive) lens.

4 METHOD AND METHODOLOGY

Parts of this chapter have been published in:

[1] "Trust in Automation: How this is shaped by the human operator and the underwater domain" paper presented at UDT 2019: Undersea Defence Technology Annual Conference, Stockholm, Sweden, 13-15 May 2019.

[2] "Building risk matrices from interview transcripts utilising HCA and IPA" in the Proceedings of Contemporary Ergonomics and Human Factors 2019, Stratford-Upon-Avon, England, 29-1 April May. https://publications.ergonomics.org.uk/uploads/Building-risk-matrices-from-interview-transcripts-utilising-HCA-and-IPA.pdf

The sections: Method on Page 3-4 of [1] and Stages of Analysis on Page 2-7 of [2] have been used to prepare this chapter.

4.1 INTRODUCTION

The objective of this chapter is to discuss the methodological preferences and the chosen approaches to address the research questions. The aims discussed in Chapter 1.3 guide the thesis and these are influenced by the theoretical framework and philosophical orientation of the researcher, discussed in this chapter (Chapter 4, Section 4.2.1).

This chapter discusses the importance of the researcher's theoretical framework in the research design; justification for the methodology used; the techniques used in assessing and analysing the data; and the ethical and security concerns associated with working with human participants.

4.2 RESEARCH PHILOSOPHY

4.2.1 Theoretical Framework and Philosophical Orientation

This section explicitly states the research paradigm of the researcher and the personal beliefs about the nature of knowledge and lens in which reality is viewed. These foundations shape the worldview in which this thesis and the research questions are approached (Koltko-Rivera, 2004). Defining the heuristic assumptions of the researcher serves to appropriately outline the motivations and bias by explicitly stating epistemic, ontological and axiological orientations.

Transparency of research paradigm helps to ground the literature, research methodology and analysis and inform research design and strategy (Creswell, 2014, pp. 36-37; Osanloo & Grant, 2016). For example, Figure 7 establishes suggested elements associated with philosophical orientations and their corresponding research strategies. The following sections outline what is meant by ontology, epistemology and axiology and the researcher's subsequent orientations. The section closes with the philosophical paradigm which contributes to the research design and the implications this theory has on the methodology.

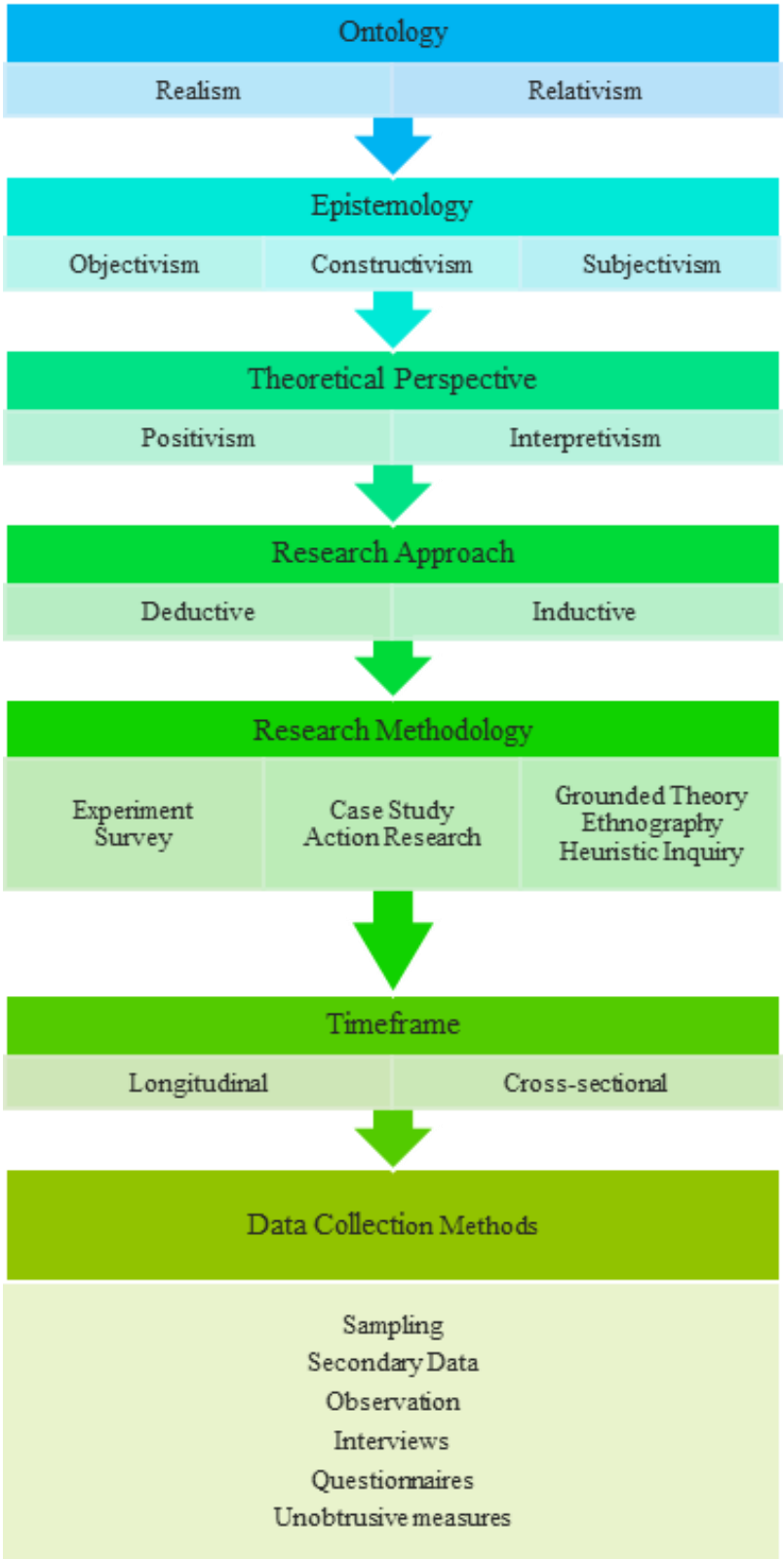

Figure 7 - The elements of the research process [Adapted from Saunders, et al. (2012)]

4.2.2 Ontology

Ontology is a system of beliefs that reflect individual interpretation about the nature of reality (Crotty, 1998). Ontology is defined in the Sage Dictionary of Social Research Methods (Jupp, 2006) as:

> *"a concept concerned with the existence of, and relationship between, different aspects of society such as social actors, cultural norms and social structures… Ontological issues are concerned with questions pertaining to the kinds of things that exist within society"*

Simply, the primary ontological perspectives are whether reality is external to human experiences and independent of human behaviour; or if it is constructed through human social interactions and consciousness. These are respectively, the realist (which include naïve realism, critical realism and historical realism) and the relativist ontologies. The perspectives represent the viewpoint of whether there is a singular or multiple reality. This has an impact on how the research philosophy and approach is then guided as ontological beliefs often steer towards certain epistemological assumptions and methodologies (see Figure 7).

The realist ontology seeks to explain phenomena through reasoning and empirical facts in the attempt to identify a singular unifying truth. On the other hand, the fundamentals of the relativist ontology (Denzin & Lincoln, 2005) is to describe worldview through the multiple perceptions of different individuals through observation. The research in this thesis seeks to explore the dominant narratives of trust in automation and emerging technology, therefore a relativist ontology is respected as this explores how the cultural environment and social interactions shape and alter participants reality. Relatedly, as the research is observed from an outsider perspective, it should be noted that the researcher is observing an etic conceptual framework.

4.2.3 Epistemology

The study of epistemology focuses on what knowledge is and how it is acquired. It looks at the relationship between the researcher and knowledge construction. Defining the epistemological

perspective is important as it clarifies the research approach and strategy outside of the data collection methodologies. It also provides structure for the interpretation of evidence and sensemaking in analyses. For example, if the researcher believes that knowledge is independent of contextual human factors and knowledge is only discovered through impartial observation, this is the foundation of the objectivist epistemology (Crotty, 1998). This often correlates with realist ontology and aligned with the positivist paradigm whereby the values and interpretations of the researcher are excluded as confounding variables.

Subjectivism, the opposition to the objectivist epistemology, is described as *"always filtered through the lenses of language, gender, social class, race, and ethnicity"* (Denzin & Lincoln, 2005, p. 21) and recognises that there is value in knowledge which differs by interpreters and their perceptions. This epistemology is influenced by the collective unconsciousness moreso than nurtured through social interactions. The Constructionist epistemology is similar to subjectivism as it believes there is no singular truth but that it is discovered in the multiple contradicting valuable worldviews within a culture. However, it asserts that social phenomena, actants and the relationship between subject and investigator are linked which construct our reality and knowledge, but that the observer does not claim to be discovering the truth through their interpretations (Guba & Lincoln, 1994, p. 112).

As this research focuses on the role of socio-cultural roles and organisational behaviour towards human-automation interaction, the constructivist epistemology is the most appropriate philosophical orientation to support the research approach.

4.2.4 Axiology

The three main characteristics of defining a research paradigm are ontology, epistemology and methodology (Guba & Lincoln, 1994). However, it has been argued that a fourth – axiology, the inquiry paradigm – should be considered (Heron & Reason, 1997, p. 287).

Axiology explores the motivations behind the research. Or more simply, what the researcher values in the research. This is important as the ethical underpinnings and incentives will impact how the research is conducted, interpreted and presented. For example, under the positivist paradigm, the axiological

perspective may view research in a value-free way where the researcher is disconnected, and an objective stance is upheld. However, the interpretivist axiological stance is value bound, whereby the investigator is intrinsically linked to what is researched and embraces the subjective interconnectivity. These stances affect the presentation of the thesis, as what makes a good researcher or worthwhile science differs between each individual. Axiology acknowledges that the researcher must comprehend and distinguish the role their opinions play in the methodology and analyses of the research.

This research takes an interpretive axiological stance, whereby instead of attempting to eliminate or balance confounding variables, it encompasses the diversity and explores the subjective differences.

4.2.5 Interpretivism

The Positivist paradigm dominated the early to mid-20th century with the explosion of behavioural psychology as a rebellion against the psychoanalysis of the 19th century. In an attempt to objectify, quantify and 'harden' social sciences with scientific theory. Experimental design and behavioural psychology provided more objective metrics to be empirically measured and quantified. Thus positivism, and subsequently post-positivism, quantifies data through structured experimental design, statistical analysis and primarily quantitative data. However, the main critique of this paradigm is that a singular observation which quantifiably refutes a theory, or proposed truth, can disprove the hypothesis (Popper, 1968). Utilising the deductive approach can be especially difficult when attempting to conceptualise a unifying theory for the complexities of human behaviour and cognition. A solution has been proposed by contemporary quantitative researchers, the post-positivist paradigm, in which reality can be studied but all observations are inherently fallible and emphasise probability and mediate individual difference (Onwuegbuzie, et al., 2009). Table 5 provides a basic comparison of assumptions between the positivist and interpretative paradigms (as adapted from Pizam and Mansfeld (2009)).

The Interpretivist stance was developed among researchers who found post-positivism ill-equipped to reflect the nuance and variation in human interactions. Therefore, the inductive process of the interpretative paradigm which is *"culturally derived…interpretations of the social life-world"* (Crotty, 1998, p. 67) seeks to emphasise the differences among people at the centre of the research approach. A

comparison of these two paradigms are outlined in Table 5. The paradigm is primarily influenced by the philosophies of phenomenology, hermeneutics and symbolic interactionism. This research advocates for the phenomenological and hermeneutic stance of interpretivism and is explained in further detail in section 4.3.2.1.

Table 5 - Comparison of positivism research philosophy and interpretivism research philosophy [Pizam and Mansfeld (2009)]

Assumptions	**Positivism**	**Interpretivism**
Nature of reality	Objective, tangible, single	Socially constructed, multiple
Goal of Research	Explanation, strong prediction	Understanding, weak prediction
Focus of Interest	What is general, average and representative	What is specific, unique, deviant
Knowledge Generated	Laws Absolute (time, context, and value free)	Meanings Relative (time, context, culture, value bound)
Subject/Researcher Relationship	Rigid separation	Interactive, cooperative, participative
Desired Information	How many people think and do a specific thing, or have a specific problem	What some people think and do, what kind of problems they are confronted with, and how do they deal with them

For this research, the interpretivist paradigm is adopted. The understanding of how socio-cultural worldviews impresses on multiple viewpoints is key for understanding different narratives. As suggested by Eriksson and Kovalainen (2008, p. 20), the *"interpretivist approach assumes that there are many possible interpretations of the same data, all of which are potentially meaningful".* Furthermore, the research approach for data collection and analysis will not be derived from the statistical analysis of quantitative data (Strauss & Corbin, 1990). The data will be transcripts of focus groups', interviews and documentary analysis (see section 4.4. Research Strategy). The issue of researcher bias are addressed in more detail in section 4.5.

4.3 RESEARCH APPROACH

4.3.1 The Narrative Approach

Narrative inquiry shares commonality with interpretivism and social constructivism by incorporating social interpretations to inform broader social narratives (Riessman, 1993). This strategy can be used to explore both individual perspectives, but also larger collective interpretations as well. This is ideal for exploring individual beliefs on trust and automation, and organisational socio-cultural structures (Casebeer, 2006).

Cronon provides a definition for how narratives can aid organisation of individual and collective perspectives through *"the storied reality of human experience…[and]…our human perspective is that we inhabit and endlessly storied world"* (Cronon, 1992, pp. 1368-1369). Storytelling has a unique way of encapsulating information in an easily digestible format for a diverse array of domains and demographics (Haven, 2007, p. 17). They are 'effective and efficient' for delivering evidence, premises and communicating theory.

Conducting narrative analysis has no best protocol. Part of its strength is that different research philosophies and methodologies can inform diverse, but valid, outcomes. For example, interview data can be interpreted through varied philosophical underpinnings, and methodologies such as discourse- or conversational analysis. Narrative inquiry recognises the individual agency and psychology of participants, as well as their socio-cultural context – and in doing so helps to preserve their integrity and 'irreducible humanity' (Etherington, 2004, p. 81). This is achieved through treating the interview technique as a two-way process with the participant, and treating this account as a collectively assembled narrative (Silverman, 2006; Gubrium & Holstein, 2009; Rapley, 2004).

Narrative analysis is often an interpretative exercise, which is why Interpretative Phenomenological Analysis (IPA) has been included in the mixed methodology used in this thesis for knowledge translation and discussed in more detail in the following section (4.3.2).

4.3.2 Interpretative Phenomenological Analysis and Mixed Methods

4.3.2.1 An Introduction to Interpretative Phenomenological Analysis

IPA is a qualitative analysis method which encompasses dual components: Phenomenology and Interpretation. These are primarily ideographic approaches which aim for an in-depth focus on their subjects, observing analyses of lived experiences and social phenomena.

Phenomenology is dually a philosophical approach and a research methodology that explores the perspectives of lived experiences. Regarding the philosophy of phenomenology, it proposes that an individual is *"embodied, embedded and immersed in the world in a particular social, historical and cultural context"* (Frost, 2011, p. 43). The research method is founded in subjectivity and examination of people's engagement with the world around them and their experiences are the fundamental principles. Subjectivism and IPA additionally have associations with embodied cognition (Larkin, et al., 2011) whereby cognizance is an intersubjective process of sensemaking and not solely isolated to the individual's internal reasoning. For example, Larkin (2011, p. 319) suggests notably IPA can usefully contribute to EASC (embodied active situated cognition) through the wide range of interrelated connections and personal meaning-making in context. IPA can be used to maintain *"personal and contextual worldliness of the body-subject…[and] offers an opportunity to put human life 'back in'"* to sense-making (2011, p. 331).

Regarding sensemaking, IPA emphasises the situated grounding of the data collected. Hermeneutical phenomenology is the practice of interpretation and makes salient the personal meaning and situated context of the participant (Heidegger, 1962). However, problems can arise in the interpretation of both parties, in what is known as double hermeneutics. This is described as:

> *"participant[s] making sense of their personal and social world and the researcher trying to make sense of the participant trying to make sense of their personal and social world"* (Smith, 2004, p. 40)

Double hermeneutics and the interpretations are thus bounded by the participants ability to express and communicate their perspective and the researcher's ability to analyse them. This process therefore can be obstructed through the researchers conscious or unconscious preconceptions or biases (McLeod, 2000). However, IPA researchers acknowledge these inevitabilities and explicitly outline assumptions and intent at the inception of the research through axiology and engaging in intro-and retrospection throughout. It should be noted that:

> *"Importantly, our prejudices should not be thought of as inherently "bad", rather we can have "good" prejudices which can be "bridled" and revised, giving rise to more useful and creative interpretations"* (Dahlberg, 2006; Eatough & Smith, 2017, p. 201)

The disadvantages of IPA are related to the philosophical origins of this methodology. For example, how 'phenomenological' is the interpretation and is it evident in the analysis. How does IPA differ from other qualitative analyses and does the researcher fit into the process appropriately? Clear and transparent communication of the iterative and creative process of reflexive engagement with the participant's dialogue to extract meaning is imperative.

4.3.2.2 Mixed Methods: IPA and Hierarchical Content Analysis

This section has been published in:

[1] (Field, 2019): *"Building risk matrices from interview transcripts utilising HCA and IPA" in the Proceedings of Contemporary Ergonomics and Human Factors 2019, Stratford-Upon-Avon, England, 29-1 April May.* *https://publications.ergonomics.org.uk/uploads/Building-risk-matrices-from-interview-transcripts-utilising-HCA-and-IPA.pdf*

Qualitative research methods are increasingly used to explore user or operator observations in human factors and ergonomic research. These involve the use of unstructured, semi-structured and structured verbal protocol, such as interviews and focus groups (Brinkmann, 2014). However, guidelines to analyse these narrative units are limited in the psycho-philosophical pedagogy (Howard-Payne, 2016)

and within ergonomics model frameworks are the primary method of communicating complex organisational relationships between actants within a system. These theoretical frameworks exist in parallel within both perspectives; for example, the macro-ergonomic Socio-Technical System Model (Rasmussen, 1997) and the constructivist Conditional Matrix (Strauss & Corbin, 1990) both seek to describe the shifting dynamics between humans, objects and human behaviour processes within a multi-levelled system. In both, it is difficult to know how to appropriately populate the frameworks with information gathered from large data sets, such as interview transcriptions. One method is content analysis, based on grounded theory (Glaser & Strauss, 1967; Glaser & Strauss, 2017). However, this technique does not explore the rich descriptive narrative inquiry useful for identifying underlying expressions in dominant themes. The importance of these are to identify barriers, faults or incidents in complex systems which may not be identified in a sweeping quantitative method such as content analysis. Therefore, the approach used in this thesis seeks to systematically analyse qualitative data through Interpretive Phenomenological Analysis (IPA) and Hierarchical Content Analysis (HCA). This is to allow us to appropriately extract information from interview scripts to build risk matrices, registers of concerns and other human factors reporting techniques, thus communicating actant or process issues in macro- and micro- ergonomics systems.

The quantification of qualitative data is often criticised as some believe it undermines the integrity of the source material. However, qualitative analysis is complex and diverse, and often has ill-defined guidelines. Whole and unanalysed transcripts are not viable for presenting data – the approach used (refer to section 4.4.2.1) outlines guidelines that presents data quantitatively without compromising the integrity of the narrative materials. This is through the use of IPA which helps the researcher stay close to the data and explore the unique characteristics of each participant in combination with HCA to identify patterns across data sets and visualise data in an organised framework

4.4 RESEARCH STRATEGY AND DESIGN

The research strategy details methodologies based on the research philosophies detailed in section 4.2. As discussed earlier, the interpretivist paradigm has been adopted, with an inductive approach to analyse

and collect data (Denzin & Lincoln, 2005). The research strategies that are implemented in this thesis are an exploratory study, document analysis and formal qualitative research.

4.4.1 Data Collection

The main sources of data in this thesis are exploratory studies, documentary analysis and semi-structured interviews, as detailed in sections 4.4.1, 4.4.1.2 and 4.4.1.3. The primary data of the responsive interviewing were captured in several UK sites in 2017. These include the University of Birmingham, and three BAE sites across southern England. Secondary data of the document analysis was captured in 2016. Research findings from the data collected are presented in Chapters 5, 6 and 7 respectively.

4.4.1.1 Exploratory Study

An exploratory study is typically performed to explore a problem to gain valuable insight on a topic of interest through scoping literature reviews or open-ended questions to a pilot participant/s. The scope of exploratory research can be large; however, it is commonly used to narrow the broad focus of a topic. Furthermore, an advantage of exploratory research can also provide the researcher better understanding of the literature, the data collection process or other strategies to be modified. The strategy is flexible and can be easily adapted with new information.

Data produced using this approach has the disadvantage of introducing potential research bias into the study if the researcher is seeking to achieve a specific result (Saldana, 2011). If research is favourable, this may in turn bias the focus group and interview data collection as the investigator could purposefully lead or guide answers to better align with their partiality. However, as discussed in section 4.2.4 about the value of research and in section 4.4.1.3.3 on methodological rigour, the bias can be explicitly addressed or mediated.

An exploratory study, was used to provide some bounding due to the vast scope and to identify unconstructive tangents which may occur. The methodology is outlined in the section 4.4.2.1 and was

used to refine this methodology. The exploration of the scope of the literatures was also utilised to appropriately ground the interview questions in theory (see Chapter 5).

This strategy aligns with the interpretivist research philosophy as there is emphasis on exploration of information, rather than explanatory or descriptive.

4.4.1.2 Document Analysis

Document, or documentary, analysis is a methodology in which documents, images and visual artefacts (Bowen, 2009) are reviewed to gain understanding of social phenomena. These are often documents which contain text or visuals which have been collected without the researcher's intervention. Some examples of these in other analyses could be letters, newspapers, radio or television sources and other public records. There is value in obtaining unobtrusive sources of information by providing context through the exploration of language and initial narrative of a desired cohort (Tashakkori & Teddlie, 2010). Document analysis can be used to highlight issues or themes which can shape constructive prompts in later qualitative analysis. The supplementary research in the case of this research was used to explore the idiosyncratic nature of submariner culture and provide a knowledge base in which to structure interview questions around (Hoepfl, 1997).

Document analysis is often used as a method of triangulation in qualitative research (Denzin & Lincoln, 2005) whereby the researcher draws on different data sets to pursue converging or corroborating phenomena. In addition, there are ethical benefits of interpreting multiple data sources as it may reduce researcher bias if artefacts are replicated over different methods. Regarding the document used in this thesis, this was protected information and thus required private access and subsequent scribing. The document was analysed in the same, though simplified, method used for the primary transcript data sets (refer to section 4.4.2.1).

4.4.1.3 Interview and Focus Group Design

For this research, interviews and focus groups are the primary data sets. The collection of this data and the participant section and cohort are detail in the proceeding sections. The findings of these data sets are presented in Chapter 6 and 7. This section discusses the justification and background for the techniques used.

Interviews were used to extract information which cannot be directly observed, but more importantly for this study, were used to explore the narrative context of the socio-cultural sphere in which individuals' schemata have developed.

To do this, there are many types of interviewing strategies. The common characteristics shared by in-depth qualitative studies are as follows: 1) They seek to explore storied responses which diverge from binary responses oft used in survey data collection; 2) Open-ended questions and discussions are encouraged, with disagreement and queries providing additional texture and richness to the analysis; 3) Questions are flexible, fluid and can be shaped by their 'conversational partner'. The term *conversational partner* is a joint collaborative process between interviewer and interviewee to explore experiences and insights (Rubin & Rubin, 2011). It communicates that each participant is a *"distinct individual and has their own cognisance"* (Rubin & Rubin, 2011, p. 33). Narratives are often gained through a selection of conversational partners via a responsive interviewing strategy. This type of interviewing technique is characterised by four main qualities, as defined by Rubin and Rubin (2011, p .38):

1. Pursuit of grounding data in context to report the lived experiences of individuals and their distinct perspective.
2. To build rapport with each conversational partner, acknowledge potential bias and how it may affect exchanges. To remain as neutral as possible and retain composure when confronting challenging material.
3. Ethical considerations must be applied with the duty of care towards conversational partners.

4. Responsive interviewing changes with each subsequent iteration of interview. The evolution of protocol, such as questions and probes, changes with repeated themes or narratives over the duration of the data collection sequence.

Overall, responsive interviewing develops through the restructuring of the discourse between interviewer and interviewee as interpretations influence future lines of questioning.

The principal motive to utilise semi-structured interviews was to establish three main avenues of discussion: perception of trust, perception of automation and socio-cultural influences. The core problem focus seeks to explore whether type of social setting influences behaviour and if there are nuances between formally organised and informal settings with the military cohort. Additionally, how do the established social practices affect attitudes, and how does power or authority interplay with these experiences.

The interplay of experiences is an additional reason for adopting focus groups for some data collection. Focus groups differ from interviews and survey data through shared behaviour, beliefs and can mediate some form of artificiality in the qualitative data collection (Wilkinson, 2011). For example, for this study the researcher additionally performed the role of the moderator, therefore by having a selection of focus groups, this allowed for some reduction of power and control as they could remove themselves from conversations and lessen bias. The following table, adapted from Braun (2013, p. 113), discusses the advantages and disadvantages of utilising focus groups.

Table 6 - Advantages and Disadvantages of Focus Groups (Braun, 2013, p.113)

Advantages	Disadvantages
Flexibility in exploring unanticipated issues	Can be difficult to manage
Good for gathering new knowledge about issues little is known about	Can easily get 'off topic' and be hard to bring back on topic
Access to everyday ways of talking about topics	Logistically difficult – recruitments and organisation
Access to interaction and meaning-making processes	Not a good method to use for busy people
Can facilitate disclosure	Not good for people who are geographically dispersed
Can lead to some level of empowerment of participants or social change	More inconvenient for participants if they have to travel to you, at a particular time
Reduce the power and control of the research, data potentially less influenced by the moderator	Focus groups are generally longer than interviews or more time consuming for participants and researchers alike
Good for groups for whom research participants might be daunting	Transcription of focus group data is very time consuming

In relation to some of the challenges discussed about with focus groups, it is important to note that a Dictaphone was used to record audio. This was so that the researcher could moderate focus groups by several means, rather than focus on transcription or note-taking. For example, having the availability to address each member on the topic for conversation to allow for each participant voice to be heard. In addition, to address different individuals so that conversation is not dominated by other members of the focus group. The limitations of verbal protocol data collection are that they are often performed in a controlled or designated space rather than the participant's natural environment which may cause discomfort or bias. This may also indirectly affect the data filtered through the lens of participants.

The following sections detail the techniques used to generate data, as well as methodological justification for sample size and participant selection.

4.4.1.3.1 Sample Size

Qualitative research uses smaller sample sizes due to the methods of inquiry and the analysis outcomes (Patton, 2005; Braun, 2013, p. 55).Table 7, adapted from Tesch (1990), illustrates a metanalysis of typical qualitative study methods and the sample size ranges common to those techniques. The numbers highlighted in bold are methods adopted or adapted from in this thesis. The number of participants in qualitative studies are often dependent on several factors, such as research approach, accessibility, adequacy and data saturation (Baker & Edwards, 2012). Regarding data saturation, this occurs when additional data does not generate new information or has reached a point where themes are suitably supported.

Table 7 - Descriptive statistics for each methodological group (Adapted from Tesch, 1990)

	No. of studies found	No. of studies after inclusion criteria applied	Range		Measures of central dispersion			
			High	**Low**	**Mode**	**Mean**	**Median**	**St. Dev.**
Action research	140	28	67	3	6	23	17	18.4
Case study	1401	179	95	1	40	36	33	21.1
Collaborative research	8	2	25	5	-	15	15	14.1
Content analysis	**213**	**42**	**70**	**2**	**30**	**28**	**25**	**14.7**
Critical / emancipatory research	6	3	42	21	-	35	41	11.8
Discourse analysis	157	44	65	5	20	25	22	15.3
Ethnographic contents analysis	2	2	52	22	-	37	37	21.2
Ethnography of communication	1	1	34	34	-	34	34	-
Ethnomethodology	7	2	55	11	-	31	31	27.6
Grounded theory	**429**	**174**	**87**	**4**	**25**	**32**	**30**	**16.6**
Holistic ethnography	1	0	-	-	-	-	-	-
Hermeneutics	**19**	**9**	**42**	**7**	**-**	**24**	**26**	**10.2**
Life history	61	35	62	1	21	23	20	16.1
Naturalistic enquiry	2	1	26	26	-	26	26	-
Phenomenology	**57**	**25**	**89**	**7**	**20**	**25**	**20**	**19.9**
Qualitative evaluation	7	1	42	42	-	42	42	-
Symbolic interactionism	22	12	4	87	-	33	28	26.5
TOTAL	2533	560	95	1	30	31	28	18.7

It is suggested that for all qualitative research, fifteen participants are the smallest acceptable sample (Bertaux, 1981, p. 35; Guest, et al., 2006). Specifically for phenomenological studies, it has been identified that at least six participants should be used (Morse, 1994), while just over two thirds identified (68%) fell within a suggested range of five to 25 (Creswell, 1998). For a doctoral thesis it is suggested that five to 25 conversational partners are suitable to gain data saturation (Morse, 1994, p. 225). Grounded theory methodology suggest between 20-30 participants, for grounded theory studies as recommended by Creswell (1998) or within the range of 30-50 suggested by Morse (1994).

For this thesis, 26 participants in total were recruited (one omitted as an exploratory study, one omitted due to data corruption of audio file). The next section discusses the important aspect of participant sampling and how the cohorts were selected.

4.4.1.3.2 Participant Selection

The diversity within the sample size of the participants is important as not all conversational partners can articulate or are as equally perceptive (Tashakkori & Teddlie, 2010, p. 191). Therefore, a range of lived experiences, backgrounds and narratives were the focus of the type of sampling used in this research. Sample size was not used to gain statistical significance but to appropriately acquire data saturation, as above, this was based on Table 7 and recommendation by Mason (2010) and Creswell (1998) to answer the research questions.

The participants were selected purposefully to aid in addressing the research questions and to facilitate in-depth and rich data sets to generate themes (Teddlie, 2007). The criteria below are from the participant information sheet and consent forms given to all conversational partners and can be additionally found in Appendix B (Participant Information Sheet (Basic); Participant Information Sheet (Full) And Consent Form). The participant requirements were as follows:

- A civilian member of the university (staff or student) who has an interest in emerging technology and/or automation. These individuals are not restricted by age or student status.
- OR:

- An individual of the military. This participants are not limited by age or status though must be familiar with command and control/operations. A diverse range of participants are sought, and experience pre- and post- service are recommended OR An individual from BAE Systems who has preferably with maritime or naval experience (whether from active duty or working in the domain as part of BAE Systems). This participants are not limited by age or status though must be familiar with command and control/operations.
- AND: Technologically literate (for example, a digital native or immigrant).
- Participants must be between the ages of 18 and 30 – however if age diversity is required this can be expanded outwards up to approximately 60 (so as within working age)

This purposeful selection was used to create a selection of adults with diverse experience, specialities and military experience in a heterogeneous sample. Both the consulting company, the forces and universities used in this research pride themselves on diverse demographics thus provided access to the heterogeneity. Furthermore, trust is inherent for adequate function of a structure hierarchical workforce in the military so identifying these cohorts had additional benefits for this research. In addition, the importance of trust is culturally significant for both the military cohort as it is a pre-requisite for security clearance to work for BAE. The values of the organisation require trust and reliance in the constantly transforming and revolutionising projects or workload. Furthermore, the primary investigator additionally required security clearance which plays a role in the credibility and authenticity of the ethical standpoint of the researcher. As the sponsor for this research project, utilising BAE Systems was an appropriate organisation for the recruitment of personnel. Access was achieved through liaison with management and benefitted from the confidentiality, as participant information cannot be shared.

Purposive sampling recruitment of participants was utilised in this research to obtain data from individuals who have experienced the phenomena under investigation per the research questions outlined in Chapter 1, section 1.3 (Gentles, et al., 2015). These semi-structured interviews (responsive interviewing) were flexible so as to engage in appropriate rapport with the individuals, as well as explore diverging and converging narratives.

Six civilian participants were recruited through voluntary purposive sampling. Nineteen military participants were collected through snowball and purposive sampling through staff contact at BAE Systems. These include: 5 Surface Navy, 8 Subsurface Navy, 4 Army and 2 RAF personnel.

With regards to studying an organisation, in addition to purposive sampling through a third party, there is the limiting and problematic factor of 'gatekeeping' (Becker, 1998, pp. 90-91). This may limit what can be studied, as well as limiting the openness of interpretation through redaction. To circumvent this potential problem, Becker (1998) suggests two solutions: *"doubt everything anyone in power tells you";* second, *"look for opinions*". This can also be achieved through systematic attempts to assess linkages between opinions, activities and interests.

4.4.1.3.3 Interview Protocol

The interview and focus group protocol included engagement, exploration and exit questions (Lincoln & Guba, 1985; Morrison, et al., 2012). The purpose of these were to encourage rapport, ask formal questions and ethically close the discussions, respectively. Ethical considerations in qualitative research and interviewing are discussed further in section 4.5. The detailed protocol grounded and bounded by previous research and via responsive interviewing techniques allows for researcher control over the line of questioning and focus on specific topics.

The interview protocol for this thesis is as follows:

1. The interview questions (see Table 8) typically start with a vague open-ended question to relax the potentially tense or formal atmosphere. These will be followed by approximately 5 questions which cover topics imperative to the research plan.
2. Prompts and probes are used to follow-up or explore ideas or opinions expressed by the participant/s. Elaboration on topics which are of interest or to address miscommunication or misunderstanding are be utilised.
3. A final statement of thanks to participants for acknowledgement of the time spent being interviewed to provide duty of care was given.

4. A research log was kept to record thoughts arising during the interview. This is also pertinent to qualitative methodological rigour as detailed in section 4.4.1.3.3.

Outlining a verbal protocol of questions to guide the conversation is important to explore crucial lines of inquiry with each participant interviewed. This can allow the researcher to freely explore questions or probes that spontaneously emerge, build rapport organically yet still explore topics that have been predetermines. It also provides a checklist for the interviewer that the relevant topics have been appropriately covered in lieu of tangents. An advantage to outlining a question guide is that it can be beneficial when there is limited time available in the data collection situation. This is especially of concern for focus groups as they are often more demanding of time and energy from both participant and interviewer.

When conducting the semi-structured interviews and focus groups, it was important to establish thematic structure to the verbal protocol to appropriately address the research questions. The following protocol outlines key questions which were utilised to address crucial themes in the research (refer to Table 8). These are trust, automation and culture. Although these questions are vague, key themes which are elaborated from this address how trust is defined; what automation and technology means to individuals; what experiences or culture shape these views; and exploration of category boundary actants and artefacts. The guide has been developed in low detail due to utilising the responsive interviewing technique. Through the subsequent iterations and interviews, the questions developed, the sequence of delivery was modified, and new or unexpected topics where covered in the framework.

A basic script prior to interview was developed so that introductions; a brief overview of the session; and items covered in the consent and information sheets could be covered and repeated to all participants. As recommended by Castillo-Montoya (2016), interview questions covered aspects of background information and discussion points to cover research questions adequately without overwhelming the conversational partners. The table below (Table 8) outlines the verbal protocol in which perspectives could be captured to yield appropriate data to answer the research questions. Dispersed within these were follow-up questions which seek to gather detail, richness, vividness, and nuance to assure thoroughness, in addition to explore further events, concepts and themes arising from

the main questions. These are designed in response to the comments or ideas introduced by the conversational partner and often worded to reflect prior answers without being loaded or leaning questions. The script is open-ended so that interviewees offer additional information, and the use of prompts and probes can allude to introspection. These include detail or elaboration probes, clarification probes, silent probes and echo probes.

Table 8 - Example Verbal Protocol

Interview and Focus Group Questions	
1a. Can you tell me what your opinions of automation are?	To identify how the terminology is defined. To explore the underlying expressions of types of technology.
1b. [follow up in terms of detail and terminology, following order of narrative]	To establish common meaning and how different people may apply different terms.
2a. What are major changes you perceive in the implementation of increased levels of automation (LOA)?	To explore the attitudes of different types and expressions of technology and which have differing or unexpected viewpoints.
2b. [as in 1b]	
3a. Can you think of something that you have read, seen or heard about which has made you think in a specific way (negative/positive/neutral) about new technology? Anything [not necessarily about automation, but emerging technology as a whole].	To discover experiences, socio-cultural factors and explore how narrow, vague or wide participants definition of technology is.
3b. [as in 1b]	
4a. Can you tell me about an experience which may have shaped the way you feel?	Explore lived experiences, underlying narratives and socially constructed views.
4b. [as in 1b]	

5a. Can you tell me how you would define trust? How would you describe trust and/or trust in automation/technology/systems?	To establish common meaning.
5b. [as in 1b] [Are the terms trust and reliance interchangeable to you, or do you think there is a difference?]	To explore trust as a concept and associated moral and/or ethical viewpoints.
6a. Can you tell me about how your peers may shape your views on these topics? How do your peers/teammates feel about trust/technology/levels of automation?	To specifically address the socio-cultural and shared narratives of participants, how it is perceived to affect and whether underlying narratives are congruent.

The questions are guided by literature review, document analysis and the exploratory study. The basic questions were used to build rapport, for example, starting the conversation with basic background so participants feel more at ease. Questions were also arranged so that difficult or contentious questions are discussed once confidence has been built with the interviewee. The colloquial wording of the questions was so that they are not leading, but also provides a subtle prompt to the conversational partner to begin talking and allows for multiple directions for them to explore naturally. As responsive interviewing was used, this allowed for flexible adjustments and immediate revisions of interview protocol through 'emergent design' (Creswell, 1998). The open-ended questions allowed for rich and textured data analysis through meandering tangents and free-form discussion on themes which may not have emerged in a more structured question. The relaxed structure of the data collection therefore required more defined and coordinated analysis methodology to confront the vast data sets. The systemised approach is detailed in the following section.

4.4.2 Data Analysis

The interview protocol, as described in section 4.4.1.3.3, covers broadly the research questions outlined in chapter 1, section 1.3. These were kept broad as to not overwhelm or lead the participants, but to also adapt to time availability of interviewees. A total of 19 individuals participated from the BAE cohort,

with 4 in standalone interviews and 15 across variably sized focus groups (between 2-3 per group). Table 9 outlines the overall participant demographics. More detailed summaries are provided in the chapters which explore the Civilian and Military cohort studies at length. The length of the interviews ranged from 48 minutes to 100 minutes for individuals; and 121 minutes to 157 minutes for groups. For the civilian cohort, a total of 7 individuals were recruited. The exploratory study was an interview of 40 minutes, followed by a 56-minute interview and two focus groups (2-3 per group) ranging between 134-138 minutes. Unfortunately, the civilian solo interview had to be removed from the study due to data corruption of the audio file.

The primary data for the interviews and focus groups consisted of 13 audio files which were transcribed using 'Routine' (orthographic) transcription. The analysis of interviews were transcribed verbatim. This produced a total of for the military cohort a total of 121438 words; including civilian cohort, a total of 143443 words; including the exploratory study and primary data, a total of 148525 words. This excludes the word count for interviewer questions and responses. For some pertinent quotes, Jefferson Transcription was also used to give further context and conversational analysis (Jefferson, 2004).

Table 9 - Interview and Focus Group Participants

Interview/Focus Group Number	Number of Participants	Occupation/Domain
Exploratory Study	1	Education Specialist [Civilian-Military Spouse]
1	2	Electrical Engineer, Physicist [Civilian]
Omitted	1	Civil Engineer [Civilian]
2	3	Social Scientist, 2 x Computer Science Specialists [Civilian]
3	2	Army, Subsurface Navy [Military]
4	3	3 x Subsurface Navy [Military]
5	3	Subsurface Navy, 2 x Surface Navy [Military]
6	1	Army [Military]
7	1	Army [Military]
8	1	RAF [Military]
9	2	RAF, Surface Navy [Military]
10	1	Army [Military]
11	2	2 x Subsurface Navy [Military]
12	3	2 x Surface Navy, Subsurface Navy [Military]

Throughout the data analysis phase, meaningful comments and phrases from the raw data were interpreted and coded as items, sets or collections of themes that were to be compared and contrasted. Interpretative phenomenological analysis and hierarchical content analysis was performed on these semantically similar artefacts. For a more in depth analytical protocol, refer to the following section 4.4.2.1 (*Stages of the Analysis*). In general, these clusters were further connected or separated with quotes of similar meaning and organised by sets of similar meaning. This coding or labelling which is used to condense a set of artefacts, produces a large number of inter-related themes. Using the coding process adapted by Creswell and Clark (2004) (see Figure 8) the units were collated into; categories

and three levels of super- and sub-ordinated themes. This, in addition to some simple statistical analysis, comprises the hierarchical content analysis by which the themes are weighted for saliency and priority. In regard to the interpretative phenomenological analysis the patterns of convergence and divergence between and within the units and clusters provides texture in the analysis and outcomes (Eatough & Smith, 2017).

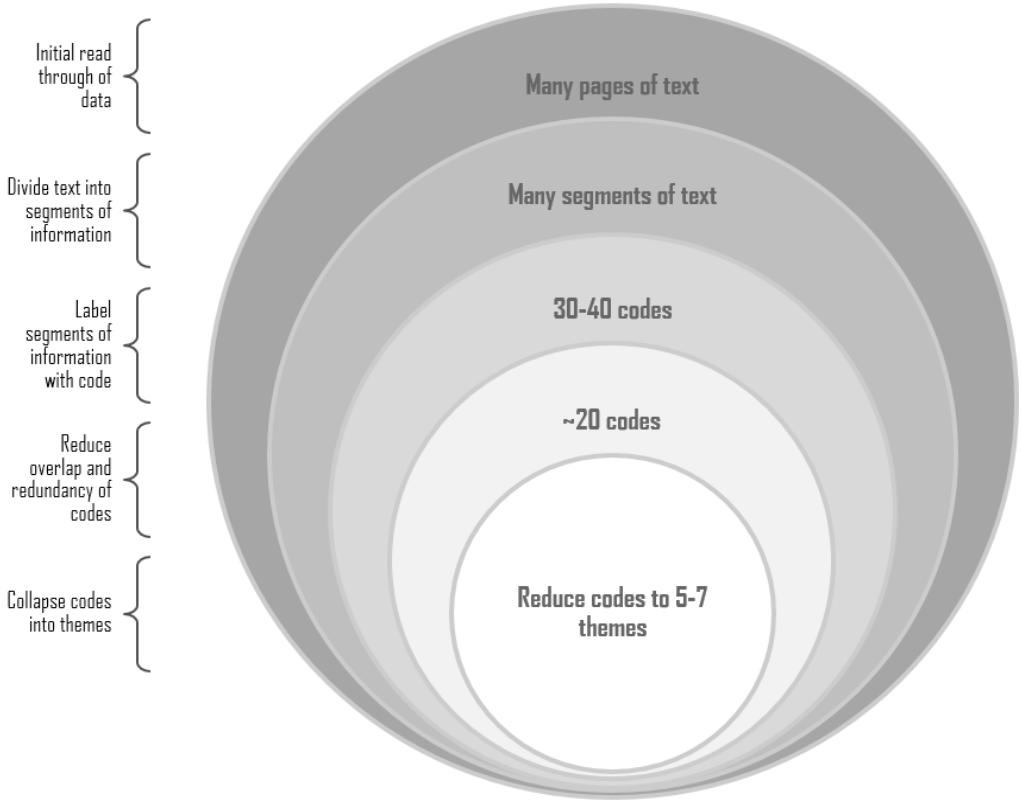

Figure 8 - Coding Process (Adapted from Creswell and Clark [2004])

In qualitative analysis, a phrase which is often used to define and explore artefacts is that of "emerging themes". This thesis is careful of using the term 'emerging', as it suggests the implication that themes pre-exist the analysis and that the researcher does not play an active role in the production or generation of themes (Braun & Clarke, 2019). In addition, 'themes emerging' suggests a realist and/or positivist discovery to the research, which is in consistent with the theoretical framework and the epistemological

and ontological orientation of the researcher and their perception of knowledge and reality. The reflective thematic analysis (in this case, Hierarchical Content Analysis) provides an analytic process in which to generate themes to "*represent patterned meaning*"[7].

The main potential of qualitative research is the *recontextualization* of the interpreted data, which develops a continually evolving theory whereby the research can be made applicable to other domains, social structures and populations. The following quote by Morse (1994), encapsulates (in the opinion of the researcher), why utilising mixed methodology where contrasting but complimentary techniques can be used harmoniously to extract the pivotal artefacts of substance:

> *"Theorising is the constant development and manipulation of malleable theoretical schemes until the 'best' theoretical scheme is developed. It is a process of speculation and conjecture, of falsification and verification, of selecting, revising and discarding. If one ever finishes, the 'final' solution is the theory that provides the best comprehensive, coherent and simplest model for linking diverse and unrelated factors in a useful, pragmatic way.* ***It is a way of revealing the obvious, the implicit, the unrecognised, and the unknown. It is a way of discovering the insignificance of the significant and the significance of the insignificant****."* (Morse, 1994, p. 32)

In summary, the transcripts utilise interpretive phenomenological analysis for initial identification, coding of raw themes and clusters of data units, and hierarchical content analysis for classification of categories, central themes and extrapolation of narratives. The analysis method seeks to systematically review themes to inform risk matrices and concern registers at both the sematic and latent thematic level (Braun, et al., 2014). The use of interpretative and descriptive approaches to classify the data provides a method of numerical communication of qualitative research to help convey recommendations in

[7] Clarke, Victoria (drvicclarke). "9/10 The researcher as sculptor works with materials (data) that delimit what they can say & they bring to the analytic process their experiences, skills, tools & techniques - they actively generate themes to represent patterned meaning in their data... Rather than themes emerged" 11:28 AM · Jul 5, 2018. Tweet.

technical (and/or quantitative reliant) domains and to appropriately communicate the in-depth rich data of narrative inquiry.

4.4.2.1 Stages of the Analysis

This section has been published in:

[1] (Field, 2019): *"Building risk matrices from interview transcripts utilising HCA and IPA" in the Proceedings of Contemporary Ergonomics and Human Factors 2019, Stratford-Upon-Avon, England, 29-1 April May.* *https://publications.ergonomics.org.uk/uploads/Building-risk-matrices-from-interview-transcripts-utilising-HCA-and-IPA.pdf*

Stage 1 – Notetaking and Preliminary Categorisation

The researcher should make notes during and post-interview regarding the areas and topics discussed throughout the interview. This is more streamlined for those who undertake semi-structured and structured interview protocol. These 'memos' (Pandit, 1996) will become initiators which will elicit how to categorise data, record ideas and aid in the development of themes in later stages.

Stage 2 – Immersion

At this stage, the researcher gathers in-depth familiarity of the data sets through immersion in the data. Through re-reading the transcripts, the researcher can identify major points of interest, general feel and context which will generate exploratory codes.

Stage 3 – Interpretive Phenomenological Analysis

The analysis process of Interpretative Phenomenological Analysis and Hierarchical Content Analysis are similar in the first three phases (see Table 10 and Table 11). In this method, the similarity provides parallel analysis in the exploration of the transcripts. This will influence the raw themes generated for the hierarchical content analysis. IPA is used as it facilitates the immersion required to adequately analyse small or homogenous data sets such as those which use interview data (whether for narrative inquiry, user experience, focus groups and more). Furthermore, IPA seeks to describe and explore lived

experiences, thus is appropriate for exploring narratives as it reports rich descriptions and characteristics identified in the data and framing the context.

The immersion in the data of this stage provides the context and knowledge of the data set which will shape the categorisation of the risk matrix consequences in stage seven.

Table 10 - Analysis Procedure for Interpretative Phenomenological Analysis (Adapted from Sparkes and Smith (2013)

Phase	Description of Process
1. **Immersion/Searching for themes in the first case**	In-depth familiarity of data through engagement with collected information (e.g. multiple re-reading of data set transcriptions). Exploratory codes generated from first case.
2. **Identifying and Labelling Themes**	Data sets are coded to capture essential quality of transcript.
3. **Connecting Theme**	Refocus analysis to identify themes from data. Cluster and connect emerging concepts that share similar meaning.

Stage 4 – Hierarchical Content Analysis

Hierarchical content analysis provides secondary analysis and seeks to explore themes inter-transcript through systematic cross-checking and confirmation. In this approach, HCA is used to outline the context in which the weighting of the themes (as described in Stage 5 and 6) is based on. The exploratory themes outlined in the previous stage are then clustered into related units (coded nodes) in a clear and concise tabulated form.

Table 11 - Analysis Procedure for Hierarchical Content Analysis (Adapted from Sparkes and Smith (2013)

Phase	Description of Process
4. Cross-checking	Examination of identified themes and clustered groupings against original data set.
5. Confirmation	Review of raw themes and clustered groupings.
6. Produce a Table	Production of data within a table of hierarchical structure to display units generated by nature of themes.

Stage 5 – Thematic Weighting

Due to the diversity and large size of transcription data sets, it can be difficult to objectively explore which themes are more prevalent than others, and whether they are relevant compared to their super-ordinate themes. Therefore, this approach seeks to succinctly divide salient themes into 'absolute' and 'relevant' themes. The term 'absolute' in this approach refers to the amount of words coded in references to the super- and sub- ordinate themes occurring in the total word count of the data set. For example (see Table 12), the number of references discussing all topics relating to cultural aspects equal 2603, which equates to 51.3% of the entire interview data set (see Table 14). However, to address the relevancy of these sub-themes within the category boundaries, the relative content weight was calculated. As references can be coded to multiple themes, the node summary report provided by NVIVO (QSR International Pty Ltd, 2012) provides number of words coded to reference nodes (total words referenced at all nodes is equal to 2603 in the given example). Therefore, the relative weighting of culture in this example is equal to 27.9%. Further breakdowns of salient themes are provided in Table 17.

Table 12 - Table Outlining the Calculation of Absolute and Relative Theme Weighting (Example)

Document Word Count	5082
Coded Word Count	9332

Category	Master Theme	Superordinate / *Subordinate Themes*	Coded Word Count
Cultural		**Category Total**	**2605**
		All Culture	663
		Culture Only	165
		Identity	0
	Culture	Age	165
		Technology Acceptance	81
		Media Influence	61
		Zeitgeist	191
		All Terminology	576
		Terminology Only	0
	Terminology	Automation	118
		Chaoplexic Definitions	458
		Trust	0
		All Schema Priming	1366
		Schema Priming Only	0
	Schema Priming	Contraditions	83
		Negative Experiences	1283
		Positive Experiences	0

Stage 6 – Content Weighting

This stage addresses how relevant the subtheme (e.g. identity or cultural roles) may be to its super-ordinate theme (e.g. cultural) by categorising these by the percentage of coded words (Table 13 and Table 14) so as to provide provisional focus to how commonly the themes occur in the transcripts and whether they require more in-depth review via IPA. The content weight (low, medium, high) provide an axis for characterising the risk matrix in the proceeding stage.

Table 13 - Table of Weighting Categorisation

Level of Weighting	Description
1. Low	Low participant weighting is allocated to issues which achieved below the 33rd percentile of relative theme to category total.
2. Medium	Medium participant weighting is allocated to issues which are within the 33rd and 66th percentile of relative theme to category total.
3. High	High participant weighting is allocated to issues which exceed the 66th percentile of relative theme to category total.

Table 14 - Table of Themes with Saliency and Content Weighting (Example)

		Salient Themes		
Category / Theme / Sub-Theme	**Consequence**	**Absolute**	**Relative**	**Content Weighting**
Cultural		**51.3%**	**27.9%**	
All Culture		13.0%	7.1%	
Culture Only	1	3.2%	1.8%	6%
Identity	2	0.0%	0.0%	0%
Age	2	3.2%	1.8%	6%
Technology Acceptance	3	1.6%	0.9%	3%
Media Influence	1	1.2%	0.7%	2%
Zeitgeist	1	3.8%	2.0%	7%

Stage 7 – Severity Categorisation

Interpretive Phenomenological Analysis is used to populate the categorisation of consequences within the risk matrix through the in-depth familiarisation the researcher will have with the data set. An example for characterising the severity outcomes are outlined in Table 15. A guideline, as adapted by Derrico et al, (2011), of characterisation of consequences are as follows: high priority outcomes can be topics that necessitate urgent resolution which may severely impede function and/or satisfaction of using a system; medium ranked severity might be concerns that are obvious but do not require immediate change; lower priority may indicate issues with simple or superficial solutions with negligible or limited risk. Input from Subject Matter Experts (SMEs) can be utilised to impart additional knowledge to appropriately weight risk factors and severity risks to actants and stakeholders.

Table 15 - Table of Consequence Severity Outcomes (Adapted from Derrico, et al., 2011)

Outcome Severity	Description
1. Minor	Negligible or Minor injury or adverse physical or mental health outcome.
2. Moderate	Moderate injury/ Moderate adverse physical or mental health outcome.
3. Serious	Serious injury/ Serious adverse physical or mental health outcome.
4. Major	Major Injury/Major adverse health physical (e.g. possibility of permanent function loss) or mental health outcome (e.g. major depressive episode).
5. Catastrophic	Possibility of death or permanent loss of function (motor, sensory, cognitive).

The traditional risk matrix (see Table 16) populated using the categorisation outlined in stage 6 and 7 can be used to apply priority weighting to the themes observed in the data set (see Table 17).

Table 16 - Matrix to Calculate Level of Priority

			Consequence				
			Minor	Moderate	Serious	Major	Catastrophic
			1	2	3	4	5
Weighting	Low	1	2	3	4	5	6
	Medium	2	3	4	5	6	7
	High	3	4	5	6	7	8

Table 17 - Table of Salient Themes with Content and Priority Weighting

Category / Theme / Sub-Theme	Consequence	Salient Themes Absolute	Salient Themes Relative	Content Weighting	Priority
Cultural		**51.3%**	**27.9%**		
All Culture		13.0%	7.1%		
Culture Only	1	3.2%	1.8%	6%	Low
Identity	2	0.0%	0.0%	0%	Low
Age	2	3.2%	1.8%	6%	Low
Technology Acceptance	3	1.6%	0.9%	3%	Medium
Media Influence	1	1.2%	0.7%	2%	Low
Zeitgeist	1	3.8%	2.0%	7%	Low

Stage 8 – Priority Categorisation

With all themes given an appropriate priority level, this can help to identify the urgency with which issues should be addressed. These can populate a table or register of issues generated by the interviews which outline the specific area of concern where the issue or problem occurs. Further useful information includes the number of participants which it affects, other feedback or quotations to support and justify recommendations (for example, see Figure 9).

Main Category	Specific Category	Issue	Additional Information	% Absolute (Relative)	Consequence	Weighting	Priority
Cultural	Zeitgeist	Views with a temporal element relative to the socio-cultural environment participants are based in.	Schema or mental models that have been shaped by the current or historical defining beliefs of that culture. Examples include introspected generational differences and cultural shifts relation to technological improvements.	3.8 (2.0)	1	1	2

Figure 9 - Example of Concerns Register

4.4.3 Methodological Rigour

Interpretivism is often criticised for being 'soft' due to its lack of objectivity. Richards (2003, p. 6) states that qualitative inquiry "*is anything but a soft option – it demands rigour, precision, systematicity and careful attention to detail.".* Trustworthiness is a cornerstone of interpretative research, thus quality assurance for qualitative research is imperative for methodological rigour.

Lincoln and Guba (1985) created a basic framework for corresponding traditional quantitative rigour against appropriate qualitative counterparts. These are as follows:

- *Objectability:* As discussed in section 4.2, there is significant debate between the merits of objectivism and subjectivism. As the interpretivist philosophical paradigm has been outlined, the theoretical process is explicit in acknowledging the inextricable involvement of the researcher as a data collection instrument to the enquiry. The outcome is a co-creation of knowledge between the interviewers and interviewees (Brinkmann & Kvale, 2015). Therefore, for rigour, the concept of **Confirmability** suggests the means to which outcomes can be corroborated or confirmed by other people. The investigator can inspect the data throughout the project; another research can take the role of "devil's advocate"; or a data audit can be conducted.
- *Internal Validity:* For quantitative research, validity refers to the truth or certainty of the research findings. Both qualitatively, and viewed through an interpretative lens, there is no single truth. Therefore, quality is associated with the **Credibility** and **Authenticity** of the researcher. The attribution of validity comes from the researcher's credibility, competence, thoroughness and integrity (Patton, 2005). In addition, credibility is reflected through the data and findings and whether they truly reflect the participants experience. This can be achieved through: observing direct connections between findings and data collected (argumentative validation); rich and in-depth context or descriptions; and use of multiple data sources for triangulation (Trochim, 2006; Yilmaz, 2013; Sarantakos, 2013). The data can be deemed

accurate through unedited audio and transcripts (communicative validation), which is further related to auditability.

- *Reliability:* Reproducibility and quality of empirical metrics or equipment increase the reliability of data analysis and outcomes in quantitative research. For the comparable qualitative context, this is difficult when the researcher is the data analysis instrument (Creswell, 2014). The consistency and trustworthiness of the researcher is thus imperative for **Dependability** and **Auditability** (Brinkmann & Kvale, 2015). As the epistemology has been defined in this thesis as social constructionist, it is impossible to replicate both the data collection and findings. However, what is offered is insight into the lived experiences and perspectives of participants in a defined context. Dependability and auditability can be achieved through clearly communicated study protocol; explicit research philosophy; and transparency of research bias and background (Trochim, 2006; Yilmaz, 2013).
- *External Validity:* There is the expectation that research must have the ability to be transferred to similar contexts and maintain reproducibility. As suggested above, empirical validity cannot be maintained when variance and interpretation is at the core of the research approach. **Generalisability** can only be applied in the sense that there is a representation of the population in the data collected. More importantly, there is **Transferability** and **Fittingness** - which aims to produce findings that can be transferred to others with sufficient similarity or level of contextual 'fittingness' (Patton, 2005; Lincoln & Guba, 1985). Appropriate transferability can be attained through clearly described context and detailed research assumptions which inform the research (Trochim, 2006; Yilmaz, 2013).
- *Utilisation:* Similar to reliability and confirmability, as statistical data or numerical outcomes are not the key focus or analysis used in interpretative research strategies, it is not possible to lift empirical data for stakeholders. Therefore, it is suggested that recommendations and theory based **Application** is utilised to holistically export the data outcomes.

In addition to the qualitative rigour suggested by Lincoln and Guba (1985), a universal criterion for qualitative research best practice has been proposed by Tracy (2010) (refer to Table 18). They suggest

that the following eight key markers conceptualise a useful pedagogical model for a variety of qualitative researchers. The rules and guidelines outlined provide a more flexible approach to methodological rigour.

Table 18 - Eight "Big-Tent" Criteria for Excellent Qualitative Research; Adapted from Tracy (2010)

Criteria for quality	**Various means, practices, and methods through which to achieve**
Worthy topic	The topic of the research is • Relevant • Timely • Significant • Interesting
Rich rigor	The study uses sufficient, abundant, appropriate, and complex • Theoretical constructs • Data and time in the field • Sample(s) • Context(s) • Data collection and analysis processes
Sincerity	The study is characterized by • Self-reflexivity about subjective values, biases, and inclinations of the researcher(s) • Transparency about the methods and challenges
Credibility	The research is marked by • Thick description, concrete detail, explication of tacit (nontextual) knowledge, and showing rather than telling • Triangulation or crystallization • Multivocality • Member reflections
Resonance	The research influences, affects, or moves particular readers or a variety of audiences through • Aesthetic, evocative representation • Naturalistic generalizations • Transferable findings

Significant contribution	The research provides a significant contribution • Conceptually/theoretically • Practically • Morally • Methodologically • Heuristically
Ethical	The research considers • Procedural ethics (such as human subjects) • Situational and culturally specific ethics • Relational ethics • Exiting ethics (leaving the scene and sharing the research)
Meaningful coherence	The study • Achieves what it purports to be about • Uses methods and procedures that fit its stated goals • Meaningfully interconnects literature, research questions/foci, findings, and interpretations with each other

It should be noted that a fixed quality appraisal or qualitative criteriological approaches, like the ones discussed in this section, can be potentially problematic. Smith and McGannon (2018) suggest the implementation of a universal criteria does not appropriately cover the researcher's theoretical framework and philosophical orientation. Furthermore, those using the relativist approach, the quality of research is not predetermined, and is open-ended and subject to change – and therefore can be limited by the characteristics outlined. For transparency, methodological rigour and explicit research philosophies are detailed in this chapter to critically engage with all methodological concerns.

4.5 ETHICAL AND SECURITY CONSIDERATIONS

Approval from the University of Birmingham's Ethics Committee was obtained prior to data collection and data analysis followed the University's Ethic Requirements (See Appendix B, page 208). The basic Ministry of Defence Research Ethics Committee ethical clearance was additionally cleared (refer to

Appendix B, also). It should be noted that as qualitative research is more open-ended and universal forms are empirically focussed; there was a significant delay in acquiring University clearance.

The supplementary information for ethical and security consideration are covered in more length in Appendix B, as stated above.

4.5.1 Informed Consent and Data Protection

As outlined in section 4.4.1.3, participants were recruited through voluntary registration which was sought with anonymous university/department email communication for civilian participants. Military participants were recruited using the collaborating partners own methods. During interview, participants were given participant information packs that concluded with a consent form. Participants were provided written information regarding consent in addition to verbal confirmation. The right to withdraw was stipulated, it explained to that participants did not require a reason for withdrawing, nor were there any consequences. Informed consent needs to be supported by the explicit assurance that confidentiality and anonymity will be provided for participants (Layder, 2012).

No participants requested to have their data rescinded in full, although some individuals requested demographic information to be redacted. This was offered for 12 months post-interview. This was as a result of published or presented data at conferences in that time and therefore was not possible to redact the anonymised information after this time. The raw data will be kept for 12 months post-completion of the thesis and subsequently destroyed along with other records. The transcription documents are additionally kept physically in a securely locked cabinet and kept digitally on a personal flash drive (and back-up external hard-drive). These documents are not kept on cloud services, apart from excerpts used in this document.

4.5.2 Personal Safeguarding

In the collection of the data, the ethical and security considerations for both participants and researcher were considered and mediated as per best practice. For example, there was small risk to military participants, as the narrative elicitation explored operational experiences. This meant it was necessary

to avoid or redact specific discourse for security or personal reasons. The participants were respected and not pressured into disclosure of information they were not comfortable with or unable to for security reasons.

Rooms were booked prior to conducting the observations to respect the site and disrupt the participants routine as little as possible. The wellbeing of the individuals was taken into concern with care to recommend rest breaks and beverages. In addition, potential power imbalances and exploitation of participants were considered in focus groups (Creswell, 2014).

The researcher entered a dialogue of closure to close each interview and focus group. The exit strategy included clear closing statements and opportunity to finish discussion on any topics discussed or have questions answered. The participants were thanked for their contribution to the study and an open invitation for further dialogue was offered (Morrison, et al., 2012).

Personal safety was also considered for the primary investigator as outlined by the Social Research Association (Social Research Association, 2017). For example, the dimensions of physical threat or abuse; psychological trauma; risk of accusation of improper behaviour or compromising situations; and others (Arksey & Knight, 1999). Further consideration needed to also be considered in conducting the research due to the physical disabilities of the researcher. Time management, appropriate pacing and travel were taken into consideration and mediated as best as possible.

4.5.3 Audio-visual Data Collection and Analysis

There are specific qualitative research ethical and security concerns when utilising audio-visual data. The audio-visual data of the document analysis, interviews and focus groups were collected using a camera and a Dictaphone, respectively. The Dictaphone was placed in view of the participants but out of direct eyeline as to not distract or cause unease. This was to allow for a more informal environment to foster easy rapport and conversation. As outlined in section 4.4.3, for transparency and authenticity, the advantages of using audio-visual data suggested by Silverman (2006, p. 350) are: 1) accessible to be withdrawn by participants; 2) can be replayed and transcripts improved; 3) preservation of sequences of talk.

Confidentiality and anonymity are a prominent concern with the type of data collected. All participants were therefore assigned numbers. Originally pseudonyms were going to be used to preserve humanness (Morrison, et al., 2012). However, due to the number of participants, this proved to be confusing in the data analysis summaries when referring to large numbers of corroborating quotes.

4.5.4 Research Background, Beliefs and Bias

To address the potential for bias, it has been proposed that bias in preconceptions exist when the researcher fails to appropriately address them (Malterud, 2001, p. 484). Approaching the study from different perspectives and paradigms are equally valid, and diversity in methodology can be used to develop a richer understanding of complex phenomena.

With this, the researcher has been associated with Defence for 15 years through placements, employment and sponsorship. Therefore, there is potential for preconceptions of military personnel through previous working relationships and experiences. These experiences are also interspersed with academic research into chronic illness discourse and volunteering in third sector consulting. The stance of the researcher seeks to abide by "empathetic neutrality" (Patton, 2005) – thus show interest, understanding to participants from a non-judgemental position. A reflexive journal (Lincoln & Guba, 1985) was kept to record any reflections of opinions during the data collection process. Furthermore, a research log was compiled with regular entries about any changes to methodological or logistics through the research study. This transparency can allow for audit, and assumptions in the data analysis can be revealed and contested through supervisory review.

The theoretical framework and philosophical orientation of the research has been explicitly stated in section 4.2.

4.6 SUMMARY

This chapter discusses the methodological choices and strategies utilised in this thesis. The research paradigm of interpretivism underpins the theoretical and philosophical standpoints of the research approach and design strategies.

The chosen qualitative research methods of IPA and HCA place explanation of social phenomena and obtaining rich, descriptive data at its core. This mixed methodology has been utilised to elicit dominant narratives and underlying expressions of trust, to answer the research questions. A diverse selection of civilian and military personnel were recruited through purposeful sampling. A total of 26 participants were interviewed, which included the exploratory study and the omitted data file. Twenty-four data sets were inductively analysed into master-, super- and subordinate themes and grouped semantically based off of existing trust literature. These are discussed in more detail in the following chapters.

As the validity of small sample size qualitative research can be met with criticism, this thesis sought to provide in-depth and robust methodological reasoning, in addition to concrete theoretical and philosophical underpinnings. As the research looks at socially constructed worldviews and culturally impacted heuristics, it was important to highlight the role of interpretivism and the guiding principles therein. Furthermore, that the axiology was identified, as the research explores the rich texture of small data sets and the nuances therein, rather than a focus on data saturation. To explore narratives and underlying expressions within the data, the ability to repeatedly immerse oneself in the data, and respond with acknowledge biases is imperative for transferable outcomes and situated recommendations. From this, all excerpts have been included with minimal edits and are kept as whole as possible, to retain context and the authenticity of participant voices.

5 STUDY 1: CASE STUDYAND DOCUMENTARY ANALYSIS

The research process for this thesis, as stated in the Methodology section (refer to Chapter 4; section 4.4) is based off responsive interviewing which adapts and fluctuates with the data collected and impressions developing with each interaction. This technique can be free-form and limitless in discussion possibilities, which with the already ill-defined subject matter of trust, imparts a challenge. Therefore, this chapter discusses two exploratory studies, a document analysis and a preliminary interview, which provide contextual grounding and appropriate interview boundaries.

The document analysis presents context of the socio-cultural idiosyncrasies within the military narrative of submariners through transcript analysis of audio-visual data provided from a restricted access speech. The exploratory study featured a participant situated in both military and civilian culture as a military spouse. This supplied insight on initial semi-structured interview questions, in addition to identifying potential conversational boundaries.

5.1 DOCUMENT ANALYSIS

5.1.1 Overview

The purpose of documentary - or document – analysis is to explore recorded data, such as memoranda, diaries and speeches and provide context bound findings about phenomena from an unobtrusive collection method. Primary sources, such as original documents, can provide a direct connection to the experience under investigation. This type of analysis is additionally informative for accessing temporally situated experiences (e.g. historical documents) or other restrictions. For example, Payne and Payne (2004) suggest that this technique can be used to investigate and identify themes which may have limitations or bias in physical sources (e.g. reporting bias in interviews). For methodological rigour, it is advantageous to explore the manifestation of actions or behaviours when participants are not observed or not situated in the investigation (Scott, 2014).

Documentary investigations were traditionally utilised by classical sociologists, with the contemporary prevalence of these methods are seen in historical and information science specialisms. In qualitative research, the current utilisation of document analysis is to use the resources to seek convergence and corroboration, supplementary to interviews or survey data (Mogalakwe, 2006). The goal of triangulation is to strengthen credibility through intersecting findings across the differing data sets and to reduce potential bias (Bowen, 2009; Wilson & Hutchinson, 1991).

For research that utilises mixed methodology, and especially those which interview using responsive interviewing techniques, Yanow (2007, p. 411) suggests:

> *"Documents can provide background information prior to designing the research project, for example prior to conducting interviews. They may corroborate observational and interview data, or they may refute them, in which case the researcher is 'armed' with evidence that can be used to clarify, or perhaps, to challenge what is being told, a role that the observational data may also play."*

Overall, the use of document analysis in this thesis is to provide triangulation and provide narrative context to contribute to the overall understanding of the military cultural perspective.

5.1.2 Data Analysis

The analysis for the document follows social constructionist inductive and interpretative techniques, in addition to the rhetorical pentagram (or pentacle) of speech analysis (Kjeldsen, 2004, p. 73). This is to specifically address the context in which to ground further analysis, more so than intensive interpretive phenomenological analysis. The rhetorical pentagram outlines areas of consideration for exploring audio-visual data such as speeches. As Figure 10 illustrates, the intentions of the artefact are impacted by: 1) the speaker; 2) the circumstances under which the speech is being delivered; 3) the audience; 4) the language rhetoric; and 5) the topic or overarching theme of the speech.

Figure 10 - Rhetorical Pentagram for Speech Analysis

5.1.2.1 Speaker

The speech was conducted by an OF-4 level individual (NATO, 2010) delivering an address on the operational context, both historically and going forward, of submariners in the Royal Navy and how performance of operations may change. Specifically, a key discourse of the speech involved their combat system engineering specialism and the portrayal of team kinship. This was presented to a collection of naval experts and industry professionals to forecast issues in the evolving battlespace, facilitate novel technologies and how to strengthen team cohesion.

5.1.2.2 Circumstances

The speech was recorded overtly with the permission of the individual for later distribution with their employer and therefore was both aware and consenting of the recording. The audio-visual data was subsequently transcribed. Therefore, there is a disconnection between the primary source collection and the secondary use of the data set, as a document analysis. This can reduce bias due to the artefact's original purpose is separate to the data analysis in the context of this thesis.

5.1.2.3 Audience

The primary source was performed in a conference setting in the UK to naval experts and industry professionals (as mentioned previously) concerned with HM Armed Forces. The primary receivers of the communication were representative and contextual to the address. For example, those of working age and of appropriate management level to implement change and facilitation. Additionally, those in attendance were of appropriate education level, security clearance and nationality to receive the information delivered. Furthermore, due to the primary recording security restrictions, the acquisition of the privately accessed data necessitates the researcher to also have appropriate security clearance, expertise and nationality to analyse the data.

5.1.2.4 Language

The speech follows a semi-formal structure, in that there is rigidity in the delivery due to the type of personnel in the original audience. However, it also contains colloquial anecdotes and sincere emotive delivery when addressing the familial kinship of teammates.

The narrative of the speech begins with a formal introduction, followed by factual context of the history of submarines. This was proceeded by anecdotes of lived experiences during maritime operations were delivered and how technological modernisation has impacted their service. This introduced the topic of emerging technology and automation, which provided a platform to pedestal the goal of the speech, which was the discussion of the benefits of improved combat system development.

The tone of the speech, as introduced previously, had factual aspects. There are moments of brevity and humour to lighten the somewhat serious tone. Primarily, there is a sense of 'otherness' and 'in-group versus out-group' language and subsequent connotations throughout. The language uses opposites and parallels to double-down on the othered narrative. The speech also utilises significant rhetorical questions to frame discussions. The individual furthermore uses connotative language to prime arguments and often uses figurative language and imagery to simplify concepts which some of the audience may not have experienced (e.g. life aboard a submarine). There are aspects of perceived superiority and hauteur, in the delivery of successes and continued excellence within the domain.

Overall, there is passionate delivery of their experiences aboard operational submarines and when discussing teammates. There is pride in both the act of service and team cohesion. The emotive delivery and the expressions of 'otherness' plays a significant role in the storytelling when providing context to the social climate of submariners.

5.1.2.5 Topic

The themes explored in the speech are covered more extensively in the analysis (Section 5.1.3). The speaker builds their argumentation ethos as a subject matter expert (SME) and a serviceperson to lend credibility to their expert testimony (Toulmin, 2003). As discussed previously, the pathos of their emotive language in their delivery is furthermore reinforced by voice inflections on key clauses, physical gestures and eye contact maintained with the audience. This hypotaxis and example constructs within the speech are also discussed further in Section 5.1.3.

5.1.2.6 Intention

The overall intention was to be both informative and celebratory of the unique capability of the Subsurface Navy and the unique team cohesion the environment fosters. Furthermore, it serves to inform individuals outside of the social climate of some of the idiosyncrasies within this echelon of the military, and how the strong team cohesion can be exploited beneficially for facilitating new and emerging team dynamics. It serves to convince the audience that the competencies and proficiencies within the subsurface domain are to be celebrated and explored further. It is unknown whether the goal was to secure avenues of funding or to emphasise and commend the members of the associated organisation or holon.

5.1.3 Document Analysis

The main expressions of the argument, or speech, were concerning the 'otherness' of submariners compared to their colleagues. This in part is connected to the backing traits in clauses surrounding team cohesion. In addition, reluctancies to adapt were discussed, both with social schemata as well as technological facilitation.

5.1.3.1 Otherness

The speech was introduced with an emphasis on the difference, or perceived distance or isolation, the Subsurface Navy has in comparison to their colleagues with this statement:

> *"At the turn of the century, the Royal Navy introduced its first submarine, the Holland One, and at the time, an admiral in the admiralty famously quoted, that he considered submarines to be "underhanded, unfair, and damned un-English", and you know what? I think he was probably right, sure a whole load of things have changed in the hundred years since we introduced that platform, submarines are bigger, more powerful, more capable than they ever were before"*

This opening puts an emotional hypotaxis on the atmosphere within the audience venue by introducing categorisation through self-depreciation, framing the holon as the 'underdog' with the persecutory comment despite the competitive advantage. Through use of this narrative, the speaker can manufacturer sympathetic identification with the group (Schuman & Harding, 1963). Furthermore, as industry personnel were members of the audience, there may be grounds that this framing may have been used as a marketing stratagem. Studies have shown this branding can increase purchase intentions for consumers (Paharia, et al., 2011).

Following this statement, there were also expressions of hostility or judgement to the out-groups in comparison to their in-group:

> *"…we believe ourselves to be better than the rest of the Royal Navy, not just because we know we are, but because we work harder, we play harder, and we deliver more punch than our general service colleagues."*

The underlying expression can be borne of both direct competition with their military colleagues, especially when monetary resources are concerned which can foster hostility, regardless of active or passive forms (Tajfel, 1979; Ministry of Defence, 2019). There is also a secondary pedagogy regarding 'Us versus Them' mentalities. This concerns the connections between identity perception, self-esteem

and group membership (Turner, et al., 1987). Some have additionally argued that important group membership provides additional psychological resources which reinforce the connection to the social group (Jetten, et al., 2015).

Other connections to themes associated with Social Identity Theory are emphasised in the ethnocentric beliefs underlying the following excerpt:

> *"We like to do things informally, because we believe ourselves to be, immeasurably professional and we're intensely proud of that. We also rely absolutely on our comradery which is born ultimately of intense proximity to one another, and on operations, everything is focused, absolutely, on delivering the mission, submariners tend to change their tone when we go on ops, so whereas we might be fairly jovial in normal circumstances, when we get serious, things become serious really really quickly"*

Another perspective is that the enculturation of the military hierarchy, strict regulatory environment, and sole focus on mission success develop the loyalty to the in-group through close relationship bonds with other associated members:

> *"We submariners are strange beasts, it's definitely not for everyone, and certainly not for the faint of spirit, but actually, it is a fantastic life, Churchill famously said that there were "none who faced grimmer perils, or showed more devotion than submariners"*

Similar to the framing of the initial introduction to the speech, ethnocentrism can be associated when referencing out-groups and the 'black sheep effect' (Marques & Paez, 1994). A common theme in the speech oft refers to comparable out-group members negatively and judge ingroup members more favourably. The preservation of social identity and ingroup uniformity can provide some substructure to attitudes towards team cohesion.

5.1.3.2 Team Cohesion

> *"You are absolutely unlikely to meet a group of people so committed, or a team so cohesive, as submariners"*

> *"…you won't see the sun for months on end. You'll work, six hours on, six hours off, seven days a week, without a break. You won't shower, or at least you will rarely shower, which won't matter, because nobody else will shower either, you'll all stink together and so none of you will notice until you get home*

It is well known that shared experiences are influential predictors for interpersonal attraction (Pinel, et al., 2006). However, recent literature has begun to explore shared adverse experiences and team cohesion through promoting group bonding through social support (Bastian, et al., 2018). Furthermore, negative experiences from an exogenous source can promote a heightened focus on fostering in-group social integration, support and solidarity (Knight & Eisenkraft, 2015). Together with physical proximity fostering traditional methods of team bonding, the unique high criticality work and living space that is a submarine can be instrumental in the development of intrapersonal relationships.

5.1.3.3 Adaptation

> *"…the reality is that in the heart of every submariner is a pirate; submarines might change, but submariners don't…"*

> *"You know what? I suspect a submariner that stood in front of you twenty years ago would give you exactly the same message, and no doubt, my successor's and their successors in twenty years' time would tell you exactly the same story. Submarines change, but submariners don't."*

Resistance to change has roots in many different sources (Pardo del Val & Martínez Fuentes, 2003). From the excerpts included here, there may be organisation connection to zeitgeist or the perpetuation of the scheme despite situational changes (Barr, et al., 1992). Whether this is positioned to connections to tradition or resistance to change in routine. Other internal factors may be that of maintenance of

sameness in behaviour, devotion or loyalty to the group or other working models of self (Eagle, 1999). The rigidity and authoritarian culture of the military and reliance on routine can also be a justification for resistance to adaption. For example, the organisational structure of the military and the non-disclosure of information or maintenance of confidentiality may cultivate 'organisational silence' through the purposeful withholding of information. Research has identified that this dissimulation can create the shared perception that speaking up, or going against the status quo, is unwise (Morrison & Milliken, 2000). Therefore, this fear of the unknown through systemic silence can have negative consequences on the organisations ability to adapt to change.

5.1.3.4 Technology

Technology was a topic covered throughout the speech, which took on a variety of forms. In relation to the previous sections, aspects of technology forming critical points in the social actor network within the sociotechnical system:

> *"The captain only gets the information he needs if the system is working, if every element of the system is working from the code fragments that make up the systems, from the cabling that runs throughout the boat, from the racking, from the bolts that connect the hydrophone, to the face of the submarine, every element is absolutely essential, and so no matter what anybody in the room does, each and every one of you are critical in enabling the submarine service to remain safe and allowing us to remain the best."*

An interesting point from this segment is the variety in which technology, and automation to a point, forms the team framework. For example, pointing out that the technological system is both a mechanical physical item, as well as an internal programmed actant. However, there was emphasis on the mechanical aspect, thus defining separation of human versus technology-based teammates. Viewing automation through a mechanical lens highlights the potential for under-reliance (Bousquet, 2008) or, that the system teammates are framed as tools only and subservient to the operator (Reeves & Nass, 1996).

One of the more telling parts of the speech with regards to human-system interaction from both a teammate integration point of view, and an authoritative hierarchical societal structure, is in the following:

> *"…You are aware no doubt that Artful has recently gone to sea, which is a massive victory for the submarine service, you probably know, some of you will know, that the ship's company were initially very nervous, about CCS and how it would perform, submariners are traditionally nervous of all change, it's a healthy response, but actually, you guys have done a fantastic job of persuading them, they are hugely flattering about the system, they recognise that it is, significantly capable, massively more robust than ever before, than they ever imagined…"*

The resistance to change seems to be an underlying expression which is embedded within the submariner culture. In addition to the comment regarding '*it's a healthy response'* attempts to justify the opposition, which is somewhat juxtaposed with the persuasive language. This can be interpreted as both genuine and sceptical. From a superficial standpoint, the speaker is actively addressing the resistance and encouraging the audience that despite hesitance, there was appropriate facilitation and interaction with the automated systems actants. However, the strategic '*but actually'* is crucial in interpreting this clause. The term 'but' has a propensity to subconsciously negate, cancel or distil positive statement and can be covert contradictions of the initial statement (Booher, 2015). This therefore can infer that although facilitation and use was accomplished and positive experiences were had, there may still remain underlying resistance and subsequent boundary to full integration with human teammates.

5.1.4 Summary

The document analysis provides a fundamental starting point to ground some of the narratives which may arise or develop from the military cohort interviews and focus groups. It also serves to bolster themes originating within the literature regarding institutionalised schema priming and the importance of team cohesion on shared social identities within the military. Especially in the submariner context.

This excerpt provides insight into the lived experiences and cultural quirks of the subsurface navy which can be explored further and compared against in future chapters. It also provides themes which can be specifically considered when interpreting the larger military cohort data sets. For example, the underlying expressions of team cohesions and how that is sculpted by the hierarchical authoritative culture. In addition, how this may differ across the different echelons of the military with the types of experiences they have and the macroergonomics of the sociotechnical systems they reside in.

5.2 CASE STUDY

5.2.1 Introduction

The characteristics of the document analysis and the exploratory study provides a spotlight on the context of the cases (Robson, 2002), allowing the study of '*real people in real situations'* (Cohen, et al., 2011, p. 289). For the exploratory study, the key protocol for this section was to refine and explore boundaries for the main studies in this thesis. The research process, as stated in the Methodology section (refer to Chapter 4; section 4.4), is based on responsive interviewing which adapts and fluctuates with the data collected and impressions developing with each interaction. Therefore, to effectively use the time wisely and craft appropriate questions, a preliminary case study was conducted.

This section explored the methodology, the analysis protocol and the potential areas which it would be advantageous to bound unnecessary (or fruitless) tangents which are unlikely to produce appropriate conclusions. The excerpts in the interview will not be explored in depth in this section, but examples will be used to provide clarification to the coding classification going forward. It should be noted that the final table used in this section underwent restructuring for clarification after the civilian interviews were performed. For example, the themes changed in terminology for clarification or were restructured to fit under a hierarchy of master, super- and subordinate themes more concisely. The multi-stage coding process and method are explored in detail in this section.

5.2.2 Methodological Approach

The technique to exploring qualitative data sets is diverse. For this thesis, as outlined in the Methodology chapter (chapter 4, section 4.4), a procedural and interpretative process for coding was utilised. In addition to the protocol used, the terminology of coding is differentiated and clarified compared to computer science and programming language in this section.

5.2.2.1 Defining Coding in Qualitative Research

Coding is a decision-making process to make sense of large data sets (Elliott, 2018; Creswell, 2015, p. 152). In qualitative data analysis, coding is a technique which a word or phrase represents a unit. These units, or codes (or nodes within the Computer Assisted Qualitative Data Analysis Software (CAQDAS), NVIVO (QSR International Pty Ltd, 2012)) comprise of words, phrases, paragraphs or sections of transcribed text expressive of themes. Saldaña (2009, p. 3) defines codes as the following:

> *"…a code is most often a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of language based or visual data."*

A descriptive code seeks to summarise an interpreted topic and condense the data, though not as a reductionist act. The process of coding is to reflect on the core meaning, to decode the data, and subsequently appropriately encode when labelled into themes and categories. The term 'theme' and 'category' are often used interchangeably. Themes in this thesis refer to the analytic reflection of the data set, and coded categories classify codes at higher levels, the aggregation of codes or groupings of common ideas (Richards, 2015). Alternatively, categories are explicit descriptors whereas themes are derived from subtle tacit processes (Rossman & Rallis, 2016).

It should be noted that coding is not synonymous with analysis (Basit, 2003, p. 145). As previously mentioned, coding is a problem-solving action and the foundation on which meticulous and interpretative analysis can be built upon. As Gbrich (2007, p. 21) states, coding is the sequence of segregating, grouping, regrouping and relinking data so as to consolidate and extract meaning.

5.2.3 Levels of Coding

This section explores the primary and secondary types of processes of coding, and the specific approaches which are utilised in the analysis protocol. In addition, the coding scheme which forms the basis of the super- and subordinate categories and overall themes explored in the main research.

5.2.3.1 Open Coding

Open coding is the first level of coding in which distinct concepts of categories are explored. These concepts are often master headings which second level categories can explore in more depth. There are three specific styles of open coding which are utilised in the analysis protocol. These are as follows:

i. Descriptive Coding: – This method of topic coding summarises the data unit or units into a short word or phrase. The topic reflects the essence of what is discussed or written in the content. It should be noted that the code is not an abbreviation of the content but an identifier (Tesch, 1990, p. 119).

ii. Values Coding: – This application of coding reflects the values, attitudes and beliefs of the participant and outlines their perspectives or worldview. Though each of these concepts or constructs are different linguistically and underlying meaning, values coding subsumes all three (Saldaña, 2009, p. 89).

iii. Magnitude Coding: – This type of coding can be qualitative and quantitative. This method enhances description through providing figures which can indicate intensity, frequency and other statistical evaluative content (Miles & Huberman, 1994). The data extrapolated is neutral in as it only provides nominal data. For this thesis, a version of magnitude coding was developed to explore the salience of commonly occurring codes and to appraise priority and risk in concerns registered (Field, 2019). Providing numerical values provides an indication for potentially important codes, however as suggested by Saldaña (2016, p. 41) *"Counting is easy; thinking is hard work"* and that frequency does not necessarily delineate significance. The use of mixed methods in this thesis is a function of communicating the scope of themes visually and simplified.

5.2.3.2 Axial Coding

In open coding, the focus is primarily to define categories from the data set. The second level axial coding uses these concepts when re-immersed in the artefacts so as to confirm whether they accurately interpret themes and to explore how they are related. Axial coding is more directed and focussed on exhuming and identifying important aspects. For example, focus coding explores the most frequented codes and is used to create sharp boundaries through data saturation. Pattern coding sews together the material in the open coding into one meta-code and this unit is often the master theme explored in the analysis (Miles & Huberman, 1994, p. 69). This also explores simultaneous coding where two or more codes can be applied to an artefact.

The coding procedure in both primary and secondary coding aids in the final coding scheme and concerns register. This allows for organising and tabulating the valuable aspects of the large data sets in a concise and easily digestible manner.

5.2.4 Coding Scheme

This section details the coding scheme which was developed using the exploratory study as a foundation through providing preliminary first level coding (see Table 19, Table 20, Table 21 and Table 22). The scheme was manipulated, modified and expanded on throughout the later studies.

This iterative process of understanding the data and coding, recoding and refinement aids in the revalidation of the coding scheme. However, as per some other researchers, there can be theoretical framework concerns in trying to identify the 'right' codes. Spencer, Ritchie, Ormston, O'Connor, and Barnard argue that "*the aim is not to produce a perfectly consistently coded set"* because *"labelling is done to manage data rather than to facilitate enumeration"* (2014, p. 278). Furthermore, others have noted caution in the use of inter-coder reliability, such as statistically discerning consistency between two coders. This is due to different understanding of the data, from potentially opposing disciplines which may impact the coding filter. This is the lens with which researchers may interpret the data and may cause conflict or bias.

5.2.4.1 Culture

Table 19 - Coding Scheme (Cultural Category)

Category	**Master Theme**	**Superordinate / Subordinate Themes**	**Code Description**	***Example***
Cultural			This major category encompasses all child units which discuss or reflect exogenous socio-cultural aspects.	*N/A*
	Culture		The code is specific to participant attributes, reflections, attitudes or behaviours moreso than social heuristics (see Scheme Priming)	*N/A*
		Identity	Code reflecting introspection of self. This can be identities related to social groups. This is a superordinate category to the following.	*so, I think with each generation it is different isn't it – each generation gets long in the tooth and it is shifting the views isn't it.*
		Age	A subordinate code to Identity. Age was included as a separate node due to the interest in age in the literature to technology resistance and adaptation.	*I think technology can always be improved but there is always going to be error isn't there, something could always go wrong and I think people of*

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
				my generation would see that more than maybe younger generations
		Technology Acceptance	A superordinate node to the following two subthemes. Technology acceptance explores reflections on technology resistance, adaptation and perceived use.	*terrifies me the fact that a car could be roaming around our streets with no driver in – my god, scary*
		Media Influence	The effect of media as a cognitive primer and the socio-cultural influence of entertainment.	*[automation is] very interesting, quite fascinating because I wasn't brought up with anything like that so a bit science fiction to me*
		Zeitgeist	Schema or mental models that have been shaped by the current or historical defining beliefs of that culture. Examples include introspected generational differences and cultural shifts relation to technological improvements.	*I know if I was driving down the motorway, I'd look and look again because it is not commonplace, you don't expect to see it so I think it could be, until it became sort of the norm almost, I think it would be very distracting. Things are changing all of the time aren't they – strange. I*

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
				suppose from when I was a child to now, technology itself is incredible what is has done.
	Terminology		Participant definitions of the core concepts explored in the research.	*N/A*
		Automation	How automation is defined by the participants. For example, the level of sophistication and what they deem automation, automated and/or robotic.	*In everyday life? Urm, when you see things maybe like robots that can take on tasks that humans would have done, for instance in factories, vacuuming your floor; that sounds a good one, erm yes just (00:00:55) doing tasks that humans would have done so things like in factories which I think are good but I do have some reservations as well*
		Chaoplexic Definitions	Nodes reflecting on Chaoplexic definitions (Bousquet, 2008): e.g. Mechanistic (calculator); Thermodynamic (Engine/Car);	*if I get stuck I would then go to the calculator so yes I suppose I do trust a calculator yes. My brain doesn't compute with that being automation*

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
			Cybernetic (driverless car); Chaoplexity (Internet).	
		Trust	How trust is defined by participants. For example, whether trust and reliance are synonymous; or if trust between human and artificial actors different.	*I: [8]Do you think that your feelings towards automation like being able to trust things to work correctly will change as technology improves?* *R: Yes I hope so, the proof is in the pudding as they say, so yes. Yeah.*
	Schema Priming		Related to the node Culture, the coded artefacts in this node reflect experiences and reflections on whether this would impact future interactions or experiences.	*Yes I think that that is human nature – if someone has a good experience, you are almost boyed with it and if someone has a bad experience it makes you fearful doesn't it so yes I think definitely, definitely, yeah that does, the experience of others would rub off on yourself yes.*

[8] I = Acronym for Interviewer / R = Acronym for Responder

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
		Contradictions	Excerpts which conflict with other statements or attitudes presented within the verbal protocol. Or exceptions due to perceived automation (or lack thereof) or capability.	*I don't know it is a very strange thing isn't it. I think driverless cars they do concern me but pilots, no not as much because they are sitting there at the front aren't they. I think the driverless car could be distracting*
		Negative Experiences	Negative experiences with using automation.	*Yes I do – Hotpoint tumble driers, I won't touch them. Yeah, yeah. That fridge freezer that set off Grenfell Tower also Hotpoint; I wouldn't have Hotpoint in the house now as you are just fearful aren't you*
		Positive Experiences	Positive experiences with using automation.	*Yes because it has made my life easier. When I was doing my first lot of studying we hand wrote all of our assignments so we hand wrote 3000 word assignments and if you wanted to pop a paragraph in the middle, you'd have to draw big arrows and then put it*

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
				at the end with asterisks and that, whereas now you just put your cursor there don't you and write your paragraph so it definitely, in academia makes life a lot lot easier I think in that way.

5.2.4.2 Interpersonal

Table 20 - Coding Scheme (Interpersonal Category)

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
Interpersonal			This major category encompasses all child units which discuss or reflect endogenous socio-cultural aspects.	*N/A*
	Self-Determination Factors		Codes exploring intrinsic psychological requirements to develop personality and motivational factors.	*N/A*

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
		Autonomy	Superordinate theme to Hubris; Ownership and Authorship; Personal Agency; and Role of Operator.	*N/A*
		Competence	Superordinate code to Risk; System Error; and their child nodes.	*N/A*
		Relatedness	Superordinate code to Organisational Responsibility and Sociocultural Scheme.	*N/A*
	Autonomy		Superordinate code exploring themes on personal autonomy.	*N/A*
		Hubris	Themes expressing topics which on servant leadership of human agents with artificial actants. For example, that human can exceed decision making software, despite physical and cognitive limitations.	*If something went wrong there is no human in there to take over I think, so in the ideal world nothing would go wrong but if it went wrong, something failed in the system, something failed and it couldn't stop or hadn't been told to stop and kept going on and ploughing into other cars*

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
		Ownership and Authorship	Codes reflecting clauses on fluid and ill-defined boundaries on ownership (such as digital downloads) or authorship (artificial intelligence).	*I'd much rather laden myself down with half a dozen books*
		Personal Agency	Codes exploring threats to personal agency, choice and identity.	*Well then you are losing yet another skill and that is the skill to drive isn't it, and if people can't drive because they are reliant on jumping in their car and telling it where to go maybe and sitting in the back seat reading the paper on the way to work*
		Role of Operator	Codes exploring theme on how the participant feels about person-in/out/on-the loop in human-system integration/interactions.	*I just touch the brake twice and the cruise control cancels itself out, so I am in control whereas a driverless car. Wonder how they do work, how do they control their speed you know if you are on a motorway and it goes down to 30, how does that happen – inbuilt computer is it?*

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
	Competence		Superordinate code exploring the ability of the actants within sociotechnical systems.	*N/A*
		Risk	Codes reflecting on risk appetite, behaviours or perceived risk allowed by actants in a sociotechnical environment.	*I suppose if there was a little robot pushing your vacuum around that is not quite as tragic if something went wrong unless they bashed into your best furniture but so it depends really what it was – a robot pushing a vacuum around or a driverless car, there's a big difference isn't there so I think I'd have to look at each individual one I think, separately*
		System Error (H-S)	Similar to the Risk code, themes explore error in H-S interactions with expressions of blame. This includes the child codes of human error and artificial actant only.	*using things like computers, anything like that, you can't blame the computer you have to blame the person that has either programmed it or operating it don't you – as I've learnt to my cost when I've lost lots of things that I've*

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
				needed! [laughter] 'Can't blame the computer, you've done something'
		Human Error	Clauses exploring attitudes on human error and attribution of blame.	*because a human if they're driving a car and everyone in front of you is suddenly grinding to a halt because something has happened, most humans would be able to, well hopefully, would be able to apply the brake and veer out of the way*
		System Error (Technological Actant Only)	Clauses exploring attitudes on artificial actants and attribution of blame.	*I think technology can always be improved but there is always going to be error isn't there*
	Relatedness		Interpersonal connections and the development of relationships and their effect on integration.	*N/A*
		Organisational Responsibility	Themes which explore the relationships in the sociotechnical system. For example, the negative attitudes to artificial agents could be	*My brother-in-law's a farmer and combine harvesters, when he was first combining, he would get in and drive a combine harvester and it was a hell of*

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
			better integrated with leadership support or responsibility. Themes which explore accountability in H-S groups.	*a skill to go up and down the fields and turn it and do whatever. Now he gets in, they programme it and then they sit back and eat their cheese and pickle sandwiches – he says it does it for them so; think of the size of a combine harvester but there is only that in a field isn't there*
		Sociocultural Schemata	Codes exploring societal pressure from interpersonal relationships and structures, and how they impact the formation of schemata.	*I: Do you think that everyone in your office feels that way?* *R: Well there is only 3 of us there and yes all of us yes, we all say the same – no one says to the other one don't be so daft, no we are all wary, all very very wary.* *I: What do you think would happen if one of you changed your mind?*

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
				R: We'd ignore them [laughs] we'd batter them down. I don't know, I don't know.

5.2.4.3 Automation

Table 21 - Coding Scheme (Automation Category)

Category	**Master Theme**	**Superordinate / Subordinate Themes**	**Code Description**	***Example***
Automation			This major category encompasses all child units which explore participant reflections on automation.	*N/A*
	Automation		As above.	*N/A*
		Anthropomorphisation Only	Subordinate codes which reflect on anthropomorphising of artificial actants.	*Satnav Sally was not cooperating, and we just didn't know where we were, well what do we do now, who do we phone*
		Perceived Permanency Only	Subordinate codes that discuss the fluid boundaries of digital and artificial actants. For example, issues around the lack of physical presence with system teammates.	*…tactile; I like to feel a book, I like to sit there and hold a book. I've not yet moved on. I do use my kindle, but I wouldn't dream of taking a kindle on holiday and sitting on a sun lounger for two weeks with a kindle – I'd much rather laden myself down with half a dozen books*

5.2.4.4 Trust

Table 22 - Coding Scheme (Trust Category)

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
Trust			This major category encompasses all child units which explore participant reflections on trust.	*N/A*
	Trust		As above.	*I: Do you think that your feelings towards automation like being able to trust things to work correctly will change as technology improves?* *R: Yes I hope so, the proof is in the pudding as they say, so yes. Yeah*
		Capability	Similar to competence – artefacts in this code explore how trust is linked to perceived and/or actual capability of system teammate.	*[The following three nodes were created in the iterations post exploratory study]*
		Decision Making	Themes which explore how trust attitudes, behaviours and beliefs have	*N/A*

Category	Master Theme	Superordinate / Subordinate Themes	Code Description	*Example*
			on naturalistic decision making. These can also include themes of expertise, gut-feeling and other related cognitions.	
		Reliability	Trust is often used interchangeably with reliance. Exploration of this concept are coded at this node.	*I: Do you think that maybe it is the manufacturer pedigree has an influence on what you think about it?* *R: Hmmm yes I do; top of the range and things because what you get for your money don't you so yes and they have different standards of cars don't they*

Table 23 - Thematic Coding with Content Weighting and Priority Weighting

Category / Theme / Sub-Theme	Consequence	Salient Themes: Absolute	Salient Themes: Relative	Content Weighting	Priority
Cultural		**51.3%**	**27.9%**		
All Culture		13.0%	7.1%		
Culture Only	1	3.2%	1.8%	6%	Low
Identity	2	0.0%	0.0%	0%	Low
Age	2	3.2%	1.8%	6%	Low
Technology Acceptance	3	1.6%	0.9%	3%	Medium
Media Influence	1	1.2%	0.7%	2%	Low
Zeitgeist	1	3.8%	2.0%	7%	Low
All Terminology		11.3%	6.2%		
Terminology Only	1	0.0%	0.0%	0%	Low
Automation	1	2.3%	1.3%	5%	Low
Chaoplexic Definitions	1	9.0%	4.9%	18%	Low
Trust	1	0.0%	0.0%	0%	Low
All Schema Priming		26.9%	14.6%		
Schema Priming Only	1	0.0%	0.0%	0%	Low
Contraditions	2	1.6%	0.9%	3%	Low
Negative Experiences	2	25.2%	13.7%	49%	Medium
Positive Experiences	1	0.0%	0.0%	0%	Low
Interpersonal		**104.1%**	**56.7%**		
All Self-Determination		1.7%	0.9%		
Self-Determination Only	2	0.0%	0.0%	0%	Low
Autonomy	2	1.7%	0.9%	3%	Low
Competence	2	0.0%	0.0%	0%	Low
Relatedness	2	0.0%	0.0%	0%	Low
All Autonomy		67.5%	36.8%		
Autonomy Only	2	1.7%	0.9%	3%	Low
Hubris	2	5.0%	2.7%	10%	Low
Ownership and Authorship	3	1.9%	1.0%	4%	Medium
Personal Agency	3	10.3%	5.6%	20%	Medium
Role of Operator	3	48.6%	26.5%	95%	High
All Compentence		25.0%	13.6%		
Competence Only	2	0.0%	0.0%	0%	Low
Risk	4	5.4%	2.9%	10%	Medium
System Error	5	4.1%	2.2%	8%	High
Human Error	5	11.3%	6.2%	22%	High
System Error	5	4.3%	2.3%	8%	High
All Relatedness		9.9%	5.4%		
Relatedness Only	2	0.0%	0.0%	0%	Low
Organisational Responsibility	4	4.0%	2.2%	8%	Medium
Sociocultural Schemata	2	5.9%	3.2%	12%	Low
Automation		**8.7%**	**4.7%**		
All Automation		0.0%	0.0%		
Automation Only	2	0.0%	0.0%	0%	Low
All Antropomorphisation		4.0%	2.2%		
Antropomorphisation Only	2	4.0%	2.2%	8%	Low
All Perceived Permanency		4.8%	2.6%		
Perceived Permanency Only	1	4.8%	2.6%	9%	Low
Trust		**19.5%**	**10.6%**		
All Trust		19.5%	10.6%		
Trust Only	3	0.0%	0.0%	0%	Medium
Capability	4	0.0%	0.0%	0%	Medium
Decision Making	5	0.0%	0.0%	0%	High
Reliability	4	19.5%	10.6%	38%	High

5.2.5 Analysis and Summary

The development process of the coding scheme was iterative and was modified over the course of the data collection and analysis. The purpose of this section was to research the possible themes and create a coding scheme to develop and organise topics to bring to the later stages. Furthermore, as shown in Table 23, the salient themes and calculation of priority weighting were conducted, it provided areas and important themes to explore in the later interviews. The exploratory study allowed for areas of interest to be underscored and highlight salient themes to explore in more depth in subsequent interviews. The participant had knowledge of both socio-technical groups without subject matter expertise in either the military or STEM. This provided a beneficial case study to explore surface level themes as well as indicate potential areas where unprofitable tangents could arise.

The case study also provided a phase where methods of abductive data analysis, inductive sensemaking and deductive grouping could be refined. This enhancement of methodology and techniques, provided an enriched protocol going forward in the thesis.

5.3 SUMMARY

The focus of this chapter was to concentrate the context to provide a descriptive narrative of the phenomena studied in this thesis and to provide a solid foundation for the main studies to reflect upon. Furthermore, the triangulation of these cases with the wider scope presented later in this thesis will provide understanding of the phenomenon and enlighten areas of interest not illuminated in the literature review (Chapter 3).

The cases covered in this chapter are sources to provide supplementary information and contextual material to support the main sources of data and analysis provided by the interviews and focus groups in the chapters to follow. Additionally, the exploratory study aided in the refinement of the quantitative data presentation. The research benefits from these data sets through re-examination and re-consideration of themes, categories and evolving narratives.

5.3.1 Summary of Concept Definitions

5.3.1.1 Trust

The definition of trust is explored at length in Chapter 2, as well as the nuances of how these theories impact the trust expressions of the participants in these studies. For clarification going forward, the concept of trust as grounded and bounded by the thesis so far, relates to the interpersonal exchanges between two actants. The nuances of more philosophical and intangible aspects of trust, such as faith, are not explored in the thesis, nor where explored in the responsive interviewing. Furthermore, using information from Study 1, this provided contextual boundaries regarding how trust and reliance are intertwined. For example, interview questions explored the themes through dissecting how the participants felt about the two terms, how the differ and how they would define them themselves. The participant conceptualisations of trust are explored more at length in the proceeding chapters.

5.3.1.2 Automation

A brief summary of the different ways automation can be defined is outlined in Section 1.2.1. For the responsive interviewing, the concepts in Chaoplexic Warfare (Bousquet, 2008) were used to frame different types of automation to the participants. This was to provide them a scaffold to explore their own contextualisation's of what they believe is automation. Although not illustrated in depth within this thesis, the interview provided participants the space to explore these themes and how different levels of complexity impacted their expressions of trust and facilitation. The scope of automation covers simple machinery, to robotics and artificial intelligence. Each of these items, as representative of Bousquet's classifications, were provided and discussed with all participants.

6 STUDY 2: CIVILIAN INTERVIEWS

6.1 INTRODUCTION

The purpose of this chapter is to identify aspects of the civilian experience of trust, definitions of trust, opinions of automation and attitudes to increasing levels of automation and emerging technology. This exploration of the narratives will focus on the first research question (RQ1: How do the narratives that people tell about automation relate to their trust in technology?). In outlining the attitudes of civilians, this will provide a basis for reflecting on the second research question (RQ2: How does trust in automation differ between military personnel and civilians?) following the second study (see Chapter 7). The methodology chapter (refer to 4) provides detailed clarification of the techniques used to obtain data in this and the following chapter (Chapter 5). The methodology utilised the case study interview to refine the conversation structure, highlight potential misinterpretations and allow for improved recovery strategies for the following responsive interviews. This bounding allowed for a greater focus on the topics in the semi-structured interviews without impeding freedom of expression and healthy tangents.

This chapter reflects on participant views and the dominant narratives and expressions of trust and attitudes towards automation communicated therein from the Civilian Subject Matter Experts (SMEs). This chapter expands on trust acceptance, the perceived concerns of more ubiquitous automation and the nature of trust and how that interplays with their position on facilitation and integration.

The distinguishing differences, focal points and thematic analysis are presented and summarised. The findings of this chapter provide supplementary comparison points and probes in the Military analysis in the following chapter (Chapter 7).

6.1.1 Participant Information

Recruitment and additional information regarding anonymity and safeguarding for participants are outlined in Section 4.4.1.3 (Interview and Focus Group Design), Table 24 illustrates the demographic information for the Civilian cohort. To maintain privacy, this table does not attribute the participant with the corresponding number used in the findings of this chapter. The age distribution ranged from early 20's (on the cusp of Generation Z and Millennial (Generation Y)) to late 40's (Generation X) which provided a wide range of digital competency and experiences. The participants all had a high level of education, which was predominantly due to the recruitment strategy of the study. As will be discussed in Chapter 9, Section 9.5, experiences of more lay persons may have provided more texture to the civilian narratives. As noted below, one of the participants listed in this table has not been included in this chapter due to a loss of data. Additional demographics which are not included in the table for privacy, were those of gender and ethnicity. The cohort included both male and female participants, though more androcentricly skewed. This is in part due to the gender diversity in STEM subjects (Beede, et al., 2011). Similarly, BAME (Black, Asian, and minority ethnic) participants were included but under-represented.

Table 24 – Participant Demographic Information for Civilian Cohort

Domain	Expertise	Level of Education or Equivalent[9]	Age Range
Civilian Military	Education	Level 7	30+
Civilian	Computer Science	Level 8	30+
Civilian[10]	Civil Engineering	Level 7	18-30
Civilian	Social Science	Level 8	18-30
Civilian	Computer Science	Level 7	18-30
Civilian	Physics	Level 8	18-30
Civilian	Electrical Engineering	Level 7	18-30

[9] As defined by The Quality Assurance Agency for Higher Education (The Quality Assurance Agency for Higher Education, 2014)
[10] Participant omitted due to loss of data

6.2 FINDINGS

6.2.1 The Conceptualisation of Trust

As discussed in Chapter 2 Section 2.1.1, the concept of trust is defined differently across disciplines (Hall & McQuay, 2010) due to the breadth of applications of trust research, but also the lens in which specialists are applying their classifications (Colquitt, et al., 2007). A fundamental narrative to clarify in the civilian's interviews, were how they define terminology and what is meant by trust and subsequently how the development of trust was influenced.

A recurring point of contention in the literature is whether the term trust can be used interchangeably with reliance (Lee & See, 2004). Many researchers attempt to classify reliance as a quantifiable behavioural measure of trust, and thus conversely define trust as fluid, subjective, internal process. In half of the Civilian interviews, nodes coded terminology expressed trust synonymously with reliance. Statements such as *"I think the more reliable something is the more you can trust it"* is echoed throughout the civilian interviews, especially when considered in terms of automation. For example:

> *"But isn't being able to rely on the system a goal of automation…" [C3]*

A noted difference however in associating trust and reliance, was that trust developed through repeated reliable use when interacting with automation. By this metric, participants acknowledged they would theoretically be more likely to trust systems than individuals:

> *"…if a machine has a definite higher reliability than a human then I think it is much easier to trust a machine than it is a human." [C1]*

Three of the participants specifically expressed that they would conceptually deem automation as more trustworthy (Rousseau, et al., 1998) as "*humans are naturally flawed"* and reflected on their assumptions when judging human error. For example:

> *"The thing is I am more likely I think, to trust a machine than I am a human when it comes to things like that because, let's say that the machine was just as reliable as a*

human so it is just as likely to make a mistake as a human but, I don't know there is something slightly more comforting to me anyway to be like ok well if something goes wrong with the machine it is because of that 1% chance that something could go wrong whereas with a human…. if they had the same reliability as a human, how would you feel about it but if a machine has a definite higher reliability than a human then I think it is much easier to trust a machine than it is a human." [C2]

Reliance, when associated with automation is discussed further in section 6.2.2. This term is an underlying theme that is recurrent across many of the super- and sub-ordinate themes. Themes in the thematic analysis (see Table 30 for full outcomes, located at pg.206 in Appendix A) which share numerous references are those in Trust (and the subordinate themes; Capability, Reliability and Decision Making) and those in Self-Determination Factors (especially those in the Competence superordinate theme and its subordinate themes).

The impression that reliance is positively correlated with better trust facilitation, conflicts with the numerous accounts where participants express attitudes that contradict vocalised statements and underlying notions, such as:

"Just thought though, that's only if you're rational and logical, you could have an illogical or irrational mistrust of something. So, it could be like really reliable, there could be no risks, but you could still mistrust it, and that's irrational but valid I guess" [C2]

There are aspects of cognitive dissonance where participants appear to place greater trust in automation when asked explicitly but have pessimistic and disapproving reactions throughout the narrative. For example, the sentiments of Participants C2 and C5 were those which featured either overt or covert negatively framed references (see Table 25).

Table 25 - Table of Overall Sentiments [Civilian]

Participant	Very negative	Moderately negative	Moderately positive	Very positive
C1	27	17	21	19
C2	24	21	19	14
C3	3	7	3	7
C4	18	7	9	12
C5	22	14	17	15

It should be noted that if this is a recurrent narrative throughout the military personnel also, that self-reporting metrics of technology acceptance and trust in automation may not be wholly or appropriately explanatory.

Category boundaries play a key role in how the civilian cohort defines trust in two distinct manners: (1) trusting the system as a tool; and (2) trusting the system as an ally. In the former, participants would express trust and reliability more positively with automation, systems or AI which were subservient, closed systems or where operators were still in-the-loop. In the latter, when the same systems were higher achieving, open and evolving or anthropomorphised, more uncomfortable underlying sentiments and attitudes relating to trust were expressed. These themes are further discussed later in this chapter (see section 6.2.3)

With regards to broadly defining the concept of trust, participants described reliance as an antecedent for developing trust whether from a knowledge standpoint or developed over time and use. However, how participants compartmentalised trust when framed with automation differed in two distinct ways which were dependent on the presentation of the system and its capabilities. Exploration of facilitation of trust in automation were carried out by examining the themes of automation capability, perception and category boundaries. This is further discussed in the next section.

6.2.2 Exploring Trust in Facilitation of Automation

With how participants frame trust established, this section discusses the factors participants associate with how increased levels of automation can be facilitated. As outlined in the previous section (see 6.2.1), a significant concomitant topic is capability. Capability can be related to reliability or the perceived capability of the system, the team dynamic or the operator (the role of the operator and other human focussed factors are discussed in the next section, see 6.2.3).

Capability was defined at nodes where the aptitude of the system was discussed in isolation or related to automation (some artefacts were additionally coded under the competency superordinate theme). For example, participants views *"...would be altered by sort of your knowledge of how well it would, is able to do its job".* The perceived capability can be associated with several factors, both within the artefacts as well as the literature, regarding development of schemata through priming, prior knowledge and experiences. From some of the discussions by the participants, capability was often associated with incongruent understanding of the type of automation and/or the level of automation the system has. For example:

> *[C3] "Yeah I often see that when say older people use, you know computing systems and they just expect the system to almost interpret what they're trying to do and give them what they expect, whereas they maybe perhaps don't understand the logical sequence nature of what the system is actually doing."*

> *[C5]"Yeah, I remember the first job I had after university and I was introducing CAD into design office essentially, and I had auto CAD and I remember once the manager came to me and he said can you get the CAD to do this, and I was like I can use the CAD so that I can design the thing that you want, but that's not going to do it for you, I do it and I use that as a tool. That was something I really had to hit home to them that the computer does not make the designs for you, the human does, and they use the computer as a tool."*

This theme is well supported in the literature regarding inappropriate reliance in a system. Participants often cited that if they have knowledge of the system, they would be able to mediate their reliance and may have overall better trust in the system. Such as:

> *"Yeah, maybe over-expectation as well, so we're told here's an automated system it's going to do these things for you and you may end up expecting it to do more and when it doesn't you blame the system, well no the system was made to do these things, you expected it to do more and you relied on it too much to do the things you expected it to do, that's on you, or it's on the person who sold it to you or it's on the person training you to use it." [C5]*

However, knowledge of a system as a proponent to trust in its reliability or capability is complicated. In questions which prompted participants to introspect on their feelings towards higher levels of automation and artificial intelligence, the trustworthiness of humans appeared higher regardless of the reliability metrics of which they aligned many of their previous statements by:

> *"Yeah if no one understands how it's working, you can still trust something if you don't understand how it works, I don't understand how [C1] works but I trust him…" [C2]*

And conversely:

> *"I don't think it's trust that's the issue, I think it's reliance, like I trust myself, if I'm using it properly, I trust it, we're grand we're going where we need to go, but I don't rely on it." [C5]*

Many of the topics surrounding capability, reliability and trust are evocative of the analysis in Baier (1986), where they propose whether reliance can exist in the absence of trust and vice versa. Additionally, they make a small distinction in the terminology of trust which helps to better define the rift between trust and reliance. The paper submits that:

> *"To entrust is intentionally and usually formally to hand over the care of something to someone, but trusting is rarely begun by making up one's own mind to trust, and often*

it has no definite initiation of any sort by grows slowly and imperceptibly" (Baier, 1986, p. 240)

Entrustment of tasks to an automated system would seem more appropriate in a lot of the topics covered by the participants as whether the individuals trust a system, technology or artificial intelligence is erratic and relies heavily on 'gut-feeling'.

To explore the underlying 'gut feeling' of participants and reflect on their instinctual reactions to technology, participants were prompted to examine their beliefs on how they defined automation. Using adapted definitions of technology from Bousquet's (2008) paper on defining socio-technical networks in the battlespace (refer to Chapter 1.1, Section 1.2.1 for a more detailed explanation). By providing a framework in which to reflect upon, participants could then discuss their attitudes towards increasingly sophisticated levels of automation. The examples that were provided to the participants were that of a simple calculator, a satellite navigation system, a driverless car and the Internet. All of these systems contain automated processes and increase in both levels of automation, the role of the operator and the capability of the system. Reflected in the following passage, documented antecedents of trust in automation, as outlined in the systematic literature review (see Chapter 3, Sections 3.3.2, 3.3.3, and 3.3.4), often rely on transparency, fidelity and that negative experiences can have an effect on negatively priming users:

"Yeah but maybe it also comes down to like, you're doing your own check as well, like something that you can and can't check, so like with a GPS, if your GPS is taking you in the middle of the ocean, you'd be like ok it's lying to me, but you might change the way you trust the GPS, because you can do a sanity check yourself, whereas with like a self-driving car, or maybe the internet is a better example, so with the internet, if something goes wrong you don't necessarily know it's gone wrong, and then it comes as a surprise, and I think that's, it's the surprise ones where we start trusting things less." [C2]

From using the categorisation of the different types of automation, the topic of anthropomorphising, whereby technology is given human or humanlike attributes, is discussed in 18% of the civilian data sets (refer to Table 30, page 206). Participants had two main discourses when discussing technology with humanlike attributes; (1) whether the system applied superficial characteristics; and/or (2) whether the system mimicked humanlike cognitive processes.

The former, often stimulated discussion relating back to the concept of trust and reliability, whereby capability of the system and the misappropriation of trusting behaviours may occur. For example:

> *…[anthropomorphising] might still help it fit into the unit but maybe not because if it looks like a human, you might expect it to have more qualities of a human…[C1]*

In addition, the use of multi-modal interface interaction, such as using voice instead of text influences, was discussed. Three of the participants agreed that vocal communication (such as GPS or Smart devices such as Amazon® Alexa or Apple® Siri) increased engagement with the information and felt that with the humanlike properties that they would instinctually cooperate with the device. An example of this as follows:

> *"[context: using a device with a multimodal interface]…but just giving a machine a voice…if it was a machine that was supposed to be monitoring all these different things and been telling you when something's going wrong or not, and it was just telling you in text on a screen, compared to if it had a voice, I think that makes a massive difference to how much you trust something….but if it was just a machine, with text going on the screen or something, or it had no human characteristics to it. [C2]*

A sentiment shared across all the participants was that the use of superficial human-like characteristics, often created an empathetic response which they felt had an impact on their interaction, even when critical of the systems they are implemented:

> *"No I think there's something about, even though you know it's not a human doing the job, there's something human about it, so you're kind of like, I can put my trust in this*

sort of humanlike thing, even though it's not human…I think there's like a natural tendency to trust because we have to to survive, work together, trust and so on. I think it helps if you have a machine that has human characteristics for sure". [C2]

There is an underlying expression which dominates many of the civilian transcripts, that technology is expected to behave akin to socially constructed rules. The Media Equation (Reeves & Nass, 1996) (for more in depth coverage, refer to Chapter 2, Section 2.5.2) advocates that humans fundamentally interact to communicative technology (especially those with humanlike characteristics) as though it were another social actor both consciously or unconsciously regardless of perception. Many autonomous technologies are electromechnical but share human qualities, and this feeling of discomfort is seeded in their placement close to *category boundaries* as these actants are not defined as binary. The second discourse of the mimicry of human cognitive processes is a theme that is noteworthy when applied to the notion of increasing levels of automation which will require more sophisticated automation through artificial intelligence:

"… I think when it comes to AI and it's starting to make decisions that question, that has its own morality, that's when it gets uncomfortable." [C2]

This theme is discussed at more length in the following section.

6.2.3 Inter- and Intra-Personal Factors

To categorise inter- and intra-personal themes, categorisation used Self Determination Theory (SDT) categories (Deci & Vansteenkiste, 2004) as they are intrinsic to human behaviours and development of personal autonomy across gender, culture and time (Chirkov, et al., 2003). Personal autonomy and competence were recurrent themes which subsequently had primary salience in the thematic analysis (72% and 66%, respectively. See refer to Table 30, page 206), in addition to relatedness which delved into viewpoints on inter-personal observations.

In relation to the closing of the previous section, discomfort at technology displaying humanlike processes exist at the category boundary as there is mimicry of human identity without human presence

(MacDorman, et al., 2009). Mori suggests that the uncanny reactions come from our expressions of self-autonomy and thus an instinctual threat to our identity (Mori, et al., 2012). This relates back to Section 6.2.1 and the participants conceptualisation of trust as many of the negative viewpoints were expressed when technology was viewed as an ally. The cultural lens in which automation is viewed is consistent with Western ethos in that '*the artificial reproduction of nature is a negative act*" (Kaplan, 2004, p. 10; Pak, et al., 2017). Culture affects the calibration of trust in automation (Huang & Bashir, 2017) and a superiority discourse that underlies the narratives of all the civilian participants supports the proposals of the Media Equation and the feelings of power and domination in the Human-Computer relationship.

The following excerpts explore the conflict whereby participants express their discomfort in the role of automation superseding ability yet must remain under the stewardship of an operator to pacify the ego of the users. Some examples of this 'hubris' are as follows:

> *"… you put a human in there, not because they have any real purpose, but just because people feel more comfortable seeing that there's a human there somewhere because they're much more trusting of that human than of the computer system, even if it's actually just a façade, even if the human isn't actually doing anything, it's just a cardboard cut-out or whatever, they'd still feel more comfortable with thinking that there's a human there".[C3]*

And:

> *"Say you had a control room(.), and you say it's fully automated(.), does everything it needs to do(.), but we have:: to put some people in there ↑just↓ in case, and <<u>you use those words></u>. At <u>no point::</u> have they said, these people have <u>any</u> idea what the system is, >what it's doing, how to fix it if it breaks, they have no idea of anything<, but <as* ***soon::>*** *as you say(.), ↓we've put some people in there to check just in case, ↑****everyone goes oh ok great****(h). >And you sit back and you go<, the automated system is so:: much more intelligent than <u>everybody</u> in that room combined(.), and ↑yet↓ you're quite*

happy to dismiss that over somebody else that you've <had no evidence of their capabilities>(.). Chances are with an automated system it's been demonstrated and whatnot (sigh)" [C5][11]

The civilian cohort expressed sentiments and situations which echoed the feeling, that despite the growing chasm between capability of human and non-human actors, the attitude that there must always be a human-on-the-loop was persistent. That humans must have a superior supervisory role, regardless of capability of the operator because of *just in case* discomfort.

An example of the shift in attitude in which a system is deemed more trustworthy, both in isolation and in comparison, to a human user, is associated with utilising a system as a tool, and the decision-making sequence is transparent and consistent to the operator:

"I would trust the machine to be….consistent in its decision making; again back to this point of consistency and that probably has a lot of value so in terms of trust I would trust a machine to make the same decision – whether or not I agreed with that decision is another question but I think I would trust a machine to make the same decision over a human, based on any amount of information." [C1]

This suggests that confidence in the automation is related to the entrustment of a task and the repeatability that there is contractual trust whereby trust exists if the operation is obeyed and this agreement ensures suspicion of ill will can be removed. This is aligned with the premise of relative trust and power, suggested in Baier (1986), whereby a power dynamic exists between trustor and trustee, where the latter is put into a position where they are vulnerable to the unknown intent or will of the trustor.

[11] Jefferson Notation conversational analysis was used in this instance as there was heavy use of tonal inflection (e.g. sarcasm) which is not reflected in the isolated context of the isolated quote.

Vulnerability can be mediated when there is a mutual exchange of trust. However, this actual or virtual contract is impeded by the complexity of technology with humanlike capabilities without the familiar underlying cognitive processes and schemata:

> *"… be that people naturally feel much more comfortable with the idea of trusting another human than with trusting a machine or a system, because they're another person like me, they think the same way that I do…" [C3]*

This impediment towards facilitation of trust can be stressed further in safety-critical situations, whereby risk compounds on the power imbalance where the final judgement supersedes that of the operator and the ego of the user is threatened:

> *"… I personally don't need to know how it works but I think that there's some comfort in knowing that maybe someone somewhere does. So if nobody knew, … if this AI was creating its own neural nets and stuff, nobody understood how it was doing its thing, even if it was like, stupidly reliable and was getting great results every time, I think I would feel uncomfortable, because no one knows for sure if it's always going to be like that, like now it's getting to the point where these things could start making decisions by themselves and, it gets science fiction-y but…[C2]*

Participants additionally expressed worry about the consequences of entrusting decision making to automatons. There were concerns about the legal recourse and consequences for erroneous and broken trust contracts:

> *"I think that is something that is different between a human making a decision and a computer making a decision at least at the moment, that even though a human might be less reliable and inconsistent, we have a culturally accepted standard of if they make a decision that causes problems or causes issues, society has as mechanism for dealing with that person's decision, and that is a legal or cultural or whatever, structure, but at the moment if we have a machine and it might be more reliable, we might trust it to be more reliable and consistent, but in that 5% of cases where it makes an error or the*

0.1% of cases where it makes an error, if that error is catastrophic, that doesn't sit well, certainly not with me, I don't know about you. We don't have a mechanism to, an existing mechanism to deal with that." [C1]

"I think that's an interesting topic though because what if you programmed a machine to be aware of the law, these laws that you've created, and then it does something that you're not comfortable with, it kills someone let's say, but it has a perfectly logical reason that fits in with the laws." [C2]

The response by participants supports Reeves and Nass (1996) human-computer relationship theory where humans equate humanlike behaviour with socially constructed rules when it comes to ethics and morals. As non-human actants cannot be held responsible in the same manner which human individuals can, discontent can be felt in the gap where punitive action cannot be performed. As disciplinary action, however, can be taken on those who are integral in the creation of the artificial actant many participants queried whether hierarchical accountability or command responsibility would play a role:

"I think that would require some sort of cultural and legal shift, but I think if there was a mechanism for dealing with flawed decisions, because ultimately if we assume we have some level of complexity that enables decision making and ownership of those decisions by the machine, then that ownership of those decisions also carries with it ownership of the consequences of those decisions." [C1]

The current cultural norm is to view error by technology and human actors as two separate events. For example, the attribution of blame is a binary decision – either the fault of the operator or a failure in the automated system and objects. However, there has been a shift within the human factors pedagogy. As the separation between human and artificial actors narrows, the web of dynamic interconnected elements means that neither party is solely responsible for fault or failure. It is suggested the term 'system error' is used to holistically describe the actor network of sub-standard design and poor functionality of both parties (Salmon, 2018). However, the difficulty in attribution of error to the system as a whole, not the actants within, is the issue of accountability and responsibility. As touched upon in

earlier extracts, assignment of liability is a factor which the civilian participants emphasise is a prominent factor that affects their desire to trust automation. An interesting discussion in the data is one which reflects on blame culture of many safety-critical domains (Pidgeon, 1998) and fragmented risk management:

> *"I think it's very easy to use the automated system as a scapegoat for these accidents… infrequently do we hear people say, there was zero training, like none of these guys were trained, and none of these guys were told about the implications of what could happen, and the amount of cost cutting that had been taken place and the amount of equipment that hadn't been maintained and was out of date and wasn't working properly and the organisational culture and how that had changed and the pressure and the stress and the fatigue and the workload, on the people within the system, are all levels, nobody talks about that. "[C5]*

In an atomistic system, the weaker entity of the power dynamic is the artificial actant (in this case the automated system) as the operator has a supervisory role. As systems are not inherently thought of as human, there is no moral dilemma if a trust contract is broken with an artificial agent. Which is why it is easy and more comfortable to assign culpability to the system, than peers. However, as we become custodians or teammates to embodied artificial, autonomous, conversational actants, a moral transaction may have to occur once technology starts to make their own moral and ethical decisions. The ability to willingly entrust these systems, regardless of reliability, with these decisions was a source of apprehension to participants. For example, C4 discussed their uneasiness in autonomous decision making in high risk situation, irrespective of the reliability compared to human performance:

> *"…statistically how many accidents happen now because [human error], against how, statistically, how many accidents can happen because a computer is actually driving, or the computer having to make the decision … However, as a society we are not that keen on accepting that a computer is driving and there can be an issue where the computer will have to kill you…" [C4]*

A critique which may explain some underlying behaviours negating trust regardless of metrics which the participants feel should facilitate it, is the concept of Othering. The term 'Othering' is a concept based in Phenomenology, that describes the 'Other' as dissimilar or opposite to the 'Self', of 'Us' or the 'Same' (Bullock & Trombley, 1999). It is also a characteristic that is ubiquitous across cross-cultural individual and group identities (Powell & Menendian, 2016). It would suggest there is internalised conflict of 'Us versus Them' as human-like technologies would be the opposing unit to humans on this theoretical spectrum. Furthermore, this also connects with the unease felt towards actants close to the category boundaries which threaten the concept of self. Trust is also affected, not just through this distrust of the 'other' but also in the power imbalance which othering implies as it often defines the opposed entity as belonging to a socially subordinate category.

However, social cognition over the last 15 years has given insight into empathic responses and unconscious bias into 'others'. Research by Fiske, *et al.* (2007), suggests that competence and warmth play a role into integration. They propose that high competence and low warmth can allow for passive facilitation. This ambivalence was expressed by civilian participants when discussing automation with consistent reliability and low risk (that these technologies have not proven harmful to the 'in-group'):

> *"… but I think automation works really well when people don't think about it. Like no one is really thinking about factory workers and a lot of the time now it is just automated machines." [C2]*

> *"Ah but we're very happy to trust automated systems in other ways [mobile phones]. It's become so natural that we don't even think about it…." [C5]*

Similar to the effect negative experiences with automation have on trust in the system, perceivers attenuate to behaviour that disconfirms, than confirms, warmth (Singh & Teoh, 2000). If warmth and negative experience are perceived as indicative of risk or harm, transparency and fidelity may assist in mediating these factors and aid in the facilitation of appropriate trust.

Perceptions of asymmetry and imbalance are underlying complexities in the attitudes and behaviours of the participants concerning trust toward emerging technologies and increasing levels of automation.

By contrast, C4 emphasised the benefits of a balanced unified partnership, where allocation of function is appropriately shared:

> *"… to have an unbeatable system is when they pair up the best computer with the best human chess player. So that's absolutely fool proof way or a fail safe way to actually guarantee that they will always win, it's impossible to beat those centaurs they call them, because it's a machine and a human operator, so I guess that you will still need to have for probably for very particular sensitive systems, so just probably healthcare, like an operational theatre or something you will definitely need to have an expert there, in place to supervise and whatever the computer is doing, so they could have a robot performing surgery very very intricate and hard and tiring, but they need to have this expert in the field to actually understand all the nuances that the computer is literally incapable of catching. "[C4]*

Computers are seen as dependable peers' when users can be supported without experiencing feelings of inferiority or superiority and '*existential anxieties in the human condition*' are mediated and an appropriate human-system relationship can be achieved (Loy, 2002).

6.2.4 Factors Influencing Trust at the Macroergonomic Level

The civilian cohort observed systemic cultural inequalities within their socio-economics spheres, with which they felt impacted on trust facilitation of autonomous actants. At the organisational level, they professed concern in the integration of increasing levels of automation on communities:

> *"Yeah I agree and it's this very selfish way of seeing things and because we know that there's very selfish way of working when we think of things like automation, we go mm no, because we know we've seen it we've seen what happens with our families and our towns and cities and whatnot when automation comes in and I think that's why you've got things like your driverless cars, people are concerned because they know as soon as you bring some automation into the Western world that's it you're gone, deal with yourself, oh what you've been working with us for twenty-five years, no don't care, on*

your bike. And so of course you're going to have people who, it's not so much the automation that they don't trust it's the intent, it's what it's going to be used for and what it means to them and their lives. Now if you had a system where they were like ah don't worry about it you're grand it's all good it's going to help, chances are you'd had a lot more compliance, enthusiasm towards it." [C5]

All participants echoed similar sentiments whereby there is a discourse that distrust at the organisational level trickles down into allocating trust at the inter-personal level with automated systems. Although the automation is not perceived as harmful, in fact often identified as a beneficial aid, mistrust still shapes their worldviews. As in the previous section, trust is oft associated with a power dynamic. From prior experience and historical zeitgeist (e.g. industrial revolution) of exploited socio-economic groups, the way individualistic societies manage organisational responsibility is an underlying cause for concern in the integration of increasing levels of automation.

However, as outlined by C5, a positive cultural shift has already been observed where the intent of integrating automation has been transparent and beneficiary in addressing systemic inequalities:

"So if you think about actually, we've already had automation making people's lives better and we [C5 and Interviewer] might notice this more than you guys, but if you think of things like dishwashers you think of things like washing machines, tumble dryers, you think of all of those things that were introduced into the home, you didn't have women having to stand there or sit there and scrub clothes for four hours each day and then having to dry them off, now you've got machines that does it all for you, shovel it all in in the morning; you can now have a career, because you're not having to sit at home and clean everybody's clothes, and other things, and so we've already had a huge shift in society and culture and human rights." [C5]

Even though automation is used as a tool in these situations, through a feminist theory lens it can also be viewed as an ally in addressing socio-cultural disparities. Where the intent overall is to address power

imbalances and address concerns within a community, as opposed to threaten the group or identity, acceptance may be better facilitated.

Social structure on a global scale may also affect how participants trust automation and develop their explanatory models. For example, if the culture has individualistic or collectivistic schemata. Research into trust in automation across cross-cultural narratives (refer to Chapter 3 for more details) explored historical zeitgeist and subsequent media influence. Humans do not develop in a cognitive vacuum and therefore how entertainment can influence social heuristics was a topic discussed in the interviews. The following conversation highlights the impact of the media we consume, how it can prime individuals for unconscious bias, and projection:

> *"But we're only basing it off, we're only looking at this potential future in an incredibly negative way, because that's what we would do, that's only based on us, based on our own very naïve and very primal way of thinking. We may create a brand new species, a brand new life form as it were that goes actually we don't want to squabble around and kill each other like you lot do, we want to go do something fun, tarah. They might just piss off somewhere else into space and go nah this is boring we want to go explore, why are you all sitting here on this ball of rock, off we go, enjoy. Then the new AI might be incredibly compassionate and go oh ok we're sharing let's try and make things nice for each other shall we. Like why does it have to be so negative?" [C5]*

> *"Absolutely right. But I think it's that we have been taught by popular culture, science fiction and so on that AI is bad. And the only reason that I think it happens is because tragedy always sells better than good things, it always will have more clicks, it has always sold more newspapers since the beginning of the printing press, it's always been like that." [C4]*

A topic which is of interest in this segment is that participants address their unconscious bias relating to Othering. For example, in acknowledging that humankind historically has acted offensively to other groups, they remark that this behaviour is projected onto artificial actants that exhibit human-like

characteristics. The subconscious defence mechanism, such as projection, (Jung, 1976) have a phenomenological association with the internalised discomfort and anxieties discussed earlier in this section regarding threats to autonomy and identity at category boundaries. The motivation to defend the Self against the uncanny, has been documented against humanlike entities since Sigmund Freud with dolls in 1919 (Freud, 1919) through to roboticist Masahiro Mori and beyond (Mori, et al., 2012). Whilst many researchers are still keen to explore facilitating trustworthiness through anthropomorphising, mind perception to examine trust in automation is a growing field of inquiry (Wegner & Gray, 2017).

To summarise trust in automation is not simply and interpersonal act between actant and actor. It is a systems problem, in that facilitation in impacted by intrapersonal processes and the social construction of schemata at the community and organisational level.

6.3 SUMMARY

The aim of this chapter was to explore research question 1 (RQ1): "How do the narratives that people tell about automation relate to their trust in technology?". The factors which influenced trust were established through analysis of the participant discourse based on semi-structured questions which evolved in the responsive interview format. The master categories of themes are as follows: (1) how participants defined trust as concept; (2) how automation impacts trust expressions; (3) how intra- and interpersonal factors impact trust at self, group, community and societal levels.

The factors defining trust were whether participants viewed the automation as a tool or as an ally. As a tool, reliance and trust were inextricably related to one another by civilians. Capability and competence of the system reflected many of the outcomes from the systematic literature review (see Chapter 3). For example, fidelity and transparency of the system provided surface level confidence as access to underlying processes were key to determine trustworthiness. However, when automation is viewed as a collaborator, participants had inconsistent opinions with the underlying expressions in the interpretive

analysis. Introspection on how they view different definitions of automation provided more exploration of these themes.

The relationship between participants and how they viewed automation were explored through questions on their experiences or reflections with human-system interactions of different categories of automated systems. Key themes which affected trust were associated with the type of relationship participants had with the actant and what their role as a user was. Participants expressed difficulty entrusting responsibility or taking risk with automation that they felt was close to category boundaries (Reeves & Nass, 1996). When automation was observed to have human-like characteristics – whether in anthropomorphised characteristics or humanlike cognitive attributes – participants could not easily reconcile or understand their innate unease.

The inter- and intra-personal observations of heavily explored internal processes in automation mimicking human-like characteristics as more disconcerting than aesthetics. Themes of morality, ethics and complex human behaviours appear frequently interviews. Participants felt threatened by artificial actants with human-like decision making behaviour and factors involving risk, culpability and organisational responsibility. Trust as a power dynamic, how otherness may play a role in trust facilitation and cross-cultural narrative differences were also discussed.

The next chapter will look at the Military cohort and build upon the information gathered with the civilian participants. The chapter will seek to explore the dominant narratives and underlying expressions of trust in relation to civilians as well as observing differences and similarities between and within the echelons of the military.

7 STUDY 3: MILITARY INTERVIEWS

7.1 INTRODUCTION

The purpose of this chapter is to identify aspects of the military experience of trust, automation and attitudes of emerging technology usage and facilitation. Furthermore, to explore their opinions and attitudes towards their own echelon of HM Armed Forces and their colleagues. This exploration of narratives will focus on both research questions: RQ1: How do the narratives that people tell about automation relate to their trust in technology?; and RQ2: How does trust in automation differ between military personnel and civilians?. The reflections of the previous study (Chapter 6) can underscore sociocultural divergence between civilians and military personnel.

The data collection and analysis methods are presented in Chapter 4. The preceding Chapter (chapter 6) of civilian interviews were performed before the interviews and focus groups of this chapter were conducted. Therefore, a more refined responsive interview protocol was utilised, in addition to conversational boundaries where unfruitful or disproportionate discussions may arise.

This chapter reflects on participant views and dominant narratives and expressions of trust and attitudes towards automation framed by a diverse selection of Military Subject Matter Experts (SMEs). This segment expands on automaton facilitation; the nature of trust development in military structures; and the impact of organisational relationship formation and has on impeding system-human team cohesion. The chapter concludes with a summary of findings.

7.1.1 Participant Information

Table 26 illustrates the demographic information for the Military cohort. As outlined in the previous chapter, to maintain privacy, this table does not attribute individuals with their corresponding participant number used in the findings of this chapter. Recruitment and additional information regarding anonymity and safeguarding for participants are outlined in Section 4.4.1.3 (Interview and Focus Group Design),

The age distribution ranged from early 30's (Millennial (Generation Y) and digital native) to early 60's (Generation 'Baby Boomer' and late digital immigrant) which provided a wide range of digital competency and experiences with technology. Furthermore, there was diverse range in length of service and level of authority with HM Armed Forces, with the shortest service at 7 years and longest tenure at 36 years. The ranks of participant were additionally equalised at the appropriate NATO rank grades (NATO, 2010) across the different echelons so as to provide additional privacy. Therefore, their rank has been provided but does not indicate which echelon of HM Armed Forces they belonged to.

As noted below, three of the participants requested to have participant information redacted. Additional demographics which are not included in the table for privacy, were those of gender and ethnicity. The cohort included both male and female participants, though heavily androcentricly skewed. This is in part due to the gender diversity in HM Forces (Clark, 2020) . Similarly, BAME participants were included but under-represented in the sample.

Table 26 - Participant Demographic Information for Military Cohort

Branch	Length of Service	Rank on Leaving	Age Range
Army	30	Captain / Group Captain / Colonel	30+
Navy - Surface	19	Lieutenant-Commander / Squadron Leader / Major	30+
Navy - Sub	7	Lieutenant / Flight-Lieutenant / Captain	30+
Navy - Sub	34	Sub-Lieutenant / Flying Officer / Lieutenant	30+
Navy - Sub	32	Warrant Officer 1 / Warrant Officer / Warrant Officer 1	30+
Navy - Surface	28	Warrant Officer 1 / Warrant Officer / Warrant Officer 1	30+
Navy - Sub	25	Warrant Officer 1 / Warrant Officer / Warrant Officer 1	30+
RAF	36	CPO / Flight Sergeant / Staff Sergeant	30+
Navy - Sub	32	CPO / Flight Sergeant / Staff Sergeant	30+
Navy - Sub	27	CPO / Flight Sergeant / Staff Sergeant	30+
Navy - Sub	25	CPO / Flight Sergeant / Staff Sergeant	30+
Navy - Surface	24	CPO / Flight Sergeant / Staff Sergeant	30+
Navy - Sub	22	CPO / Flight Sergeant / Staff Sergeant	30+
Navy - Surface	10	CPO / Flight Sergeant / Staff Sergeant	30+
Army	26	PO / Sergeant / Sergeant	30+
Army	15	PO / Sergeant / Sergeant	30+
Army[12]	-	-	30+
RAF	-	-	30+
Navy – Surface	-	-	30+

[12] Participants requested this information be redacted

7.2 FINDINGS

7.2.1 The Conceptualisation of Trust

As with the civilian participants in Chapter 6, the clarification of how respondents define trust terminology and the perceptions of how it develops is crucial for exploring their narratives, how this impacts their worldview and trust facilitation with emerging technology.

As with the civilian conceptualisations of trust, 18 of the 19 participants remarked on how trust and reliance are inextricably interlinked, especially with regards to automation facilitation. Though it is more nuanced than this, as will be discussed, many did introspect on whether they do linguistically differentiate the two terms. For example:

> *"Yeah, but I don't think you can unlink them. To be honest I don't think it's something anybody would actually think about; until you ask that question, it's not something that I would have thought about. You got reliance and trust, but when you say it, nearly everything you do, you've got either trust in it, whether you've got to rely on something or not or you've got to rely on something to trust it; they're linked to varying degrees."[S2]*

> *"Trust and reliance. I would say reliance is taken over a period of time, because it's constant and it's constantly working properly, so you are more reliant on that particular automation. Trusting is something else, trust is an emotional thing where you can, because of its reliance, you tend to trust it a lot more". [A4]*

A common expression in the military cohort, compared to the civilians, is the option of choice. For many, due to the authoritative sociocultural structure of the military, many did not have the choice not to interact with technological actants. The facilitation of inappropriate reliance on an untrusted or untrustworthy system, especially in high risk areas, had an impact on their schema priming. This could be through negative experiences in adverse environments, or psychological distress arising from lack of control, stress and autonomy:

"I think from an operator you'll always blame the system; the chances are you are right it could be an operator fault somewhere, but nine times out of ten they will always blame the system first. I was, in my earlies I was a troop commander for, I was a royal engineer, and I was reliant on a piece of kit that was pretty fundamentally unreliable and one hundred and ten blokes, as it was at the time, and they had fifteen of these pieces of kit…and you know that you actually, you couldn't rely on the kit. I mean even without the enemy; you had the enemy it would be a disaster. Having to trust it and not being able to rely on it". [N2]

For many, reliance was not the only metric of trust. Personal autonomy and knowledge, for example with capability and decision making, were identified by a significant number of participants as a modifier of trust. Expertise and experience were significant areas of conversation for those in the Submarine service. Section 6.2.3 explores this further, that there are collective- and self-determination factors associated with pride in intelligence or aptitude in the subsurface holon (Thomas, et al., 2017) which can be associated with team cohesion or incohesiveness. In this section, however, the following excerpts explore how trust is often defined by experience and expertise:

[S7] "Do you know I think there may be, you know with these relational triangles…if you have knowledge, so you've got reliance, reliance may be enforced, i.e. you're not trained to do what the systems doing, you're reliant on that system because you have a deficiency in knowledge. If you've got low knowledge and the system tends to do the right thing then your trust will increase. If you've got high knowledge then you may notice smaller inaccuracies in which case that will be detrimental to trust…"

[S3] "…am I trusting the maintainer or am I trusting my gut instinct as an experienced user of the equipment, and you can be looking at it from different aspects, looking at it as a maintainer or looking at it as an operator?"

[S6] "Well you're constantly on your feet, and that's where experience comes in, that's where the gut feeling we were taught, and that's only learnt from experience."

However, there are contradictions with the proposed speculation that trust is mediated by experience and expertise in the participants. One of the themes explored is regarding seniority and the regimented hierarchy within the military sociocultural structure. The clause below explores how experience can be problematic when it comes to trust facilitation explanatory models towards technology:

> *"No, so it's interesting I think, and I think that's the area where there's no logic we apply to trust. I mean some of the MOD junior guys that have just come straight from training, actually we might not trust them, because we say they haven't got the experience, but if they've just came from training, they haven't got the bad habits, they haven't got the, oh just do it this way even though the instruction book says so, I think we're actually not very good at picking the right part." [N5]*

Similar to themes discussed in the civilian cohort, the topics around capability, trust and reliability can be associated with task entrustment discussed in Baier (1986). This proposes that entrustment is formal and intentional, whereas trusting behaviours are often unconscious and instinctual from cognitive priming. This is reflected in the Naval participants significantly (seen in participants: A2, N1, N2, N4, N5, S3, S6, S7), that they entrust themselves to their human teammates by the higher allowance for failure, error in decision making and unreliable behaviours compared to a system teammates identical behaviour:

> *"Don't know why it gives you hnnuuughs [indicating gut feeling/discomfort] because he could be the worst [teammate] in the world, but it gives you confidence doesn't it…We tend to be less trusting of computers than we are of humans, because we understand that humans make mistakes, it's human nature everyone makes mistakes, no one's perfect, we expect the computer to be perfect, because we've designed it to be perfect, so when it goes wrong, your trust goes straight away because it's designed to work. Not not work. [A2]*

An interesting development, differing from the civilian cohort, is that which aligns with team cohesion development within the military (as explored in Chapter 1 (section 1.2.3)), and the impact of threat to

life has towards interpersonal trust development. For example, team cohesion and potentially inappropriate entrustment is reinforced over all else:

> *[N5] "Yeah I mean it's quite interesting, we've all talked about trust trust trust with automation and probably your instinct would naturally say I trust my teammates, but actually we all know they fail, they fall asleep, they miss things, they've been out the night before, and you've got very inexperienced, but we do tend to put more stock into human emotion, I suppose because for years and years you've relied on it, and all your training was based on team training, but the reality is, automation is entirely logical, whereas teammates aren't."*

> *[S2] "I wouldn't have thought for a second that the guy sitting next to me wouldn't be able to do his job if something went wrong, I would just automatically know that he could do his job, and I'd be doing mine and he'd be doing his etc etc. I don't think it's ever crossed my mind that anybody wouldn't be able to do what they've got to do in a real emergency, that just wouldn't cross my mind."*

Furthermore, those whose experiences include special operations with increased threat to life, opined that irrespective of knowledge of human error, they may still prefer to over-rely on human teammates over a system actant:

> *"I recognise the risks, but I'm reasonably comfortable with them. I also understand the stats that supports it, the probability of that happening is incredibly low, and therefore I am comfortable with it, I've done far more hazardous things for a living so I'm reasonably sanguine about that risk to my own life." [N1]*

> *"That's very true. I think there is varying levels of trust in automation as you go through the spectrum and I think you've hit it on the head there actually, if it's to do with a life, your trust is going to be a little bit, you're going to be a bit more weary of it and you probably wouldn't put one hundred percent trust into it." [S2]*

When it comes to ownership and accountability of decision which may result in loss of life, many participants felt more comfortable if the responsibility were attributed to a human actor. This is discussed further in Section 6.2.3 with regards to morality, communicative technology and discomfort in cognitive process mimicry.

One recurring theme within the Naval subset, were the narratives which placed significant trust expressions as a result of lived experience and priming of explanatory models, whether this is attributed to human or system teammates:

> *[S2] "The submariners coming through, they're not going to have that past experience to call on. So they're going to, I won't say they wouldn't understand, but they probably wouldn't even think about what could go wrong with automation; whereas I'd be thinking, what happens if that was to go wrong, what would I do. They won't think like that, I'm sure they won't, I'm positive they won't in fact…"*

> *[N3] "Yes for me because you learn from mistakes, if you never make a mistake, for me you will never learn and for me I've learnt a lot…"*

The interviews further expressed that maladaptive attitudes towards technological reliability or capability can be associated with the lack of positive reporting. For example, as there are limited protocols for systems working accurately but significant protocol for reporting negative incidents can introduce bias in explanatory model formation.

With regards to priming explanatory models through lived experiences and shared schemata, the following section explores how participants conceptualise their own culture and explores the impact of this unique environment has on trust expressions.

7.2.2 The Conceptualisation of Culture

> *"I left the military about thirty years ago, and my wife tells me that I am still a military person, because I still do things in certain ways, and I've never lost the habit of doing*

different things in different ways. So you've got a completely different animal from the military person…" [R2]

The conceptualisation of military culture is crucial for grounding the narratives, underlying expressions and lived experiences of the personnel interviewed for this research. The regimented, authoritarian hierarchy developed in high risk and criticality environments can have a significant impact on how trust mechanisms are formed and influence explanatory models. The social schemata development is further unique as there is significant perseverance of these attitudes and behaviours post exit from the service (Binks & Cambridge, 2018):

> *"It can be challenging when you come out of the MOB[13] because I think mostly you think that the people you are with have a sense of duty and comradeship and teamwork ethic if you like, that you can depend on. So people would tend to be innocent until proven guilty and those guilty elements you try and temper if you like by training and information and things like that. When you come outside, you find that it can be the other way round; some people are very much about what they're doing rather than being part of a team, as a group endeavour we're trying to do a thing, they're just trying to do what they need to do." [S7]*

As explored in Chapter 1 (section 1.2.3) and Chapter 5 (section 5.1.3.2), the formation of trust attitudes and team cohesion is one of the primary factors which can cause difficulties in transition to post-service civilian life (Kirke, 2010). This is imperative in exploring expressions of trust in the military cohort, as highlighted above, the implications of inappropriate expectations in human teammates. Whether through over-confidence in their colleagues or a detriment to system facilitation.

One theme of interest to explore is how the crystallisation of military identity plays a role in human-automation interaction. Grimell (2017) suggests that in military identity their dominant cultural

[13] Main Operating Base

narratives are established during service, to some extent through *'the perception of being involved in something meaningful and significant'*.

A pervasive narrative is regarding the strong attachment bonds formed with their 'military family' (Mikulincer & Shaver, 2007). In a focus group exclusively of submariners, the passionate evocation of the closeness they felt with their colleagues illustrated some of the instinctual trustworthiness and propensity to trust their human teammates. Furthermore, how these relationships impact their perception of system teammates as a threat to this team cohesion. That a relationship not developed through sacrifice and adversity would be viewed as 'less':

> *[S6] "You have to be, because your living environment is inherently dangerous; you've got to cope with all that along with your job and your watch keeping and constant tiredness and that's where the training kicks in. I understand where automation would take that, may take some of that away, but it's still, you trust, you know if you're in the Dolphins you trust that man instinctively, because he's done the same training you've done. Or women."*
>
> *[S5] "The submarine world is a long way behind in the automation room compared to other things, a long way, and a lot of us in that family don't trust machines, they trust the people."*

Military identity also plays a role in the hesitation towards automation use through threats to self-autonomy. Similar to the threat felt by civilians regarding artificial actants involving culpability and organisational responsibility. For example, significant pride was, and is, attributed to job identity and personal capability and inappropriate implementation had a distressing effect on both individual and group schemata:

> *"And that was a massive mind-set change because suddenly you noticed it because suddenly the people that replied upon as subject matter experts (SMEs) in certain areas, so they go to the two people that could understand how to fix things or what to do if this happened or what to do if that happened, they suddenly became the great men*

and women because they just weren't needed anymore and they had no role in that area. And because of the change to, we only rely on the information that's given to us, this system is provided to us and it's all automated, therefore we have no input into it, there was that drive for a business to absorb that work and not the troops on the ground, and that created a gulf, and that created a major imbalance in most of the units because these people suddenly had no key skill, and they felt, a lot of them, anger and resentment to the fact that they were suddenly being side-lined; and you saw it in their work, they went downhill, very rapidly, because that one key segment disappeared, they'd had no other area to build themselves from, so other areas of their work would go downhill as well. So it was a massive effect on them, because these people who were, used the term 'Comms Gods' who could walk around and fix everything and do everything suddenly had no role because the system had become automated, and they couldn't effect the automated system in the same way," [A3]

The sentiment in the above quote was echoed in a number of participants, whereby they remarked on the discord they felt when this perception of self- or group- identity was threatened and oft responded in derogation or criticism (Petriglieri, 2011). The distress fed back as a result of redundancy through changes to military identity is explored in the research by Pak *et al* (2017) and can be seen here with regards to automation. They examined the reconstructing of identity from military to civilian self and found the erosion of their military identity through the loss of structure, role and important military relationships, led to persistent feelings of 'redundancy' when relieved of their operational role. The impact of sociocultural dynamics in group identity are previously developed in section 6.2.3 with civilians. More specifically, how perceptions of identity and professional expertise are a critical obstruction to group cohesion with automation facilitation with the high subject matter expertise of submariners.

The next section uses some of the concepts explored in the preceding sections to frame and explore how these expressions of trust and military culture impress upon attitudes towards technology, implications to appropriate facilitation and integration with system teammates.

7.2.3 Attitudes Towards Technology, Facilitation and Integration

> *"Yeah it always has, that fear of being replaced by a machine, that's always been there." [S1]*

The sentiments explored in Chapter 6 are also expressed significantly in the military cohort. The historical zeitgeist of the integration of automaton has negatively affected groups within individualistic cultures management of organisational responsibilities. As explained in the previous chapter, this unconscious Othering bias headed by inequalities, job loss and unknown structural impact, can have a persistent negative attitude towards automation.

The definition of how participants viewed automation was markedly telling on where they viewed system teammates in power dynamics. As explored throughout this thesis, power imbalances are key in how trust is conceptualised and acted upon. All participants viewed artificial actants as the weaker entity within the power dynamic (see Media Effect; Chapter 2, Section 2.5.2). Such as with the reflection below, many viewed automation as simply a tool, an asset, not a viable teammate in which to entrust tasks, decision or emotional relationships with:

> *"They're a tool to do a job…You're trusting a machine and that takes time to get the trust in the pedigree of the software which is driving it…[S1]*

The excerpt below similarly expresses objectifying language viewing automated teammates as a *"piece of kit"* rather than convey the actant in any social role:

> *"I think it's when the, when something happens that is completely out the ordinary that you've not expected, where you've had a lot of trust in a bit of kit or a system…" [A3]*

> *"From my side of the world, I would trust the guy next to me, over a bit of machine, especially if it was telling me something I didn't want to do, or there were other options, it might be the best option, but there are other options. For me on the ground, in Afghan, I would trust the person next to me over a piece of equipment." [A2]*

Once more, category boundaries are fundamental in how the cohort experiences social responses with automation in: (1) propensity to trust the system as a tool; and (2) unwillingness to trust the system as an ally. As the former suggests, participants more readily express trust and reliability with subservient automation, systems or AI in closed systems or where operators were still in-the-loop. Subsequently, when systems are humanlike in cognitive processes or decision making, open or anthropomorphised, this caused expressions of distress.

Specifically, regarding anthropomorphism of technological agents, there were a number of semantically similar themes centred around parasocial trust mechanisms discussed by a number of participants (A1, N1, S1, S4, S5, S6, S7, S8). Some felt that the implementation of anthropomorphism as a method to evoke trust behaviours akin to human-human interaction, were dishonest which has organisational implications for distrust in authority:

> *"If you know that something doesn't work and you think that by giving it a human face or a pet's face that you're going to overcome that then that's not overcoming it, that's concealment so…" [A1]*

Other participants opined that creating human like avatars may be influenced by socio-cultural prejudice and underlying personal bias. As discussed in the Civilian Study (Chapter 6; Section 6.2.3) out-group denigration, socio-cultural bias and 'Othering' is embedded in human trust interactions:

> *"I think if. I don't know, if you can choose your avatar, then you can choose someone who potentially you would trust, a figure that you might. Unfortunately, people have prejudice views, particularly based around race and unfortunately it's kind of…It's a personal preference isn't it. Even though it's a machine. It's bringing emotions into it." [S1]*

In the desire to introduce artificial agents through use of humanlike information communication (MoD U.K, 2014), a noteworthy viewpoint from the experiences of Submariners, is the following:

"There's no emotion, no supposition, no anecdotal in there, all he's doing is relaying facts and he's trying to do it as mechanically as possible and then we found that we can actually do it not mechanically but electro-technically and now all of a sudden we're trying to put a human back into a system that we originally tried to de-humanise." [S7]

In the exploration on how humanoid system teammates may be used in future command and control spaces, the participant voiced the potential futility in utilising novel user interfaces in knowledge transfer. Anecdotally, four participants (A3, N1, S7 and S8) verbalised comparable narrations whereby teammates would change from colloquial vernacular to robotic delivery of data, details and commands when communicating orders or critical information:

"…So you may know that individual very well, but when he reports that, he reports it in a standardised form without any banter, without any emotion, so that you can assimilate that data. So it's the reverse of what we're saying so you're taking the human element out of it and trying to impart data in as quick and succinct and standard format as you can, that was then replaced by an automated system. What you wouldn't want to do is take the automation and try and add the human element onto it, because we've always been trying to avoid it even when we did have a human element. [N1]

Furthermore, the trend to accelerate trust facilitation through human-human teammate synthesis may be ineffectual, as the delivery is the antithesis of the lived experiences and data exchange within these environments:

"No, I don't trust anyone. You know, and most people in the military that are going, I trust two people and you're not either of them. So it's like, from that perspective if you try and make everything more acceptable, more friendly, more human, or more understandable from a human perspective; if you give it a personality or in some way make it more approachable in that way, then you're automatically going to cause distrust, because people don't trust anyone…"[A3]

For a number of participants, they viewed anthropometrics and avatars as negative and superfluous user-interfaces for data communication and C3I operations. Specifically, for the potential to maladaptively influence perceptions of capability, system-teammate integration and facilitation of automation due to the inherent human-trust issues developed in everyday life.

Furthermore, an additional hesitation for integrating emerging technologies and artificial actants were into the wider training of recruits using Augmented and Mixed Reality Virtual Environments (AMVE). This was discussed at length with OF-2 and OF-3 ranked Submariners (NATO, 2010) with a diverse age and time in service. Firstly, they see the operational benefits for training purposes due to the impossibility of using a functional submarine in training operations, however they also reflected that the tangibility of threat is difficult to replicate in a safe environment:

> *"You're not going to risk the submarine itself or the crew and everything else yeah quite rightly so. So there's the realism of training, is it can only really be done ashore, and you never get the same level of realism in a training place than you would on board the submarine. There's various things, even if it's just a pillar in the way or a bit of kit in the way that you can't get the valve, it's things like that." [S2]*

Why this is of interest is that there is the underlying expression, that training can be useful for priming technology facilitation and appropriate use, but themes of reliance, capability and trust in teammates are seemingly more critical:

> *"I've been involved in real emergencies on submarines, I'm sure you have as well and the way you react and the way the crew reacts is completely different from when you're doing training or anything, it's completely different. Everybody is more, they're more tuned in, they know this is real, they know they've got to get it right. Whereas when you're just doing it for not real, yeah you're going through the motions but I'm sure [S1] will agree with me, when it happens for real, you see a completely different attitude in everybody, it's completely different. Everybody just completely changes, I mean your training kicks in, very much so. I'm not aware if it's the same outside the*

military or submarine service but I would imagine there's not many people getting in that situation, where you're under the water x amount of hundreds of metres and you've got water coming in, that's not good. You do something about it, and I think that's just the way we think. You're relying on your comrades." [S2]

In addition, participants N4, N5 and S3 further echo that prior experience may prime explanatory models for appropriate facilitation with automated teammates, but social schemata and cultural exposure were perceived to impact trust mechanisms more.

Related to training and experience, the role of expertise and trust of those in decision making roles (also discussed further in section 7.2.5) is entrenched within the cultures explored in this research. The most clearly defined attitudes of binary views of automation are seen within the subsurface personnel. For example, participants either accepted in-the-loop automation where SMEs and command have the primary decision-making directives or that operators were fully out of the loop. The following excerpts are an example of these impressions:

"And there is at least one case I know of where people's lives have been saved because the person in the chair did not follow the drill because he knew the drill was wrong. If it had been down to a machine, then a number of people would have been k…[pause], would not be here now, and that was the guy on the chair, at the time made a decision that from the evidence he had to him, yeah, meant that he shouldn't follow the drill as was laid down, and went a different route and saved all those lives; and that's what worries me, is a machine will just do what it's programmed to every time, even if it's the wrong thing, so you still need the person in the loop." [S5]

And:

I don't think they're ready for it and it's another one, it's either you have humans in it and then you keep humans very closely in the decision making thing, or you take the humans out because then they start causing the problems, because humans don't follow rules like computers do. [S7]

A key thread which links many of the themes presented in this section are of risk, capability and self-determination factors. As the military is one of hazards and peril, the culpability of those in decision making directives is associated with trust dynamics such as power distribution, predictable consequences and organisational responsibility. However, artificial agents currently exist out of certain trust based social mechanisms such as the legal system or rules of engagement. These are designed to *"guarantee or secure . . . trustworthy conduct . . ."* (Weber, et al., 2003) and therefore with automation and emerging technology, the need for trust *"becomes proportionally greater"* as the components of trust, predictability and evaluation are undetermined or unidentified (Barber, 1983).

The next section explores trust at the inter- and intra-personal system ergonomic level. For example, to explore the personnel subsystems and the intrinsic needs and relational bonds of the individuals and how this impacts trust and technology facilitation narratives. Section 7.2.5 will look at the organisation and managerial subsystems which impact human factors at the macroergonomic level.

7.2.4 Inter- and Intra-Personal Factors

> *"You have to trust implicitly the man next to you, that he knows how to do his job and yeah you're a big team family, you're in it together. It's different, there aren't many equivalent industries in the civilian world really. And there's not many industries where you're so, you've all effectively sacrificed your freedom, but you know that you can be called to service and you don't have a say in it, to do stuff." [S1]*

To continue the themes of team cohesion and close interpersonal relational bonds as introduced in section 7.2.2, the 'military family' narrative is a core element of trust formation and development within the Military domain (Mikulincer & Shaver, 2007). Group identity and kin selection in research has evolved from early cooperative group behaviours (Smith, 1964) to a diverse collective of ideas exploring the nuances of in-group dynamics and their impact on collective behaviours. For example, research by Smith (1983) into combat and comradeship explored group loyalties resembling familial love (or stronger) was identified in military narratives as a reason for their perseverance behaviour in high risk, high mortality environments. More recently, social identity theory and the in-group identity

fusion of 'military family' has begun to explore themes of sacrifice and sacred leadership (Whitehouse, et al., 2014). They suggest that the family-like social bonds within the military *"which indissociably wed personal identity to group identity, may well enable combatants to fight on, even in the face of death and defeat"* (Atran, et al., 2014).

Cross-cultural studies propose that most resilient forms of group identity are bounded by sacred values (Caspi-Berkowitz, et al., 2019). These values are unconscious decision-making tools that a community may explicitly or implicitly treat with protected or infinite significance (Tetlock, et al., 2000, p. 853). These themes and intuitive decision making, especially as a supplement for traditional decision-making methods in in military doctrines are discussed further in Eriksen (2010). The sacred community values used to strengthen interpersonal bonds and bolster trustworthiness is a key underlying expression which can impact the social priming and explanatory models for both human-human and human-actant interactions.

Sacred, or Heroic, leadership as a method of organisational management, is one that is fundamentally intrinsic to the authoritarian hierarchical leadership structure of the military, but also naturally occurring in the high-criticality domain in which it is strengthened. The term heroic leadership, is often based on military style leadership that is often used in command and control situations and has basis in influencing others in accomplishing any task or objective whilst putting individual wellbeing at a lower importance to those of the mission or comrades (Lebel & Ben-Shalom, 2018; French & McCain, 2004). Cohen (2010) defines eight common traits in Heroic Leadership, which *"To summarize, …is doing what is right while for the most part ignoring the potential benefit or harm to the leader…leading a group with absolute integrity while raising individual performance to a personal best, and building a team spirit of sacrifice for the common good".* Together with identity fusion, familial attributed ties and themes of self-sacrifice (Swann, et al., 2014) the group members share core characteristics, and this social response priming shared within the culture can increase trust perceptions by way of the 'family like' ties and sacrifice. Grint (2010, p. 100) theorises that in alignment with shared sacred values, "*…sacrifice is an essential mechanism for the performance of all forms of leadership".*

This is of interest to addressing the underlying expressions in the Military cohort for two reasons; 1) how sacred leadership is impacted in the unique environment of Submarines?; 2) how the role of the 'heroic leader' influences trust and the subsequent effect on system interaction?

To address the first query, there is the regimented organisational structure within the commanding officers of the submarine. However, as has been touched upon in the previous sections and throughout this chapter, the role of the SME's are unique in being somewhat characterised as a decentralised leader through working collaboratively with the group (Harris, 2009). For example:

> *"You know on a submarine control room, they don't talk about individuals, it's the command team, it's everybody in the control room when we do a thing, everybody is responsible, everybody can raise an issue, everybody can say stop or don't do that because such and such a system is not aligned and these kind of things. So yes the trust is earned, but the thing is you can do things to modify it or compensate where there is a, you trust them to do the right thing, or to have the right intent and you can compensate where you think they may not be sufficiently knowledgeable or practiced and things like that." [S7]*

As the boundaries of this leadership are porous, this allows for other individuals' participation within the organisation to strengthen social trust responses within and between individuals in the community (Gronn, 2003, pp. 27-50). Most importantly for the SMEs, knowledge dissemination is a key proponent of the creation of an informal leader as an emergent property of the group (Uhl-Bien, et al., 2007). The trustworthiness of an integrated heroic leadership, is neatly encapsulated by a veteran Royal Navy Artificer:

> *"…submariners are definitely one for all or all for one." [S6]*

Going forward into more automated, complex battlespaces, the role of these self-organising, emergent, positive, non-linear leaders can correspond well to evolving Chaoplexic warfare (Bousquet, 2008). Furthermore, they also express a dominant narrative which runs throughout the participant interviews and focus groups, that expertise is revered and respected significantly:

A qualified submariner gives you that sort of level of credibility that you don't get in other branches of the military, because not only do you know your job, you know the systems and the jobs of everybody else in that boat, you know the system, if you are the only person in the compartment and something goes wrong, you know what to do to save everybody on that boat, and that's the level of knowledge. And sometimes that means you've got to think beyond what the drill says you should do, because you're going to face situations that are not covered by your training and you've got to reason through it and I don't think the machines can do that yet. The point is though, if you rely on a machine to do it, ninety-nine point nine percent of the time, point one percent left of people to do it, they don't have the knowledge or the skills or the experience to actually make the right decision. [S5]

The subject matter expert role (SME) is valued and somewhat sacrosanct (refer to 'Comm Gods') in the authoritative hierarchy. The role within the organisations of the SME at the interpersonal level is important to the facilitation of trust with two main reasons: 1) their role is embedded in inspection rituals, and 2) charismatic or heroic leadership indirectly creates trustworthiness.

Explored by Bluhm (1987), inspection rituals create trust through the accreditation or testing of complex systems by an individual with expertise. Thereby, allowing others to participate in everyday activities under the assumption that the individual was trustworthy in their assessment, and that the organisational structure is trusted through accountability (see trust power dynamics in Chapter 2, section 2.2). Another example of inspection rituals and accreditation is embedded in the regimented ranking system intrinsic to the military authoritarian structure. The following excerpt is the highest ranked individual (OF-5 (NATO, 2010)) regarding trust development within a unit:

"Yeah, I think from me it's the immediacy of the supervision…there's the connection between, you've got somebody who is trained and experienced and licensed by somebody else who is independent and regularly trained and updated, and they are the

person there when you're trusting the organisation that they wouldn't be there if they weren't all of those things. [A1]

In addition, the removal of responsibility by a leadership role can additionally influence trust behaviours through the *"[silencing of] anxiety of followers in and through the provision of safety and security"* (Bauman, 1993, p. 73). This can be linked to the theme of trust, regarding power dynamics and obedience, explored in Chapter 2, section 2.3 by Milgram (1963) and Zimbardo (1969) involving agentic shift. Recent research has also identified traits in subordinates, that leaders are often attributed greater sensemaking abilities than followers (Alvesson & Sveningsson, 2003) which is key for exploring the narrative surrounding military identity and professional expertise (Woodward & Jenkings, 2011). More importantly how there are negative expressions towards both automation and new recruits due to the actual and perceived impacts of skill fade in these subject matter experts:

"You're underwater, and it all goes wrong, the last thing you want to do is have to wait for a computer to reboot itself to put it right, so that's where, in the submarine world we still do pride ourselves in our knowledge, we know how things work and everyone has a level of knowledge, it's probably getting diluted now if I'm honest with you, because when I did my qualification first time round back in the nineties, it was learning everything, what valve did what, what the consequences of shutting that, what are the consequences of doing this, and you spent weeks, sort of like ten weeks, going round and that's all you did, you had books, you had drawings, you could draw systems out, you could do the lot. Now you go to one of these A class platforms, actually the computer does it for you…so the automation is there and it does it, yeah brilliant that is a really good system, apart from when that one fails; … unless you know where it is, you've always relied on a computer, because whereas before you went round manually moving a valve, moving a flap or whatever, computer says no and that's it, you just sit on your hands then, I don't know where it is, it's in here somewhere. So unless you can build in any fail safes or it wouldn't be appropriate to override some things, or

overrides, or yeah maybe it would work, but that's the human element, because if you've then got the expertise to decide, actually, you don't need that to be in that position, it will still work, you can teach a computer to do that but it's the human element of, yeah I know this system I know it well, I've used it like this before, it will work.[S8]

Skill fade is addressed more in the following section (section 7.2.5). Regarding the above excerpt, there is an underlying expression of mistrust with automation in comparison to human actors through two dimensions of trust. Firstly, there is the aspect that trust propensity can be attributed towards prediction of risk and their consequences. In the submarine environment, as well as the socio-culture of the military, there are inherent aspects of social structure which can mediate these cognitive processes. For example, the relational bonds between colleagues and schema priming through lived experiences. With technology, the transparency and thus power dynamic are a lot more clouded as it is unknown if they will reflect the same social responses. Secondarily, the other aspect is that of personal consequences and/or self-interest factors. As highlighted in both the civilian and military cohort, the legality of error responsibility is a major concern and that distrust can be fostered by the psychological tension or fear of the inability to predict the event or judgement within the task (Deutsch, 1958):

"…it has to be a gradual process, people fear change in a big way anyway, especially in the military, you know we trust what we already know, and to bring in big changes too quickly, then we're known to moan a little bit, aren't we…Yeah I think the only way it's going to work is gradual, people fear change, despite what they say, ohh technology, but they do fear change." [S6]

The reflection of mistrust arising from uncertain social response predictions and judgement is also reflected in the use of ubiquitous technology and increased levels of automation, through concerns about surveillance:

"It does give you the big brother mentality though doesn't it, look who's watching that sort of thing, you have to think about, is it a total division of your privacy and everything else." [S3]

Whereas there is balance in trust and power dynamics with human actors, developed in the close bond formation and reciprocating support for one another, this is not apparent with system teammates. There is concern that the system actants exist outside the organisational structure and thus cannot be trusted as a predictable or known trusted entity. Furthermore, there is also the rhetoric that technology can be used in applications such as *"unannounced inspections, external audits, quality control sampling or 'sting operations'"* (Bluhm, 1987, p. 340). There may be impact on explanatory models as artificial actants may be seen not just as unpredictable teammates, but actively nefarious and malicious, despite the explicit neutrality of a non-sentient system. The trend to proliferate levels of automation at an increased rate of change may also be detrimental to the facilitation of trust in automation. The reduction of predictability in HAI may increase H-H inter-dependence, which is already heightened in the military 'family'.

The following section uses some of the themes explored in this section and extrapolates on them at a larger sociotechnical systems level. The interpersonal relationship and the intra-personal cognitive processes impact how individuals and teams interact with management, governmental and other regulatory structures.

7.2.5 Factors Influencing Trust at the Macroergonomic Level

As suggested in the previous section, the impact of unpredictable consequences without a H-H social contracts impacts the power dynamics and trust of the individuals. From a larger perspective, as the military is a punitive authoritarian organisation, the mistrust arising from unpredictable consequences to actions is more prominent than it would be with civilians:

"*I suppose there's potential for an element of mistrust in the level of intelligence, or perceived intelligence that avatars have and if they seem to be smarter than the user, and they're connected to the outside world and they can potentially feedback*

information on that person, I can see an element of mistrust from that perspective". [S1]

There can also be mistrust at a macro systems level, as increasing levels of automation and facilitation of system teaming interactions is perceived by many of the participants as an organisational demand:

> *"But it's being driven from the top rather than being driven by user demand…I mean I'd have no trouble as a result of what I do for a living embracing the use of automation because they go hand in hand; I can't do what I do without embracing automation because obviously for a start I've rammed it down the throats of people in the navy already, having been on after I left, and I've been involved in having it rammed down my throat when I first went to the fleet after training anyway" [N4]*

And:

> *"Well innovation in the military is driven from the top down rather than the bottom up like it is everywhere else. The people making the decisions about what, where are we going to go, are those experienced, because like human nature you're resistant to change, they will stick with what they know rather than embrace new technology. Whereas in the technical world it's the young innovators coming in that are driving the change, I don't think we are very, we don't embrace change, we tend to lag behind everybody else". [S5]*

Research has suggested that task autonomy and freedom of decision-making can reduce dispositional resistance (Battistelli, et al., 2013), however, both of these variables are assignments which are commonly reassigned to automated systems. From an organisational psychology standpoint, the restructuring of the operator role is a key concern in facilitating interaction and addressing maladaptive attitudes and behaviours towards resistance to change:

> *"I suspect there probably will be … Joe Bloggs who's writing the algorithm for one system, that person may be very experienced, very knowledgeable, understands the*

majority of the issues that may arise, and yet on another system that needs to interact, that software engineer hasn't got that level of experience or knowledge, whatever. So even then, what data you put in like RO said, it could cause conflicts, it probably will cause conflicts". [N2]

There is also resistance to change as a result of perceived and actual capability of system in use. As highlighted in Table 27 (See Table 31 in Appendix A (page 207) for full context), significant discussion is categorised around self-determination factors and system competence. More specifically, there was higher iterations of dialogues surrounding systems error concerning HSI than human error or technological error solely.

Table 27 - Interpersonal Salient Themes and Priority Weighting

Interpersonal		**71.4%**	**42.2%**		
All Self-Determination		71.4%	42.2%		
Self-Determination Only	2	0.8%	0.5%	1%	Low
Autonomy	2	25.5%	15.1%	44%	Low
Competence	2	31.6%	18.6%	54%	Medium
Relatedness	2	13.5%	8.0%	23%	Low
All Autonomy		25.5%	15.1%		
Autonomy Only	2	3.7%	2.2%	6%	Low
Hubris	2	6.1%	3.6%	10%	Low
Ownership and Authorship	3	5.4%	3.2%	9%	Medium
Personal Agency	3	3.2%	1.9%	6%	Medium
Role of Operator	3	7.1%	4.2%	12%	Medium
All Competence		31.6%	18.6%		
Competence Only	2	9.9%	5.8%	17%	Low
Risk	4	6.2%	3.7%	11%	High
System Error	5	7.4%	4.4%	13%	High
Human Error	5	4.9%	2.9%	8%	High
System Error	5	3.2%	1.9%	6%	High

A key theme of some of these discourses, especially in the Maritime domain, were around the competence of the inclusion of Commercial of the Shelf (CoTS) technologies in tandem with the concern about skill fade:

"The fact we're getting more equipment on board that's CoTS just compounds the problem even more." [S4]

"It's more engineering to make it compatible. It moves away from the intention in the first place, plug and play, at the minute it's anything but plug and play" [S3]

The competence of CoTS seems most apparent with submariners, and this may be in part due to the rhetoric explored in this chapter, that there is significant pride in expertise aligned with their military identity compared to their compatriots in the other military echelons. Many of the submariner participants relayed distress around skill fade as a pioneering organisation:

> *[S5] "But then when you're talking about the software, with automation it is not going to be what it needed to be, so we'll have to adapt it which means we'll have to change it and in doing that change, will it still work as it's meant to? Will there be something underlying that we won't be aware of until it all goes horribly wrong?"*
>
> *[S6] "Because that's not the design intent"*
>
> *[S5] Also there's the rate of development of that is by the time we decide to use this technology, by the time we actually got a working version of it on the platform, the industry has already moved on to something new, something else, and then we don't have the support to maintain what we've got, we're always going to be, at the moment at least, behind the curve, working with stuff that is, last year's technology, not what's coming in new, not the cutting edge anymore.*
>
> *[S6] Whereas we used to be the cutting edge.*
>
> *[S5] So we're not driving the technology, we're just trying to adapt.*

The issues surrounding the adaptation of CoTS/MoTs (Modified off the shelf) and the adaptation for use in a high-criticality domain, as proposed above, has significant human factors implications. For example, the discrepancies in the design process with equipment not fit for purpose for the high criticality domains. The reduction of reliability, capability and competence of the system can foster mistrust and under reliance, especially if the dissent is sewn by trusted leader figures such as SMEs. Furthermore, the perception of organisational 'cutting corners' through financial incentives to uses CoTs/MoTs may elicit organisational distrust. Research has shown that perceived Machiavellianism in

military environments can negatively impact generalised trust at individual and group levels, as unit cohesion is of significant value (Karakatsanis & Swarts, 2015).

The wider holistic view of the macroergonomic system, and the social units therein, is important for developing appropriate trust in systems. With trustworthiness assigned to unit leaders, this can have consequences on the ability of their prior experiences and perceptions to infiltrate the distributed team mental model via negative schema priming:

> *"Yeah, if people have had a bad experience in the past, but also when you've got it on board a submarine, if the automation has fallen down like what's happened before then that's going to negatively affect people's perceptions of it and then people start spinning discs about how poor it is and all the issues it has and things do tend to spiral a little bit as well. But when people start spinning discs about things, telling stories sorry. Perceptions can be very easily changed". [S1]*

> *"Used to do that thing where, I used to be a marine engineer officer, so I used to watch keep on the nuclear reactor and propulsion systems and we used to play that game during the quiet hours overnight, kind of what happens if that goes, if that failed right now what would we do? And then we used to do that with various different things, playing the game of what would you do if that went wrong?" [S1]*

People consume storytelling readily (Smith, et al., 2017) and narratives are central to human cognition and learning (Ramiller, 2020). They are a means to support social cooperation and dissemination of social norms, and further nurture community support. Therefore, if trustworthy, 'heroic' leaders create stories of their lived experiences the perceptions are likely going to be assimilated by their unit. With the rate of technological improvement ever increasing, and competence and enhancement of systems rising, these maladaptive worldviews and narratives may persist.

The next section explores converging and diverging narratives between the echelons of the military explored in this thesis. For example, some of the persevering impressions the holon culture has had on

them post-service, but also the perceived differences each echelon has about each other, and how this may affect organisational trust.

7.2.6 Cultural Narrative Convergence and Divergence

> *"I think it's variable isn't it, I mean, I suppose not so much the navy but if you look at different armed forces, I don't [pause] people get very focused and slightly frustrated at people in the military, they have this very set idea that we only follow the rules and it depends on what branch you're in so I think there'll be different effects with different people…"[N5]*

A significant amount of military culture literature observes the organisation from an integration perspective, whereby the values and assumptions are shared by all members of that social structure (Martin, 1992; Redmond, et al., 2015). However, as explored in Chapter 1 (section 1.2.2), culture is developed through the shared values, ideas and experiences of its members (Soeters, et al., 2006). Compared to the individualistic context of civilian western culture, the military domain is somewhat countercultural as a collectivist hierarchical organisation (French & McCain, 2004). In addition, the diversity of each echelon can have implications on the cultural narratives within them. For example, the roles, command structure and encounters can influence how the member's worldview is socially constructed. Furthermore, the workplace environment, variation in deployment and reintegration, and whether the culture is institutionally orientated, can impact the ethos and shared principles of the echelon.

From exploring the data sets, there is a distinct perceived hierarchy within HM Armed Forces. This is primarily attributed to perceived levels of funding and investment, for example:

> *"…You are in the pecking order because the Navy being the senior service, the RAF always have expensive equipment to keep their planes in the air, and the Army always gets third bite of the cake so to speak" [A4]*

"In the military, you're given, or you get what you're given, and you're always going to be behind, always. Now you've got to accept that in the military. Well I say that, in certain branches of the military, they get [whispered] The Red Arrows, they want for nothing, it's because it's a showcase." [R1]

"Yes special forces, want for nothing." [N3]

But also, there is an underlying expression of elitism, whereby those in more focused expertise domains receive preferential treatment:

"If you thought about a pilot, then actually they're going to be matriculated to death…. But if you think about a soldier, sailor, the level of investment is less in them…" [A1]

"I think again it will come back to what walk of life you're in. If you're involved in industry where there's a lot of automation I think you'd be quite happy with it, but I think the military, yeah I would probably say, not so much the Army, but the Air Force and the Navy." [S2]

Overall, there is a clear distinction felt by all participants that there is a cultural chasm between civilian and military personnel, especially with the introduction of automation and emerging technology. Participants S4, S5, S6, of which all have a long service record and subject matter expertise viewed submariners as less adaptive and sceptical of introducing new technology as a result of capability, reliability and competence:

"Numbers, there are thousands of aircrafts, tens of thousands of cars that will be used as testbeds gathering evidence, there's a handful of submarines, to get the same amount of evidence would take decades, so decisions are going to be made on the imitation of this, on a lot less evidence than there was in other industries." [S5]

As previously mentioned, especially with submariners, expertise is valued both as a safeguarding measure but also as a trustworthy leadership attribute. Participants S1 and S2 additionally opined that

expertise is a factor in whether unit members are open to adapting to automation or emerging technology, rather than resistant.

With regards to implementing emerging technology, one of the key differences that was noted is that the Army has a different viewpoint compared to the RAF and the Navy. The Army cohort viewed technology predominantly as tools due to the disposability and usage of systems within their experience:

> *"Ground forces are limited to what technology they can use anyway in a battlefield, because it's got to be cheap and it's got to be reliable and it's got to be something you can throw away. Whereas you look at the RAF from the Navy, their kit is supposed to last a long time, and when I say last a long time, you've got ships, submarines, you've got aircrafts, ground protection, so those are built to last. Whereas the Army, it's on the field, his feet on the ground, tanks are supposed to last but the expectancy is not very long." [A4]*

There is also the underlying sentiment, that some of the participants viewed themselves within the macroergonomic structure, similarly to that of these straightforward and austere tools:

> *"I need to know facts, I need to know numbers, I need to know the details, beyond that I don't need to know anything else. I've been trained to be ruggedized, I've been trained to be hardened, I've been trained to be capable in these environments, I don't need someone to keep telling me everything's is going to be alright, I need someone to tell me that the information is correct or the information is not correct…". [A3]*

However, in part to the simplistic black and white, binary views within the Army, the OF-5 participant proposed that this attitude is beneficial for technological adaptation compared to civilians as the chain of command can eliminate the allowance for doubt, and aid facilitation:

> *"…They're a tightly knit group of individuals, then if the new thing is a piece of kit, but actually it's got to overcome the scepticism of all of the individuals and the kind of, whatever the group view is of this sort of thing, you can easily see that they're quite*

enough, …. So you get a lot of bottom up adoption…there are some areas that have really jumped ahead;. Get them to do what they're really good at, which is experiment, play and adapt. ." [A1]

There are meaningful attributions of trust to the chain of command. These charismatic leaders (Cohen, 2010) can sublimate the burden of choice, through the predicable actions of their command chain, through the shared sacred values of that culture of which they are all strongly committed to. These relationships have an exceptionally high power imbalance (for example, high risk and threat to life), therefore can create total trust commitment with their followers. Therefore, it is important to observe that these leaders can establish and exploit this loyal social structure to impart support into facilitation of trust in technological systems and automated teammates. The trust mechanisms of power dynamics and trustworthiness can be used to aid in facilitation via schema priming by recapturing the narrative framing into trusting system teammates.

7.3 SUMMARY

The aim of this chapter was to explore research question 2 (RQ2): "How does trust in automation differ between military personnel and civilians?". Similarly to the previous chapter (Chapter 6), the main categories extracted from the semi-structured responsive interviewing were: (1) how participants define trust as a concept; (2) how automation impacts trust expressions; (3) how intra- and interpersonal factors impact trust at self, group, community and societal levels. Furthermore, there is an additional category which arose from the analysis, which were the convergence and divergence of themes and narratives between the different echelons of the military cohort interviewed

Participants primarily viewed automation as a tool, not a social actant. Thereby the emphasis on reliance, capability and culpability were significantly propagated with regards to whether they would trust the system. Furthermore, although not as heavily discussed with this cohort of participants, the unease and discomfort of artificial actants close to category boundaries were still apparent within the military personnel. However, the role of social cues and responses are far more apparent in the above

cohort. For example, the risk and culpability of trusting a system teammate was significantly more difficult to justify compared to their human teammates. This can be attributed to the intensity at which team cohesion, military identity and authoritarian structure influence the trust dynamics and social responses of the participants. Once more, the themes of out-group denigration and in-group favouritism bias are magnified due to the psychosocial development of worldview in military units:

> *"…because the trust in your shipmate is about a human choice as well, it's about shared values, shared experiences, sacrifice, that you know that the person next to you will go the extra mile for you, you know that he's fallible because he shares the same frailties as yourself, but you know what you would do for that person, and therefore you know what that person would do for you, and that's built on generations of tradition. It's not purely you, it's a corporate memory and its corporate values; you are part of an established organisation that is based on those values. [N1]*

Furthermore, the psychosocial attitudes and behaviours developed through service were observed as persistent post-service, regardless of time elapsed away from military culture. This was observed with all echelons, that the concepts of trust, inter-personal relationships and social interactions were somewhat crystallised in the intense environment that is warfare and the battlespace. There could also be a psychophysiological basis to the manifestation of the persistent behaviours through the effect of stress activation on neural pathways (Marek, et al., 2018):

> *"We got the fire alarm at [Redacted} that goes off at half past nine every Tuesday morning and I still make buttons every time because it sounded (sic) just like a manual flood alarm on board a submarine. No! Jesus Christ, honestly me and my colleague are both submariners and we both look at each other and [audibly comical sigh in relief]." [S3]*

The social structure of the military is both the obstruction and the solution for improving facilitation of automation through utilising the immense trust mechanisms already reinforced in the unit cohesion and heroic leadership in the organisational structure.

In summary, the conclusion of this chapter, is that the conceptualisation of trust is strongly associated with reliance, capability and culpability in terms of automation interaction. Similar dynamics are fostered in team cohesion, however the impact of the socio-cultural structure in the military has a significant impact on trust formation with substantial outgroup bias. The ‘othering’ and the distrust of system teammates, is of interest as social responses to this discomfort are also apparent with the civilian cohort. The consideration of these associations are further discussed in the subsequent Discussion chapter (Chapter 8).

8 DISCUSSION

8.1 INTRODUCTION

This chapter is divided into three subsections which address the primary outcomes of this research. The first section offers a discussion of the trust concepts explored and their impact on emerging technology. This is to address RQ1: How do the narratives that people tell about automation relate to their trust in technology.

The latter sections both discuss the research in terms of summarising RQ2: How does trust in automation differ between military personnel and civilians? In the first of these sections, the narratives from the civilian cohort are explored. In the second section the focus is on the military cohort and the inter-echelon variations.

8.2 TRUST, AUTOMATION AND AUTONOMY

As emphasised throughout this thesis, there has been a significant rise in research to address how the concept of 'trust' is adjusting to cope with the radically changing sociotechnical environment and ever-decentralising systems. Underlying this spread of research, however, is the problem of creating a definition of 'trust' that is sufficiently robust to allow it to be useful, and that captures lived experience of people exposed to these technologies. This thesis contributes to contextualising trust in sociotechnical systems through grounded participant experiences and narratives.

The following diagram is an amendment to the key concept diagram posited in Chapter 2. The following trust concepts and theories within the literature were expressed and supported by the narrative explored in this thesis.

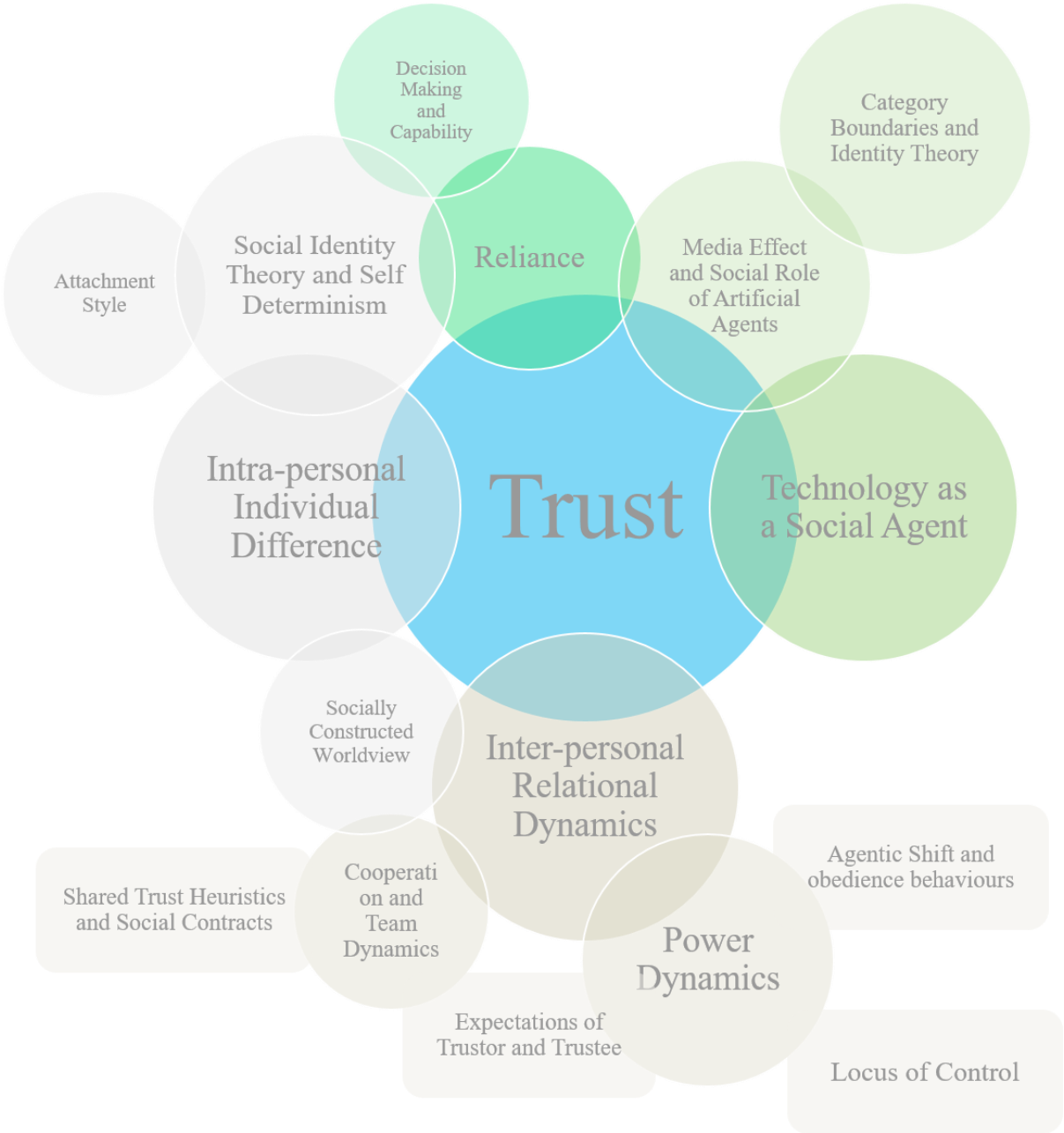

Figure 11 - Key Concepts of Trust Expressed in Participant Narratives

As outlined in Chapter 2, these concepts explore the multiplicity of trust definitions and approaches in a broad selection of literature. An example of this is with regards to exploring trust as synonymous with reliance which is used widely in TIA research. In both cohorts, there were shallow associations with reliance, such as:

"…if a machine has a definite higher reliability than a human then I think it is much easier to trust a machine than it is a human." [C1]

However, with introspection and discussion, these reflections belied one of the core interests of the research, which was to explore the underlying expressions of trust which are removed from reasoned behaviour:

"…So it could be like really reliable, there could be no risks, but you could still mistrust it, and that's irrational but valid I guess" [C2]

"Having to trust it and not being able to rely on it." [N2]

As proposed by Baier (1986) and Colquitt *et al* (2007), the concept of trust is closely aligned with trust propensity of the individual (trustee), the trustworthiness of the actant (trustor) and the stakes involved in the entrustment of the task (power dynamics). However, factors which can influence, such as reliance and capability, are not wholly reflected in trust mechanisms when interacting with system-teammates:

"I suppose because for years and years you've relied on it, and all your training was based on team training, but the reality is, automation is entirely logical, whereas teammates aren't." [N5]

The primary outcome of defining trust is that there is a significant underlying discomfort when interacting with artificial actants. The narratives reflected this regardless of reasoned decision making, reliability in the system as there was an underlying issue impacting integration within the group unit of the military cohort:

"Don't know why it gives you hnnuuughs [indicating gut feeling/discomfort] because he could be the worst [teammate] in the world, but it gives you confidence doesn't it…We tend to be less trusting of computers than we are of humans…" [A2]

The deeper, underlying tacit forms of trust were of interest in this thesis, and how the social environment shapes the construction of these schema. Reflected in both cohorts, the chasm between trust and reliance were explored. This can be applied to both human-teammates and system-teammates. For example, an

underlying narrative which had influence on trusting behaviours was the socially constructed compartmentalisation of in-group favouritism and out-group bias. An application of this to personnel, is as follows:

> *"…I think that's the area where there's no logic we apply to trust. I mean some of the MOD junior guys that have just come straight from training, actually we might not trust them, because we say they haven't got the experience, but if they've just came from training, they haven't got the bad habits, they haven't got the, oh just do it this way even though the instruction book says so…[N5]*

Similarly, with automation, if the actant was not identified as a teammate (or within-group) there were expressions of underlying discomfort:

> *"I would trust the guy next to me, over a bit of machine, especially if it was telling me something I didn't want to do, or there were other options, it might be the best option, but there are other options. For me on the ground, in Afghan, I would trust the person next to me over a piece of equipment." [A2]*

A dominant trust narrative through both cohorts is the relationship associating trustworthiness of an actant and trust formation. The literature explored in Chapter 2, which posits that trustees are associated with vulnerability and expectation, are imperative for defining trust concepts (Baier, 1986; Colquitt, et al., 2007). A collection of concepts from the literature were used to scaffold interpretation of participants concepts of trust, and subsequently culture and relational bonds.

One of the theoretical concepts of trust formation which was used to strengthen intrinsic self-determination factors and trust mechanisms is Social Identity Theory (Tajfel, 1979). The theory posits that an individual's self-concept is derivative of their perceived membership within a relevant social group. From this trust dynamics between in- and out- group members are impacted, as is tacit trust mechanisms (Tanis & Postmes, 2005). Trust narratives in this thesis were expressed with consideration to self- and group- identities. From this investigation, identity theory impacted trust in automation in two specific circumstances. First, if the automatous actant was perceived as the out-group, such as:

> *"… be that people naturally feel much more comfortable with the idea of trusting another human than with trusting a machine or a system, because they're another person like me, they think the same way that I do…" [C3]*

Furthermore, group identity is closely associated with individual identity. A key component of personal motivation and trusting behaviours are associated with self-determinism factors.. One of which is personal agency and autonomy. The existence of a human-like agent, whether in socio-physical form (e.g. anthropomorphisation) or cognitive process (e.g. artificial intelligence) is identified in this thesis as a category boundary which participants are uncomfortable in engaging in trust facilitation with:

> *"I think yes, yes you do have to trust your teammates; I think that's the first thing, yes you have to trust them and over a period of time you do trust them. With automation, I'm not convinced that you could trust automation at the same level as you would trust your fellow shipmates …" [N2]*

Anthropomorphism is discussed separately in the following subsection. However, regarding trust and threat to self-identity, multidisciplinary research has explored themes of category boundaries. Category boundaries are areas of perception between two categories which are on a spectrum that have classically been described as binary. This often results in cognitive conflict creating a strong, negative affective response (Mori, et al., 2012).

The field of Queer Theory, which explores the social construction of sexuality and gender, has explored at length how concepts beyond binary definitions, encounter resistance as some individuals perception of self is maintained by fixed and defined beliefs about their worldview (Garelick, et al., 2017). That these individuals are resistant to incorporating contradictory beliefs through their psychological need for closure. As presented in this thesis, how participants primarily defined automation as a tool, rather than a teammate is reflective of this compartmentalising. Furthermore, cognitive rigidity and flexibility have implications in the adoption or facilitation of technology. Barak (2018) found those with flexible thinking and less cognitive rigidity, were more likely to interact and integrate with technology than their unyielding counterparts. The implication proposed is that trust facilitation for newer recruits into the

military may be more malleable, whether as Digital Natives or the sociocultural visibility of non-binary representation.

One of the key trust constructs to evolve in this thesis, is the importance of perceived and actual power dynamics between actants (Baier, 1986; Luhmann, 1979; Kramer & Tyler, 1996). A power dynamic exists between trustor and trustee, where the latter is put into a position where they are vulnerable to the unknown intent or will of the trustor. As suggested above, othering and discomfort from ambiguous social responses from artificial agents play a role. An illustrative point which is perceived in both military personnel and civilians, is the following:

> *"I suppose there's potential for an element of mistrust in the level of intelligence, or perceived intelligence that avatars have and if they seem to be smarter than the user, and they're connected to the outside world and they can potentially feedback information on that person, I can see an element of mistrust from that perspective". [S1]*

This excerpt encapsulates the concepts of trust in relation to vulnerability and entrustment from a number of standpoints. Firstly, there is mistrust of the artificial agent through othering and is perceived as an outsider which can disrupt trusting behaviours from uncertainty and unpredictable expectations. Secondly, that there is discomfort associated with identity – both cohorts attribute qualities of their identity to expertise. As mentioned previously, self-determination factors are closely aligned with trusting behaviour (Chirkov, et al., 2003). Thirdly, that distrust of the artificial agent through the perceived impact of unpredictable consequences. That without a H-H social contract, this impacts the power dynamics and trust of the individuals. Subsequently, the threat of disciplinary action may have negative impact on trust behaviours at an organisational level.

This imbalance of power in the dyad of trust can influence facilitation with automation, the social roles and responses in the dynamic are important for the behaviour of the organisation. This theoretical standpoint and description of trust provides insight at the intra-personnel, inter-personal and community level. This therefore has implications for the literature regarding systems thinking and

macroergonomics. The interviews herein show support for utilising aspects of social identity theory and trust power dynamics to forecast trust in automation facilitation issues as they become more complex and are integrated in social units.

Weber (2003) suggests that as systems become more complex, the need for trust becomes greater. That components of trust, such as predictability and evaluation (Barber, 1983), will invariably change the locus of trust and associated mechanisms (Jones & Shah, 2016). For example, Chapter 7 explored the 'terrorist principle' (Bluhm, 1987) in terms of technological actants, through the lack of organisational transparency whether system teammates could covertly audit their colleagues. The lack of predictability and transparency required for entrustment or trustworthy behaviour associated with social responses, may impact future facilitation of trust in automation. The exponential growth and ubiquity of technology may cause psychosocial strains with both H-H and H-S interaction. Complex systems require somewhat opaque implementation of internal working for data management and information communication. For example, cognitive load is a major concern for operators:

> *"Thinking about information overload on people who haven't had highly specialist knowledge. If you thought about a pilot, then actually they're going to be matriculated to death, and you're really going to understand the limits of what a pilot might be able to, could or couldn't take…" [A1]*

The reduction in transparency to manage the strain on working memory capacity comes with greater dependence on these complex systems, thus impacting on power differentials in trust dynamics. In addition, themes surrounding automation error support the literature in that mistakes have lasting beliefs impacts on trustworthiness (Barg-Walkow & Rogers, 2016). Error reduces perceived reliability, capability and competence of the automation. Conversely, responses to human error did not incur the same disdain:

> *"Yeah I mean it's quite interesting, we've all talked about trust trust trust with automation and probably your instinct would naturally say I trust my teammates, but actually we all know they fail, they fall asleep, they miss things, they've been out the*

night before, and you've got very inexperienced, but we do tend to put more stock into human emotion." [N5]

Research into human error and trust find that vulnerability, through human flaws, can facilitate trust through the perceived trustworthiness of imperfection. This 'Pratfall Effect' of human error is perceived as an equaliser by reducing power dynamics through vulnerability to facilitate trust has been observed significantly in organisational and governmental leadership (Aronson, et al., 1966). This can additionally impact organisational behaviour within the team unit and cohesion when developing relational bonds and social contracts of trust between members.

8.3 CULTURAL SIMILARITIES AND DIFFERENCES

The current research into implementing anthropomorphism in high criticality domains such as Defence and Healthcare trend towards positive short-term trust facilitation. However, as suggested in this research, there is need for long-term implementation to manage sociocultural change in trust dynamics with artificial actants. For example, a significant amount of the research (refer to Chapter 3 for more information), utilise civilian participants (such as undergraduate students) and generalise the outcomes across domains and cultures. As presented in this thesis, this generalisation may yield inaccurate and potentially harmful recommendations. For example, a study which used university participants by De Visser *et al* (2016) recommend that "…*anthropomorphic automation is associated with greater trust resilience. Designers therefore should consider the incorporation of human-like features as a deliberate design choice".* Despite this, they note that social roles and responses with artificial actants may have implications on *"dependence, vulnerability and intimacy"*. Below is an example, of how introduction of inappropriate anthropomorphism may impact how operators interact with their system teammates, as well as further implications for organisational mistrust:

"If you know that something doesn't work and you think that by giving it a human face or a pet's face that you're going to overcome that then that's not overcoming it, that's concealment so…" [A1]

Shallow trust resilience at an individual level cannot be generalised to inter-personal and organisational behaviour. The surface trust resilience can be somewhat associated with Media Equation theory (Reeves & Nass, 1996). Whereby, humans as social beings will often react to humanlike responses and roles in a socially constructed manner. However, this may have negative repercussions in the long-term facilitation.

Outlined in this research, there are distinct differences in how trust is formed within the military cohort. For how critical these trust dynamics affect individual's identity formation and team cohesion, the suitability for these recommendations may have precarious outcomes. In addition, the risk involved in over-, under- or maladaptive trust or reliance in a system can have potentially tragic consequences in the high-risk environment that is the military. For example:

> *"…in a military environment it's very easy because you can tell someone to do something and they do it, takes a lot of the guess work out of situations that way if you tell someone to do something they do it…. because they're entrusted, so they have the responsibility, and when you have the responsibility of ten, one hundred, six hundred, a few thousand people's lives, even though you walk through it like it's nothing to worry about, because that's who you are at the time, unless you can confidently turn round and say yeah that's all gravy and we're all happy with that, the stress is unbearable, massively unbearable; and you have to have the ultimate faith that it's going to consistently work." [A3]*

The role of organisational responsibility, morality and accountability is concurrent between both civilians and the military cohort. Participants voiced hesitation and discomfort at the unpredictability in trust negotiations, whereby the artificial agent or technological actant has greater power in the dynamic as culpability is an unknown factor.

> *"I wouldn't have thought for a second that the guy sitting next to me wouldn't be able to do his job if something went wrong, I would just automatically know that he could do his job, and I'd be doing mine and he'd be doing his etc etc. I don't think it's ever*

crossed my mind that anybody wouldn't be able to do what they've got to do in a real emergency, that just wouldn't cross my mind." [S2]

However, a pertinent difference with the military cohort, especially those in the Ground Forces, were negated in part by organisational responsibility law (such as the Yamashita Standard (Parks, 1973; Lael, 1982; Rowland, 2003)). The command structure and the development of trusted leaders through sacred values (as discussed in Chapter 7, section 7.2.4) also play a role in the formation of how trust is conceptualised in these institutions. Military participants sublimated entrustment to the command structure and the Rules of Engagement which govern volatile interactions:

"…it's always the commanding officer's problem, not just the person who did the deed, it's everyone up the chain to the very top, and that's always bred in especially in the military environment, it's not one person's fault, you know, it's never one person's fault no matter what happens, it's everyone in that chain of command and everyone in that link…." [A3]

With regards to artificial actants, the discomfort towards artificial intelligence was far more pronounced in civilian narratives. For civilian participants, artificial intelligence and perceived cognitive processes in actants created discomfort through threats to identity of self via category boundaries (Mori, et al., 2012). As civilians exist in an individualistic culture, threats to self-identity are starker, for example:

"… I think when it comes to AI and it's starting to make decisions that question, that has its own morality, that's when it gets uncomfortable." [C2]

Whereas identity fusion of the collective identity within the military and their close units provides a more resilient and more rigid self-concept (Chirkov, et al., 2003; Swann, et al., 2014):

"in a military environment it's very easy because you can tell someone to do something and they do it, takes a lot of the guess work out of situations that way if you tell someone to do something they do it…. because they're entrusted, so they have the responsibility, and when you have the responsibility of ten, one hundred, six hundred, a few thousand

people's lives, even though you walk through it like it's nothing to worry about, because that's who you are at the time,." [A3]

Conversely however, when group identity was threatened, the intensity of expressions were more significant than the individual civilians. For example, expertise was a dominant narrative in the military, especially those in the Submarine service. Participants would admonish a complex system for failure yet have hubris in the capability of a human operators despite human error accounting for significant amounts of system failure across domains (Reason, 1990).

In reflection, a cultural difference and implication which may impact these underlying expressions is the punitive action in the authoritative institution, blame attribution in organisational structure and command responsibility. Furthermore, leadership and expertise entrustment is presented intensely in the military cohort which can be associated with team cohesion and collectivism which in turn can create underlying tension and discomfort when self- and group identities are threatened (French & McCain, 2004).

Overall, civilian participants introduced the narratives of in-group and out-group bias in terms of human and artificial actants. The military cohort subsequently magnified this theme in addition to introducing partiality dynamics between echelon culture. The next section discusses this premise further.

8.4 NARRATIVE PERSONAS AND INTERECHELON DIVERGENCES

The narratives explored in the analysis provide insight into their conceptualisations of trust and technology. The participant experiences also establish considerable culture related matters that can shape concepts and attitudes towards H-H and H-S teams. The dominant expressions presented in this thesis concerning cultural implications are those which surround the identity fusion with their echelon of HM Armed Forces and the impact this has on team cohesion and trust mechanisms.

As presented in the systematic literature review in Chapter 3, mental models are a technique used to explain or represent cognitive processes, such as decision-making and reasoning. Schemata are similar in that they are socially constructed explanatory models influenced by interactions, the individual's

environment and their lived experiences. Explanatory models and schema can guide perceptions, attitudes and behaviours. In the military cohort, a common denominator across the echelons, were how these are shaped by service experiences and comradeship. The latter is especially influential as construction of explanatory models of the self and group identity have origins with Attachment Theory (Bowlby, 1973; Ainsworth & Bell, 1970). Introduced in Chapter 2 (section 2.3) an individual's attachment orientation can shape their *"relational expectations, emotion, and behaviour"* (Fraley & Shaver, 2000). Research has sought to explore attachment styles and how they apply and impact military personnel and their responses in active situations (Caspi-Berkowitz, et al., 2019; Mikulincer & Shaver, 2007; Mikulincer & Florian, 1998). Table 28 outlines a working model of adult attachment adapted from Bartholomew and Horowitz (1991).

Table 28 - Working Model of Adult Attachment Styles (Adapted from Bartholomew and Horowitz (1991)

		Model of Self	
		Positive (Low Dependence)	*Negative (High Dependence)*
Model of Others	*Positive (Low Avoidance)*	Secure	Anxious-preoccupied
	Negative (High Avoidance)	Dismissive-avoidant	Fearful-avoidant

Emerging literature has explored the use of attachment orientations to study the unique comradeship and team cohesion in military personnel and how that might influence trusting and altruistic behaviours. For example, Yip *et al* (2018) observed that individuals with higher avoidance conveyed a more individualistic orientation, whereas dismissive avoidant participants were less likely to align themselves with groups (Rom & Mikulincer, 2003). In Defence observations, anxious-preoccupied personnel often would act as sentinels, with increased protective behaviour towards their unit and shared mental models (Ein-Dor & Tal, 2012). Furthermore, though who exhibited high dependence developed rapid protective

behaviour and schema (Ein-Dor, et al., 2011). The investigation into trust facilitation, found the social role and responses in developing entrustment and appropriate relationships are imperative.

As reflected in Chapter 7 with participants remarking the military family is *"one for all or all for one"* [S6] the importance of trust facilitation is far more than mediating reliability, as trust is a concept rooted in self-, group and philosophical identity. The relationships formed in trust dynamics and their underlying mechanism effect the individual beyond decision-making. As Caspi-Berkowitz *et al* (2019) proposes; *"…People scoring higher on attachment anxiety are likely to be more ready to self-sacrifice for a group or a cause, the ultimate test of their group investment…".*

Therefore, to illustrate dominant trust narratives towards automation, the following table (Table 29) adapts existing validated adult attachment styles (Besharat, 2011) to present examples of personas corroborated by the outcomes of the interpretative analysis of this thesis.

The 'Model of Self' categorisation codifies the relational dependence of the presented persona. For example, those with low dependence represent more individualistic identities whereby group cohesion is not prioritised in self-identity. Similarly, the modified "Model of Others" category is the relational interaction with artificial actants or automation. For example, those with low avoidance are more secure in their self-identity, or flexible thinking, that cognitive rigidity surrounding category boundaries has a lesser impact.

Table 29 - Table of Example Personas from Dominant Trust in Automation Narratives

		Model of Self	
		Positive *(Low Dependence)*	*Negative* *(High Dependence)*
Model of Others ***(Artificial Actants)***	*Positive* *(Low Avoidance)*	**Persona 1**: Often Civilian; readily accepting of technology with expressed positive sentiments; Exhibits mild hesitance, however supportive of LoA and integration of artificial actants if appropriate culpability and capability measures are accounted for; Mild discomfort at uncanny category boundaries, though this expression is not necessarily associated with threat to self-identity.	**Persona 3**: Often Military; adaption hesitance with regards to capability and culpability; will present positively if acceptance benefits team unity – for example, if H-S teaming will protect teammates shallow facilitation is possible; can be dependent on the trustworthiness presented by roles of responsibility and aligns with the shared explanatory model/schema of the unit. Mild threat to identity but has greater feelings associated with identity within the group as opposed to the individual self.
	Negative *(High Avoidance)*	**Persona 2:** Often Civilian; presents with cynicism and discomfort. Associated with threat to self-identity and cognitive rigidity. May logically or pragmatically reason acceptance of system teammate; shallow technological adaption; underlying expression is that of avoidance; may present with critique of trustworthiness of organisational responsibility and individualistic mentalities.	**Persona 4:** Often Military; hesitant to integrate with artificial actants; higher trust and reliance with human unit members.. Unit identity high, threat results in fearful-avoidant behaviour towards automation; additional threat to self-identity (such as expertise and skill fade); concerned with technological use and negative effect on social cohesion (socialising). Mild distrust at the systems-level with regards to organisational responsibility

In support of themes and narratives presented in this thesis, research has begun examining human attachment theory in long-term human-robot interaction (McDorman, et al., 2016). This and others seek to apply trans-disciplinary theories to explore human-robot interaction via social cognition and how it

can develop richer more robust methodologies (Collins, 2019). Furthermore, a growing number of researchers in the field of human-robotics/human-automation interaction are emphasising the utilisation of human-human trust literature. By grounding research in a solid foundation, this can provide advancement in knowledge into trust repair and trust violations and how complex social interaction can include social artificial actants (Baker, et al., 2018).

With Submariner participants there was a specific emphasis on skill fade and the importance of expertise with these individuals. This impacted on trust formation through inspection rituals and knowledge-based trustworthiness. For example, there was resistance to increasing levels of automation through both skill fade of their colleagues and issues regarding capabilities of the system:

> *"… unless you know where it is, you've always relied on a computer, because whereas before you went round manually moving a valve, moving a flap or whatever, computer says no and that's it, you just sit on your hands then, … unless you can build in any fail safes … or overrides… maybe it would work, but that's the human element, because if you've then got the expertise to decide" [S8]*

A secondary divergence in trust formation was that of the Ground Forces. Although there was emphasis on team cohesion, identity fusion and entrustment with all domains within HM Armed Forces; there were narrative differences in the Ground Forces that were not echoed in the Surface Navy or RAF participants. Army participants reflected on their experiences in combat and how their unit cohesion and concept of inter-personal trust was shaped by the high-risk environment they were deployed in:

> *".... Culture, ground forces are limited to what technology they can use anyway in a battlefield, because it's got to be cheap and it's got to be reliable and it's got to be something you can throw away. Whereas you look at the RAF from the Navy, their kit is supposed to last a long time, and when I say last a long time, you've got ships, submarines, you've got aircrafts, ground protection, so those are built to last.. …my life expectancy was something like about ten minutes on the battlefield if I was lucky…. Culture, yes definitely, culture differences." [A2]*

There were similarities of this narrative with the Submariner cohort. A suggestion to expound on why there were narrative convergences between these groups despite different combat environment, is that of mortality:

> *"If a ship, if the engine's failed on a ship, it still floats. If the engine fails on a submarine invariably it will sink, eventually, ends up with the fishes." [S5]*

The Ground Forces suffer higher death rates on deployment than their colleagues (Ministry of Defence, 2019). Although fatalities on Submarines are rare, the survivability within a sunken vessel is statistically higher than surface ship or airborne vehicles (Clayton, 2011). Historically, the losses of submariner crews were significant, and this zeitgeist of risk and loss may be persistent in modern Submariner narrative (Evans, 1986).

The literature and the analysis within this thesis suggest that the perceived and actual risk, sacrificiality inherent in their deployment and team cohesion can act as a trust facilitation binder. The similarities between these two echelons and the trust discourse within their worldviews are more convergent than their colleagues despite the opposing physical environments.

8.5 SUMMARY

The principle generalisations insofar are that, if trust is defined as positive expectations but negative experiences are oft used as example interactions in the data, can trust as a technology acceptance metric be appropriately used? A clear dissonance between verbalised sentiments and underlying expressions of trust attitudes was also observed across both civilian and military cohorts, more specifically exploring expressions of mistrust and distrust of technology which occur closer to category boundaries. This is important when considering the aspiration to integrate human analogues or human-human teammate mimicry in future C3I centres and Human-System Interfaces.

Within Civilian narratives, there is a specific focus on the morals and ethics of artificial intelligence and technological actants that are within the category boundary and are deemed invasive to human perceptions of self. Within the military cohort, these are mediated by the knowledge of rules of

engagement and the authoritative hierarchy of HM Armed Forces. The trust in the socio-culture and team cohesion play a role in this. One of the main findings of the research is with regards to power distribution and trust - with perceived power inequality affecting team cohesion and trust facilitation. This is seen more clearly in the military cohort as regimented structure of roles and power are somewhat institutionalised within the worldview and technological teammates do not readily fit with any schema - in addition to threat to self-determination factors, personal autonomy and identity fusion conflict.

The inter-echelon cultural difference are unique to the idiosyncrasies in their field. An interesting cultural outcome is the similarities between Submariners and their Army colleagues, based in lived experiences of high criticality and high-risk environments. A differentiation is the more informal authoritative structure and higher specialism, for example, the Artificer branch, within the subsurface echelon. Threats to self-determinism factors and pride in expertise is a major issue with technology facilitation.

Overall, trust in automation is inextricably linked to core trust mechanisms which are socially constructed. These include the development of relational trust through adult attachment styles, which are developed through lived experiences of human-human social responses. One of the major contributing factors affecting trust development and facilitation are related to power dynamics. Transparency and predictable outcomes are not yet able to be socially communicated by artificial actants. Tacit trust has technology adaptation issues regarding social schemata thus embedding negative models of technology, as well as cognitive conflict arising from category boundary discomfort arising from threats to self-identity. Furthermore, for military personnel social identity and identity fusion of being part of the ‘military’ family has far-reaching implications of how military personnel define trust, as well as impact their trust mechanisms and dynamics when required to trust an out-group actant such as a system teammate.

9 CONCLUSION

9.1 INTRODUCTION

> *"Cannot you see, cannot all you lecturers see, that it is we that are dying, and that down here the only thing that really lives is the Machine? We created the Machine, to do our will, but we cannot make it do our will now. It has robbed us of the sense of space and of the sense of touch, it has blurred every human relation and narrowed down love to a carnal act, it has paralyzed our bodies and our wills, and now it compels us to worship it. The Machine develops – but not on our lives. The Machine proceeds – but not to our goal. We only exist as the blood corpuscles that course through its arteries, and if it could work without us, it would let us die" (Forster, 1909, p. 20)*

The narratives of discomfort, skill fade and inappropriate reliance in technology, in automation, in artificial agents have persisted seemingly unchanged for a century. The story of *"The Machine Stops"* by E.M. Forster explores themes such as parasocial relationships affecting social roles and responses; over-reliance in competent systems; and the roles of personal autonomy and threats to self-determinism factors. This is not the first story to explore technology resistance, nor will it be the last. These are narratives embedded in the cultural zeitgeist[14] (Kaplan, 2004). As reiterated throughout this thesis, humans are natural storytellers, and have used narratives to convey knowledge, social responses, and complex themes for millennia. To expect mediating decision-making tactics such as increasing reliance, competence, fidelity and transparency to provide crystalline changes to social schema and explanatory models is naïve. Especially with regards to the rate of change with emerging technologies and computing capability (Moore, 1965; Neven, et al., 2019).

[14] In western culture. Cross-cultural dimensions affecting HAI and artificial actants are an area of growing interest. The following selection are an introduction to cross-cultural differences between Western cultures such as the UK and Japan (MacDorman, et al., 2009; Sone, 2016; Šabanović, 2014; Bolton, 2002; Bolton, et al., 2007).

This thesis has explored trust expressions and how it is conceptualised when interacting with human and artificial actants. By utilising both a civilian cohort and diverse pool of military personnel, the convergence and divergences of these cultures could be investigated. The research presented contributes to the literature into both trust and automation facilitation (RQ1). Furthermore, explores how the unique sociotechnical system of the military and its individual echelons can transmute these trust concepts (RQ2). This insight offers scope for exploring the fragmented and limited knowledge of military narratives and how they affect the development of persistent social schemata. The nature of these relationships is crucial for developing appropriate trust, both in human and system teammates. This thesis furthermore presents conceptualisations of military culture within, and between, the echelons. Whereas narrative inquiry has been somewhat explored in the literature, the lived experiences of the Submarine service is exceptionally limited further.

The remainder of this chapter explores the theoretical and practical implications of the research. It additionally states limitations incurred in this research and presents suggestions for future work.

9.2 SUMMARIES ABOUT RESEARCH QUESTIONS

Research Question 1 (RQ1): Facilitating Trust

1) How do the narratives that people tell about automation relate to their trust in technology?

The narratives of both civilian and military participants conceptualise trust align with mechanisms explored in the trust literature, though aligns closer to social psychology and philosophical interpretations. This is contrary to the primary concept of reliance, and its associated behaviours, used extensively throughout TIA research (Lee & See, 2004). The most pointed observation in this thesis, is how inter- and intra-personal relationships impact trust dynamics. In addition, that category boundaries and cognitive rigidity play a noteworthy role in social responses, colour worldview and impact underlying expressions of discomfort and distrust.

Research Question 2 (RQ2): Military Culture:

2) How does trust in automation differ between military personnel and civilians?

Although concurrent themes and narratives are woven into both civilian and military narratives, the emphasis is substantially different. For example, civilian narratives are more unsettled by the morality and justice of legal accountability. Whereas, although this is a concern in military narratives, a significant portion of this is mediated through the trustworthiness of the heavily regulated regime and inspection rituals of their authoritative hierarchical system. The commanding structure, organisational responsibility and unit cohesion provide reprieve of a number of these concerns. However, the impact of individual and group bonding effects trust mechanisms and dynamics significantly. The collectivism of the military culture, and how it plays a central role to the individuals and group identity is of considerable interest. Furthermore, the role in which technology, and the subsequent expertise to interact with complex systems, further alters these dynamics. The types of technology, risk and environment each have roles in the formation of fluid and crystalline trust subtleties.

9.3 THEORETICAL AND PRACTICAL IMPLICATIONS

Trust research has been conducted extensively with university student participants and military cohorts in the literature (refer to Chapter 3: Systematic Literature Review). Through exploring the trust expressions within both of these institutional structures, there is a noteworthy difference. With military participants, trust mechanisms are developed in an authoritative social structure with an intense attribution of team cohesion and inter-personal relationships, compared to the individualistic freedoms of civilian academia. There are several implications with this outcome. The most critical of those being the generalisation of trust mechanisms and associated decision-making inferences from civilians to defence domains. The Defence sector will most likely have the most significant increase in levels of automation and ubiquity of decision-making technology. Facilitation of trust in automation is required for operator safety and mission success. By applying solutions generalised to a different socio-cultural structure may have subtle, but potentially catastrophic implications.

There are common themes between both cultures, as discussed in the previous chapter, however the significance of military culture on the development of explanatory models and shaping social responses is of concern. Trust in automation research is generalised trans-disciplinarily, which may not be

appropriate as the concepts of trust differ. There are practical implications with facilitation, as reliability and confidence in the system are only mediating to a point. This may be exacerbated with the inevitable introduction of artificial actants, system teammates and anthropomorphism in emerging interfaces for Human-Machine teaming (Ministry of Defence, 2018).

Through use of interpretative phenomenological analysis, the investigation of trust themes explored concepts of trust in a way which participants may not have been aware of (Lyon, et al., 2015). The use of qualitative methods such as verbal protocol reduces the imposition of existing concepts to participants. In addition, this methodology utilises participant lived experiences and their sensemaking of how trust is defined. Exploring the myriad of how participants construct trust has allowed for theory development by weaving diverse literature. In telling stories, their narratives naturally surfaced giving colour and texture to their social reality and provided a foundation to build upon meaning-making.

Through use of whole quotes, implementing minimal editing or reductions, supports genuine participant representation through maintaining integrity and retaining authenticity of participant voices. For narratives and exploration of worldview, it was important to preserve the conceptualising reflected by participants, especially the lived experiences of the military cohort on their culture and the sociotechnical system. Furthermore, the use of the documentary analysis case study and exploratory study provided interpretations grounded and bounded in the literature and theory, without actively interrupting the verbal freedom from responsive interviewing techniques.

This research has explored how trust and reliance are associated, but that they are semantically different concepts. This has implications for academic and organisational research. As outlined in the Military Study (Chapter 7), there is significant value of employing the existing rigid trust dynamics and mechanisms in the leadership construct and inter-personal relationship bonds of military personnel. Team cohesion is one of the militaries greatest assets and therefore should be nurtured, not restricted or mediated. Technology acceptance and adaption can be supported from a micro-ergonomic level and can affect change through the macroergonomic structure. For example, Zak (2017) proposes that organisation with high-trust continually outperform those with low-trust. These impact productivity,

innovation, engagement and retention. These are currently factors under scrutiny within the Defence sector and therefore can have significant implications on the functionality of the domain going forward.

9.4 LIMITATIONS

A limitation of this research, and similarly with trust research overall, is that tacit forms of trust are difficult to isolate as they cannot necessarily be explicitly communicated verbally. For example, participants may have rationalised their interpretations of trust interacting in the verbal protocol of the interviews and focus groups. Although IPA was used to aid in retrieving underlying expressions in the conversational analysis, the cognitive processes and emotions of the participants can never be truly known.

For qualitative research, generalisability is often cited as a limitation. For this research this has been appropriately addressed and justified in the methodology section (Chapter 4). However, a limitation which has occurred is the subpar representation. Although the research made attempts for participant heterogeneity, and the respondents were somewhat representative as both institutions and domains explored in this thesis are androcentric (STEM and Military), there could have been better gender representation. Although these themes were not explored in the research, there were gendered difference in participant interaction within the focus groups. However, the heterogeneity of the participants and focus groups were mostly representative with a diverse age, service and rank of participants included.

With regards to limitations in terms of methodology, IPA is focal for immersion within the phenomena explored in the type of research in this thesis it is limited by the context interviews were situated and thus can be reinterpreted differently. Furthermore, due to academic limitations, the auditability was reduced in this specific thesis. Although addressed in the methodology regarding authenticity, and the theoretical or philosophical implications of ‘perfect data’ having reductionist overtones. Additional verification and auditing would have provided the research presented in this thesis with additional validation.

Overall, the focus in this thesis has been on the lived experiences of participants and extraction of how meaning-making of trust is affected by worldview. Furthermore, how culture shapes their social reality and trust responses. The aspiration is that the despite the limitation presented, the overall credibility of the thesis can provide welcome themes to build upon the knowledge base within the literature and provide new avenues for exploration.

A final comment, which is not a limitation but a noteworthy inclusion into this section, is that of persistent framing issues with technical jargon. As noted in Section 3.6 and Section 5.3.1, the contextualisation of phrases and concepts within the broad scope of the interdisciplinary research remains an issue. For example, the academic non-consensus of the term Mental Model. In this research, the term 'explanatory model' was applied to describe the narrative or storied experience of participants with technology. This differs slightly with Mental Model, as this is often attributed with detailed explanations of naturalistic decision-making in Human Factors and Ergonomics pedagogy.

9.5 IMPLICATIONS FOR FUTURE RESEARCH

As proposed by Möllering (2001), trust research should not be *"subsumed under decision-making and exchange theories"* and that *"interpretation captures the idea that human experiencing of the life-world gives basis for trust*". That trust research "*calls for reflexive, hermeneutic approaches, both conceptually and empirically*". (Möllering, 2001, p. 417)

This research has been presented as exploratory, to examine at a macro-level the diversity of culture in the military. To propose that each of the microcosms have subtly different trust concepts, and underlying narratives regarding team cohesion can impact attitudes towards technology adaption. However, due to the vast size of the data sets, it was difficult to tell the specific narratives of each echelon at length. To answer the proposed research questions, and provide information specifically with a focus on Submariners, other narratives were not reported in the same depth. The discourse and narratives of participants in the different echelons provided different discourse. Especially how the Ground Forces were described by others, but also by themselves comparatively to the Sea and Air

Forces. Future work into HM Army and HAI resistance or team cohesion would be of substantial interest. Furthermore, the leadership and skill fade themes in Submariners is of specific interest in HAI and human-human teaming. This would be of significance to explore further, especially with the concern of personal technology, reduction of centralised group socialisation and increased isolation with colleagues disrupting social cohesion. This could yield more information regarding trust facilitation in future command, control and intelligence centres.

9.5.1 Key Points

With regards to how to translate the findings from this thesis into generalisable recommendations, this section outlines some key points and interpretations which could be implemented in future work or research. These are as follows:

- Civilian Studies (Extended): As the sampling of the civilian cohort were primarily subject matter experts and highly educated individuals, it would be of interest to explore the expressions of trust and attitudes towards emerging technologies in a range of ages, education and socio-economics stasis levels. For example, the participants focused on matters such as legality and morality, which were not explored in much depth in the comparative case study. Furthermore, highlighted by Study 1 and a participant from Study 2, there is trust related issues on a macro-organisational or governmental level regarding responsibility and job replacement. The impact of social class, unemployment and community responsibility may be of interest in exploring barriers to facilitation and technology implementation with civilian narratives.
- Armed Forces Studies (Ground Forces; Expanded): As noted in section 7.2.6 especially, there are some interesting narratives and expressions of trust and technological facilitation or adaptability in the experiences of HM Ground Forces. Those in the Army displayed similar narratives regarding unit cohesion and team identity as the Submariners, though there were interesting nuances that could not be explored in this thesis. The has been recent exploratory studies into the lived experiences of those in the Army, in addition to the issues exiting the forces (see Section 1.2). Similar to the key point above, the class or socio-economics status of

those entering and within the Ground Forces may provide some interesting narratives. This may be in part attributable to the lower eligibility requirements into HM Ground Forces (Education/Fitness) compared to the other echelons (UCAS, 2020).

- Training: This thesis supports the doubts and criticisms of implementation of technology which utilises superficial human-like attributes as a method of human-system interaction or human-system interfacing. Therefore, it can be proposed that the use of AMVE and novel interfaces should be used with caution. Furthermore, training can be effected by the stories those leading the sessions tell, and may prime the explanatory models of the trainees. Reframing of technology and utilising the impact unit leaders have on the team for improving human-system relations is key. In addition, reporting positive interactions, rather than solely negative experiences can have an impact on operators socially constructed heuristics.
- User Interface: Similar to the above point, the use of anthropometrics and human-like mimicry should be used, only when the social roles and responses have been appropriately considered. Avatars can be useful in communicating multimodal information; however, the underlying prejudices or bias of the operator will need to be contemplated. Designers need to ask themselves, "Why am I implementing this?" – if the social or cognitive psychology of the user, or the contexts in which it is going to be used, is not adequately considered, there will remain a barrier towards appropriate facilitation.
- Design and Evaluation: As explored in the literature review chapters, and findings reported in the thesis of trust definitions with participants, the role of evaluation of trust is of interest. Reliance is used extensively in TIA research, especially in the evaluation of design and interfaces. However, as explored in this thesis, the capability of the system and the amount of trust given are not in parallel. Therefore, the over-dependence of reliance-led psychometrics to analyse Trust is a problem. By using qualitative research in conjunction with quantitative methods may provide more representative outcomes of the internal processes for naturalistic design making.

Page Left Intentionally Blank

10 APPENDIX A – CONTENT WEIGHTED OUTCOMES

Table 30 - Content Weighted Outcomes for Civilians

	Salient Themes		
Category / Theme / Sub-Theme	**Absolute**	**Relative**	**Content Weighting**
Cultural	**59.4%**	**27.3%**	
All Culture	27.8%	12.8%	
Culture Only	7.5%	3.4%	13%
Identity	2.6%	1.2%	4%
Age	0.9%	0.4%	1%
Technology Acceptance	6.5%	3.0%	11%
Media Influence	5.3%	2.5%	9%
Zeitgeist	5.0%	2.3%	9%
All Terminology	15.2%	7.0%	26%
Terminology Only	0.8%	0.4%	1%
Automation	3.2%	1.5%	5%
Chaoplexic Definitions	3.9%	1.8%	7%
Trust	7.2%	3.3%	12%
All Schema Priming	16.4%	7.5%	28%
Schema Priming Only	4.2%	1.9%	7%
Contraditions	6.5%	3.0%	11%
Negative Experiences	5.6%	2.6%	9%
Positive Experiences	0.0%	0.0%	0%
Interpersonal	**105.0%**	**48.3%**	
All Self-Determination	105.0%	48.3%	
Self-Determination Only	0.0%	0.0%	0%
Autonomy	42.5%	19.6%	72%
Competence	39.3%	18.1%	66%
Relatedness	23.2%	10.7%	39%
All Autonomy	42.5%	19.6%	
Autonomy Only	18.0%	8.3%	30%
Hubris	2.8%	1.3%	5%
Ownership and Authorship	5.8%	2.7%	10%
Personal Agency	1.2%	0.6%	2%
Role of Operator	14.8%	6.8%	25%
All Compentence	39.3%	18.1%	
Competence Only	8.8%	4.1%	15%
Risk	5.0%	2.3%	8%
System Error	11.4%	5.3%	19%
Human Error	8.1%	3.7%	14%
System Error	5.9%	2.7%	10%
All Relatedness	23.2%	10.7%	
Relatedness Only	4.1%	1.9%	7%
Organisational Responsibility	11.8%	5.5%	20%
Sociocultural Schemata	7.2%	3.3%	12%
Automation	**18.6%**	**8.6%**	
All Automation	8.0%	3.7%	
Automation Only	8.0%	3.7%	14%
All Antropomorphisation	10.6%	4.9%	
Antropomorphisation Only	10.6%	4.9%	18%
All Perceived Permanency	0.0%	0.0%	
Perceived Permanency Only	0.0%	0.0%	0%
Trust	**34.1%**	**15.7%**	
All Trust	34.1%	15.7%	
Trust Only	19.0%	8.8%	32%
Capability	4.9%	2.2%	8%
Decision Making	4.4%	2.0%	7%
Reliability	5.8%	2.7%	10%

Table 31 - Content and Priority Weighted Outcomes for Military Cohort

Category / Theme / Sub-Theme	Consequence	Salient Themes: Absolute	Salient Themes: Relative	Content Weighting	Priority
Cultural		**58.1%**	**34.3%**		
All Culture		26.4%	15.6%		
Culture Only	1	6.2%	3.6%	11%	Low
Identity	2	4.3%	2.5%	7%	Low
Age	2	4.1%	2.4%	7%	Low
Technology Acceptance	3	4.5%	2.6%	8%	Medium
Media Influence	1	4.0%	2.4%	7%	Low
Zeitgeist	1	3.5%	2.1%	6%	Low
All Terminology		16.8%	9.9%		
Terminology Only	1	0.0%	0.0%	0%	Low
Automation	1	3.9%	2.3%	7%	Low
Chaoplexic Definitions	1	5.1%	3.0%	9%	Low
Trust	1	7.7%	4.6%	13%	Low
All Schema Priming		14.9%	8.8%		
Schema Priming Only	1	6.7%	4.0%	12%	Low
Contraditions	2	2.3%	1.4%	4%	Medium
Negative Experiences	2	4.5%	2.6%	8%	Medium
Positive Experiences	1	1.4%	0.8%	2%	Low
Interpersonal		**71.4%**	**42.2%**		
All Self-Determination		71.4%	42.2%		
Self-Determination Only	2	0.8%	0.5%	1%	Low
Autonomy	2	25.5%	15.1%	44%	Low
Competence	2	31.6%	18.6%	54%	Medium
Relatedness	2	13.5%	8.0%	23%	Low
All Autonomy		25.5%	15.1%		
Autonomy Only	2	3.7%	2.2%	6%	Low
Hubris	2	6.1%	3.6%	10%	Low
Ownership and Authorship	3	5.4%	3.2%	9%	Medium
Personal Agency	3	3.2%	1.9%	6%	Medium
Role of Operator	3	7.1%	4.2%	12%	Medium
All Compentence		31.6%	18.6%		
Competence Only	2	9.9%	5.8%	17%	Low
Risk	4	6.2%	3.7%	11%	High
System Error	5	7.4%	4.4%	13%	High
Human Error	5	4.9%	2.9%	8%	High
System Error	5	3.2%	1.9%	6%	High
All Relatedness		13.5%	8.0%		
Relatedness Only	2	0.5%	0.3%	1%	Low
Organisational Responsibility	4	8.9%	5.3%	15%	Medium
Sociocultural Schemata	2	4.1%	2.4%	7%	Low
Automation		**19.2%**	**11.3%**		
All Automation		19.2%	11.3%		
Automation Only	2	11.6%	6.9%	20%	Low
All Antropomorphisation		3.2%	1.9%		
Antropomorphisation Only	2	3.2%	1.9%	6%	Low
All Perceived Permanency		4.4%	2.6%		
Perceived Permanency Only	1	4.4%	2.6%	8%	Low
Trust		**20.6%**	**12.2%**		
All Trust		20.6%	12.2%		
Trust Only	3	6.1%	3.6%	11%	Medium
Capability	4	3.8%	2.3%	7%	Medium
Decision Making	5	6.6%	3.9%	11%	High
Reliability	4	4.1%	2.4%	7%	Medium

11 APPENDIX B – ETHICS BASED SUPPLIMENTARIES

PARTICIPANT INFORMATION SHEET (BASIC)

PROJECT PROPOSAL

Megan Field
Primary Investigator

PROJECT TITLE

The Facilitation of Trust in Automation: A Qualitative Study of Behaviour and Attitudes Towards Emerging Technology in Military Culture

DESCRIPTION

The purpose of the study is to understand the attitude and behaviours of people across both civilian and military domains towards emerging technology and automation. The project is seeking to explore the 'narrative' in these different subcultures and how they differ because of personal experiences, acquired knowledge and social influences. The goal of the research is to understand "why" and what effects these differences have on people's interaction with future technology and their trust in automation. The outcome will be used to provide support in training and policy in submarine military culture and help facilitate effective and appropriate human-system interaction

OBJECTIVE

Current research of trust in automation has reached an impasse through data saturation as research often focuses on physiological measures in response to training based simulations or use of metrics which are not necessarily appropriate for the diversity in human cognition. Therefore this study seeks to explore the underlying expressions and narratives in a more immersive depth than qualitative or mixed method studies.

This study is part of a doctoral research project supported by the University of Birmingham and BAE Systems.

WHAT WILL I BE ASKED TO DO?

You will simply be asked to engage in a 60-90 interview (either solo or with another participant) where we would like to explore your life experiences, thoughts towards emerging technology and automation. The more open you are, the more enhanced your data will be. This will take place in a private and appropriate room. This will be audio recorded. The personal information disclosed in the research collected is at your discretion. Qualitative research is enriched by the personal aspect of the stories told in them and all the information collected will be kept as anonymous and confidential as possible. After the interview concludes, you will be asked to fill out a short form to inform demographics.

PARTICIPANT INFORMATION SHEET (FULL) AND CONSENT FORM

Study Title

The Facilitation of Trust in Automation: A Qualitative Study of Behaviour and Attitudes Towards Emerging Technology in Military Culture

What is the Purpose of the Research?

The purpose for the study is to understand the attitude and behaviours of people across both civilian and military domains towards emerging technology and automation. The project is seeking to explore the 'narrative' in these different subcultures and how they differ because of personal experiences, acquired knowledge and social influences. The goal of the research is to understand "why" and what effects these differences have on people's interaction with future technology and their trust in automation. The outcome will be used to provide support in training and policy in submarine military culture and help facilitate effective and appropriate human-system interaction.

Who is Doing This Research?

The lead researcher specialises in qualitative research and is seeking to explore the psychosocial aspects influencing trust in automation. Current research of trust in automation has reached an impasse through data saturation as research often focuses on physiological measures in response to training based simulations or use of metrics which are not necessarily appropriate for the diversity in human cognition. Therefore, this study seeks to explore the underlying expressions and narratives in a more immersive depth than qualitative or mixed method studies. The researcher hopes to uncover aspects which can then influence further research (whether qualitative or quantitative) and training protocol in military culture. At the core of trust in automation, there is the human component – and maladaptive thoughts and behaviour cannot be accurately observed with quantitative methods and survey metrics.

This study is part of a doctoral research project supported by the University of Birmingham and BAE Systems.

Why Have I Been Invited to Take Part?

You have been invited to take part as you have met the inclusion criteria which are as follows:

- Participants can either be:
    - A civilian member of the university (staff or student) who has an interest in emerging technology and/or automation. These individuals are not restricted by age or student status.
- OR:
    - An individual of the military. This participants are not limited by age or status though must

be familiar with command and control/operations. A diverse range of participants are sought, and experience pre- and post- service are recommended.

- OR:
  - An individual from BAE Systems who has preferably with maritime or naval experience (whether from active duty or working in the domain as part of BAE Systems). This participants are not limited by age or status though must be familiar with command and control/operations.
- AND:
  - Technologically literate (for example, a digital native or immigrant).
  - Participants must be between the ages of 18 and 30 – however if age diversity is required this can be expanded outwards up to approximately 60 (so as within working age)

Do I Have to Take Part?

No. This study is entirely voluntary.

What Will I Be Asked to Do?

You will simply be asked to engage in an approximately hour long interview where we would like to explore your life experiences, thoughts towards emerging technology and automation, so the more open you are, the more enhanced your data will be. This will take place in a private and appropriate room (for example, on the UoB Campus for civilians, or a meeting room at BAE systems). This will be audio recorded. The personal information disclosed in the research collected is at your discretion. Qualitative research is enriched by the personal aspect of the stories told in them and all the information collected will be kept as anonymous and confidential as possible. After the interview concludes, you will be asked to fill out a short form to inform demographics.

What are the Benefits of Taking Part?

In taking part, you will be able to reflect on your experience, attitudes and behaviours towards emerging technology and automation which can provide insight into your working environment as well as your academic or military working culture.

What are the Possible Disadvantages and Risks of Taking Part?

There are no risks in taking part, however, if you have concerns about the study, please contact the main investigator for any comments or concerns you may have. Please contact the lead investigator if you feel you

have been impacted negatively by this research. Concerns will be addressed by the primary investigator and if required, the lead supervisor who is British Psychological Society accredited.

Can I Withdraw from the Research and What Will Happen If I Withdraw?

Yes! After you have read this information and asked any questions you may have we will ask you to complete an Informed Consent Form, however if at any time, before, during or after the time 'interview' period and you wish to withdraw from the study please just contact the main investigator. You can withdraw at any time, for any reason and you will not be asked to explain your reasons for withdrawing.

If you would like to rescind your data, this is possible for 12 months post-interview. After this point, data collected from your interview may have already been published in a journal article or thesis. Therefore, would not be possible to redact your anonymised information.

Are There Any Expenses and Payments Which I Will Get?

Due to the qualitative nature of the study, any monetary incentive used in recruitment and throughout data collection would be deemed unethical. The standards of the 1947 Nuremberg Code (The Nuremberg Code, 1949) state that no persuasion or pressure of any kind should be put on participants (Fry, et al., 2005) (Alderson & Morrow, 2011). Incentive payments can be seen as coercive – or as exerting undue influence on potential participants' decisions about whether to take part in research.

Whom Do I Contact If I Have Any Questions?

If you have any questions, please contact the lead researcher.

Whom Do I Contact If I Have a Complaint?

If you have any complaints, please contact the lead researcher.

Will My Records Be Kept Confidential?

Data will be confidential and anonymous. The information will be kept for approximately 12 months or until the completion of the Doctoral Thesis before destruction of original documents and however, data may be retained for up to 10 years as per the University of Birmingham Research Charter rules and regulations.

Who is Organising and Funding the Research?

EPSRC and BAE

Who Has Reviewed the Study?

This study has been reviewed and given favourable opinion by the lead supervisor and support of BAE Systems.

Further Information and Contact Details

Name: Megan Field

Address: [redacted]

E-mail: [redacted]

Supervisor Contact Details: Professor Bob Stone; [redacted]

Email: [redacted] (Tel: [redacted])

Compliance with the Declaration of Helsinki

This study complies, and at all times will comply, with the Declaration of Helsinki [15] as adopted at the 64th WMA General Assembly at Fortaleza, Brazil in October 2013.

Title of Study: The Facilitation of Trust in Automation: A Qualitative Study of Behaviour and Attitudes Towards Emerging Technology in Military Culture

Please Initial or Tick Boxes

☐ **The nature, aims and risks of the research have been explained to me. I have read and understood the Information for Participants and understand what is expected of me. All my questions have been answered fully to my satisfaction.**

☐ **I understand that if I decide at any time during the research that I no longer wish to participate in this project, I can notify the researchers involved and be withdrawn from it immediately without having to give a reason. I also understand that I may be withdrawn from it at any time, and that in neither case will this be held against me in subsequent dealings with the Ministry of Defence.**

☐ **I consent to the processing of my personal information for the purposes of this research study. I understand that such information will be treated**

[15] World Medical Association Declaration of Helsinki [revised October 2013]. Recommendations Guiding Medical Doctors in Biomedical Research Involving Human Subjects. 64th WMA General Assembly, Fortaleza (Brazil).

as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

☐ **I agree to volunteer as a participant for the study described in the information sheet and give full consent.**

☐ **This consent is specific to the particular study described in the Information for Participants attached and shall not be taken to imply my consent to participate in any subsequent study or deviation from that detailed here.**

☐ **I understand the compensation arrangements that have been provided.**

Participant's Statement:

I ……

agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Information for Participants about the project and understand what the research study involves.

Signed: **Date :**

Investigator's Statement:

I ……

confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the Participant.

Signed: **Date :**

12 APPENDIX C – SUMMARY TABLES FOR SYSTEMATIC REVIEW

Table 32 - Summary Table of Included Sources (Human Related // Ability Based)

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
Aydoğan, Sharpanskykh, & Lo (2014). (Aydoğan, et al., 2014)	Proposed computational agent-based model of SA to integrate SA with trust (interpersonal). Case study to validate applicability of the model. (Bruce (2011). Understanding decision-making processes in airline operations control.)	SA of agent represented by a belief network (mm) and observed or communicated information (in addition to context) activate nodes within this network.	N/A	The Netherlands (EU)	Case Study	Trustworthiness of information sources has a statistically significant influence of decision makers SA.
Birkmeier, Korn & Flemisch (2011) (Birkmeier, et al., 2011, October)	Full study protocol can be found in Marcus et al (2011) (*Operational Feasibility Of Sectorless ATM*)	Proof of concept discussion on sectorless air traffic management and the changes to ATCos mental models and SA.	N = 8	DLR German Aerospace Centre – ATCo simulation test bed (TRafficSim) Germany (EU)	Verbal protocol and UX questionnaire	Due to the high LOA, trust in automation and fidelity of the systems decision making/conflict detection is key for improved HIS and HMM. Issues of deskilling and over-reliance due to LOA/out-of-the-loop. Noted levels of retroactive memory interference with ATCos affecting HMM and SA.
Chua, Storey & Chiang (2012)	2 independent raters mapped subjects' verbal protocol to a coding scheme. Transcripts were analysed using the coding	Use verbal protocol and the application of the theory of mental models to empirically determine a	15 Recruited, N= 14 Ps total.	Adult Engineer (high skilled) population	Repeated Measures	Inter-rater reliability was at acceptable thresholds (Kappa=0.832); Correlation between simple counts and average

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
(Chua, et al., 2012)	scheme. Coding then reached statistically satisfactory inter-rater agreement and reconciled coding differences. A problem behaviour graph (PBG) was generated the coding results to depict the thought processes of subjects task.	basis for understanding: "(1) *how knowledge engineers extract domain knowledge from textual sources; and (2) the cognitive mechanisms by which they engage various knowledge representation schemes to represent that knowledge acquired*".	Age and Gender N/A	North America	Protocol Analysis Verbal protocol Problem Behaviour Graph generated from Qual/Coded transcripts	percentages = 0.750. Aggregate level correlation is 0.980.
Fallon, Murphy, Zimmerman & Mueller (2010) (Fallon, et al., 2010, May)	N/A	Proposal of a sense-making theoretical framework to describe trust calibration with automation.	N/A	N/A	Experiential User Guide	Sense-making is used as a tool for operators to understand and re-frame automation errors/human error to aid in appropriately trusting the system.
Hawley, Mares & Giammanco (2005) (Hawkley, et al., (2005). .)	Review	Discussion and evaluation of factors affecting future C2 with increasing LOA through loss of SA and skill decay.	N/A	N/A	Review	Misappropriation of function in both Human and System tasks often create unreasonable task which do not form a coherent set. These residual tasks result in fragmentation of residual function for which human operators cannot suitably frame within their mental models. Furthermore, likely to cause erroneous behaviour if attributed to automation, in addition to potential under/over-load.
Hoffman & Woods (2011) (Hoffman & Woods, 2011)	Discussion/Proposal	The use of fundamental trade-offs issues within macrocognition in work systems can help facilitate improved human performance and system	N/A	N/A	Discussion	Laws of Cognitive work explore the multifaceted approach and influences of learned trust (mental models and SA). The fundamental bounds in human/system performance as discussed in this paper are

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
		capacity in complex systems/HSI.				as follows: bounded ecology, bounded cognizance, bounded perspectives, bounded responsibility, bounded effectiveness. The aforementioned laws are covered within the bounds. For example. Stretched systems is entailed by bounded ecology; Reductive Tendency is entailed by bounded cognizance. The trade-off model can be used in future to assess performance and adaptive capacity of macrocogntive work systems.
Lim, Dey & Ayrahami (2009, April) (Lim, et al., 2009, April)	Online user experience study with a 3 stage procedure. *Stage 1:* Learning *Stage 2:* Understanding (a. Fill in the blanks sections; b. reasoning test) *Stage 3*: Survey.	Different types of explanations may result in changes to user understanding and perceptions of the system (provide appropriate mental models)	N = 55 (of the 39 who completed demographic statements, 51;49%; 29.8)	North America	Tukey HSD Pair-wise. Task performance: (time complete-ness and correctness) User understanding: (correctness, level of detail and guess/unintelligible)	The overall mental model/perceptions of the system and explanations did not have any statistically significant difference. However, Ps in transparent conditions (Why/Why Not) produced more partially correct reasons compared to the None condition ($F[1,50]=27.4$, $p<.001$). The Why condition also produced more fully correct answers than the Why Not/None conditions $F[1,50]=10.8$, $p<.002$).
Lo, Pluyter & Meijer (2015) (Lo, et al., 2015)	Within subject experimental design (high- and low-experience groups) with a Human-in-the-loop simulator with two scenarios.	Explore the impact of skill/expertise variation levels on operator's goals and mental models and the impact on performance.	N = 22 (18%;82%) 8 Excluded in the duration of the experiment.	The Netherlands (EU)	Operator goals measured with performance indicators (Popova & Sharpanskykh, 2011).	There was no statistically significant in the diversity of completion strategies within the high-/low-experience groups. However, more experienced Ps prioritised critical maintenance (e.g. free track order, unplanned stops) as more important that

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
	Scenario 1: Light disruption to train traffic flow *Scenario 2:* Moderate to severe disruption to flow. Operator goals assessed through self-reporting questionnaire, followed by enacting the two 40 minute scenarios outlined above.		Train traffic controllers with previous experience at Utrecht Central Station (capped at 10 years for tech/mgmt. system implementation).		Simulator logs were analysed by SMEs (knowledge probes and video analysis) to assess strategic mental models (knowledge of what and how, applied to a context).	low-experience Ps. Whereas low-experience Ps attributed arrival punctuality to good performance.
Neerincx et al (2008) (Neerincx, et al., 2008, January)	4 Scenarios with reasonable workload which include DRI of threats to contribute to workload. Two scenarios were developed for air/surface warfare in high-tension peace-enforcing situations. The other 2 were situated against a civilian smuggling background. Paper reports core functions and their preventatives for out-of-the-loop problems.	To evaluation a Combat Management System (CMS) prototype built and for adaptive identification design	4 participants (warfare officers WO) and 4 participants (warfare officer assistants AWO)	Royal Netherlands Navy (EU)	Automation mode 2 variables, adaptive or fixed (fixed comparable to lowest adaptive mode). 7 DVs: Subjective Workload, Workload, SA, quality and timeliness of actions, performance and communication	Specifying adaptive automation modules helps to reduce workload and maintain adequate situation awareness during critical naval missions.
Piccinini, Rodrigues, Leitão & Simões (2015) (Piccinini, et al., 2015)	Matched pairs of participants were to drive on a simulated route twice (with ACC vs non-ACC). ACC condition was preceded by both prior information (manual), training and practice with the simulation with the ACC. A critical	To investigate the impact of the driver's mental model of the ACC and of the driver's trust in the system to compare behaviour at critical events with ACC.	N=26. 100% M. 13 ACC users (age 42.2 (SD=9.9), 13 non-ACC users (age 26.7 (SD=9.9).	Faculty of Engineering at the University of Porto. (Portugal, EU)	Mixed design ANOVA	A negative behavioural adaptation to ACC was observed as a result to the critical situation (mistrust with ACC, over-reliance and misunderstanding of ACC capabilities) and negative large correlation between the driver's mental model of ACC operation in the critical situation and the safety margins

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
	simulation (collision) was presented in both conditions.					maintained by the ACC users during the same situation. Increased risk of collision in the ACC condition compared to manual for both groups. Trust was not an influencing factor.
RTO/NATO (2007)	Literature Review/Report	The report is to address key human factors issues surrounding human-system integration and augmented, mixed and virtual environments (AMVE) in both realistic and abstract environments and military conditions	N/A	N/A	N/A	Military education is a key outcome of this report. Specifically, embedding training into operational equipment to support appropriate and contextual mental modes of technological functionality and capability. The report also covers important human factors integration with technology and systems regarding cognitive limitations, capability, situational awareness and other factors related to AMVE.
Wilkinson, Frisk & Rogers (2007) (Wilkison, et al., 2007, October)	The navigational task (with automation in the form of a route recommendation aid) experiment was a 3 (acquisition level: high, low, and no) x 2 (accuracy of automation support: 70% vs. 100%) mixed design. The critical measures were response time and subjective ratings of trust, confidence, and team identity.	To explore how varying levels of user mental model quality influence performance with an imperfect automated system.	12 Undergraduate participants (58%;42%;18-30)	Georgia Institute of Technology (North America)	N/A	Due to the small statistical power, no significant conclusions can be drawn so only trends in the data collected are discussed. Low acquisition groups with 70% automation had the higher rates of misuses and disuse with the automated aid. No acquisition, although had a high number of misuse errors, did not have disuse errors comparatively. Conversely, both high/low groups tended to overestimate the accuracy of the automation compare to the no acquisition group. The main outcome is that

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
						high acquisition primed the mental model of the task and automation most appropriately with greater mission success and detection of automation failure.
Zhang, Kaber & Hsiang (2010) (Zhang, et al., 2010)	A Virtual Locomotion Environment (VLE) simulator with a physical moving treadmill was used during a DRI task. The three proposed mental models of event structuring/counting were homomorphic, complex homomorphic and isomorphic.	To assess an empirical approach into characterisation of mental models in multi-tasking scenarios.	12 (all male participants); 25 (SD=2.4)	North Carolina University (North America)	Cognitive Task Analysis (NASA-TLX); MANOVA; Logistic Regression.	SA response patterns, confidence ratings and task workload were examined against hypothesised mental models. Overall, the study concluded that SA may reflect the structure and content of an operator's mental model in dynamic environments where maintaining SA is a continuous process.

Table 33 - Summary Table of Included Sources (Human Related // Characteristics)

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
Arkin, Ulam, & Wagner (2012). (Arkin, et al., 2012)	N/A	Use human mental models and psychology to influence the design of future military automated robotics/ systems with moral traits (e.g. compassion, guilt, deception) for ethical and moral automated decision making.	N/A	Georgia Institute of Technology (North America)	Modal Logics – (Partner modelling, Independence model matrices)	We have already described formal methods for the generation and enforcement of such ethical constraints and have demonstrated their feasibility for military applications
Beggiato & Krems (2013) (Beggiato & Krems, 2013)	Longitudinal driving simulator study with repeated measure mixed method design. Matched sample approach.	Divergent mental models about trust in automation in ACC. *"how different initial mental models of ACC affect system trust and acceptance over time, as well as how a user's mental model evolves with experience "*	51 (51%;49%; 24)	The fixed-base driving simulator (STISIMDrive 100w) at the Chemnitz University of Technology Germany (EU)	ANOVA five BFI factors	*"System trust and acceptance show an effect of initial information: More potentially critical situations presented in preliminary information, the less the initial trust and acceptance of the system is, after reading the description. While the over-informed incorrect group shows the lowest scores for trust and acceptance in the beginning, the incomplete group tends to trust and accept the ACC the most, initially."* Initial information has an enduring effect on trust and acceptance of ACC. *"The more cognitive effort needed to update the mental model, the lower trust and acceptance".*

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
Beggiato, Pereira, Petzoldt & Krems (2015) (Beggiato, et al., 2015)	Exploratory longitudinal study. On-road field test, repeated measures design. Drivers without prior ACC experience to drive pre-prepared route, repeated 10 times using ACC across a 2 month period.	Investigate the learning process of trust and acceptance of automation in human mental models with ACC.	15 (47%:53%:28 (SD=1.82))	The 37km route was identical for each P, all drove the same type of vehicle (BMW525d) Germany (EU)	Mental model questionnaire developed by the authors -6 point Likert scale. 12-item unidimensional scale for trust in automated systems SPSS: Factor Analysis (Cronbach Alpha), ANOVA.	Learning and development of trust/acceptance of automation follows the power law of learning. However, limitations not experienced are not activated within the mental model. M = 5.84, (SD = 0.39, mean values from session 5 to 10), power function explains 44% of the variance in the data (R2 = .443)
Bruemmer, Gertman & Nielsen (2007) (Bruemmer, et al., 2007)	Protocol analysis (UX verbal protocol wit simulation). SAR vehicle with high LOA – user provides system with attention hotspots and the UAV navigates to these hotspots without the operator driving the vehicle. UI reduces needs for panning camera (operator workload) by constant focus on target area regardless of distance from robot. 3 condition design (joystick (low LOA) mouse (higher LOA with auto assistance), high LOA with mouse utilising target mode and	Exploring metaphors/simplified approaches for priming trust in HHM for HIS in novel/emerging interfaces.	153	Idaho National Laboratory (INL) (North America)	N/A	'Push back' and disparity between autonomous intentions (increasing with LOA). Focus on human capabilities rather than robot/system capabilities. 20% increase in locating targets with icons condition with 90% increase in camera use thus improving task performance. Utilises cognitive mapping which can facilitate trust in automation can be communicated effectively through narrative or symbolic means. For example, the interface design should build upon a metaphor that supports the interaction (e.g. trust in automation/ human-animal trust comparatives).

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
	visual hotspot UI). , randomly generated start point in locating targets within a maze (hidden).					

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
Bunt, Lount & Lauzon (2012) (Bunt, et al., 2012, February)	Qualitative (Exp1: Interview study; Exp2: Diary study) Exp1: Ps recruited through volunteering for low-level computer usage course. Verbal protocol using ubiquitous IIS systems such as Windows Start/Google Suggest and other low level recommenders) Exp2: Two week diary study. First phase, semi-structured pre-interview, followed by 14-day field study with on-line diary entry for each IIS used on each specific day, followed by a post-interview at the conclusion.	How transparency and trust with Intelligent Interactive Systems (automated decision aids) affect HMM.	Exp1: 21 (67%;33%) Exp2: 14 (50%;50%) Majority university students, though a number of Ps were employed (ranging from data mgmt., childcare, civil service)	University of Manitoba (Canada)	Exp1: Contextual Inquiry and inductive analysis	Exp1: Transparency on why certain items were recommended was useful (e.g. YouTube/Amazon) but other sites with more sophisticated search functions, this was not the case (google search and streamlining usage). Trust affected by limited transparency were commercial motivation is noted (e.g. Amazon) or lack of sophistication (YouTube may recommend same song, multiple versions). Transparency (or explanations) required for understanding anomalous behaviour to increase trust. Thought transparency is not as important as fidelity in the system increases. Explanations were desired when (1) to understand poor or inconsistent behaviour; (2) a desire for improved interactions (even if that meant disabling the feature); (3) wanting to understand atypically good behaviour.
Dehais, Causse, Vachon, & Tremblay (2012) (Dehais, et al., 2012)	UGV scenario with interface designed to trigger special hazards within the experimental scenario (e.g. failure) within a DRI task. Heart Rate (stress level), eye tracking (attentional focus)	Conflict in human-system interaction is a pre-cursor to degradation of human operator performance and misappropriated attentional focus.	13 Mean age = 27.84, SD = 6.53 Mean level of education = 17.15, SD + 1.86. Experience with operating robots.	Institut Supérieur de l'Aéronautique et de l'Espace (ISAE) France (EU)	ANOVA	Perseveration behaviour was noted in ~70% of the participants in reaction to the 'autonomous surprise'. In those perseverative Ps, HR progressively increased indicating continued psychophysical stress in comparison to the non- perseverative Ps who's HR returned to baseline. Perseverative Ps also had inappropriate fixation and did not glance at the battery status (imperative cue on the

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
						UV) whereas all non-perseverative Ps did. Fur her miscomprehension of the situation led perseverative Ps to focus on the panoramic video in an excessive manner. **NB:** Related to mental models additionally through psychosocial attitudes towards goal attainment and sociotechnical demands on the operator in addition to future use of 'cognitive countermeasures' in adaptive automation.
Groom & Nass (2007) (Groom & Nass, 2007)	Critique of current literature regarding trust in automation in human-robot teams (systems and robotics used fairly interchangeably).	Current understanding of Human-Robot (System) does not recognize the diversity and psychosocial aspects of human teams, rather H-R (H-S) teams are 'partnerships' in part due to lack of shared mental models.	N/A	N/A	Literature Review	Asserts that research focus has been geared towards mimicking HMM in robots to allow for better team cohesion, which can be problematic in appropriating trust (and high expectations of team behaviour) as it is implausible that automation with successfully manifest 'humanness'. While group and individual mental models are structured similarly in human teams, a robotic/system agent cannot develop a HMM. In high-risk situation (where many systems/robots have been constructed to replace humans) trust is critical to succeed in both achieving the goal and as teammates (to protect their welfare and interests).
Hancock *et al* (2011)	Meta-Analysis	Evaluation of the effect of human, robot and	N/A	N/A	Meta-analysis	Correlation of $r = +0.26$, with an experimental size of$d = +0.71$.

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
(Hancock, et al., 2011)		environmental factors have on perceived trust with H-R interaction.				Robot (system) related factors (such as performance and attributes) had the greater association with trust, followed by a moderate association with environmental factors, then little evidence related to human-related factors.
Hoff & Bashir (2015) (Hoff & Bashir, 2015)	Systematic Literature Review	Explore the existing research on factors influencing trust and human-automation interaction with human operators.	N/A	N/A	Systematic Literature Review protocol	There are three layers of variability which can affect human-automation trust. These are dispositional, situational and learned trust. The latter is well supported by mental model (initial learned) and situational awareness (dynamic learned) research.
Hoffman, Johnson & Bradshaw (2013) (Hoffman, et al., 2013)	Literature Review	Explores the differences between interpersonal trust and trust in machines/intelligent systems.	N/A	N/A	Literature Review	Research into trust with HSI often utilises human-human trust attributes which may be inappropriate as technological/system failures are more impacting. However, mental model frameworks are consistently similar in predicting human behaviour with HSI (i.e. re/establishing trust)).
Lee, Lau, Kiesler, & Chiu (2005, April) (Lee, et al., 2005, April)	Ps told the aim was to investigate how people communicate with robots and make judgments of a robot. The robot, they were told, was equipped with various speech recognition and speech production functions. It could understand English, Cantonese, etc. It could answer questions posed in speech or typing.	Origin of robot and language (Hong Kong/Cantonese & New York/English) would create different mental models. Expect Ps to infer both robot scenarios would have greater knowledge	Exp 1: 60 (68%;32%;21.15) Exp 2: 48 (69%:31%:21.35)	Hong Kong; University of Hong Kong; Student Population. (Asia)	MANOVA	PS estimated knowledge of the robot based on what they knew about people. HK Familiar 83% vs HK Unfamiliar 48% (F [1, 28] = 132, p .05) New York Familiar 76% VS New York Unfamiliar 55% (F [1, 28] = 61, p < .05) Condition X Familiar versus Unfamiliar to New Yorker interaction (F [1,28] = 17, p < .05). Ps thought the robot was more likely to identify landmarks that were familiar to

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
	Ps saw four groups of landmarks: landmarks familiar to people from both cultures, landmarks; familiar to people who live in the U.S., landmarks familiar; to people who live in Hong Kong, and landmarks unfamiliar; in both cultures.	of famous landmarks than obscure.				people living in Hong Kong (F [1, 41] = 110, p < .0001 and landmarks familiar to New Yorkers, (F [1, 41] = 58, p< .0001. Also, there was a significant Condition X Familiar versus Unfamiliar to New Yorker interaction (F [1, 41] = 9, p < .01) New York, they estimated the robot to be 80% likely to know the landmarks that were familiar to New Yorkers and just 61% likely to know the landmarks that were unfamiliar New Yorkers (t [20] = 6.6, p < .05). Hong Kong similar but smaller differentiation than in the US condition (80% for familiar landmarks and 71% for unfamiliar landmarks, t [22] = 7.1, p < .05).
Nachtwei (2011) (Nachtwei, 2011)	Review	Guidance and recommendation for designers to plan and implement allocation of function appropriately using demonstrative analogies of operator characteristics type.	N/A	N/A	Literature Review/Qualitative	The review isolates twelve analogies of human operators in control centres. These narrative types both illuminate potential mental model approaches and can be used to demonstrate appropriate vignettes for designers.
Nothdurft,, Lang, Klepsch, & Minker (2013) (Nothdurft, et al., 2013, April)	Preliminary Study Web-based simulation of a nuclear power plant control room with a virtual	To explore how different goals of explanation (such as justification and transparency) influence the bases of trust in	60 initial applications; N=48 valid subjects were used when male/female Ps were evenly distributed (50%;50%;23-35)	Germany (EU)	SPSS Control received no explanation compared to experimental group.	The experiment utilised explanations to align mental models of participants in order to achieve a particular objective or goal via mapping explanation goals to appropriate mental models of Human-Computer Trust.

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
	anthropomorphic assistant with university students	Unexpected failures in Human-Computer Trust (HCT)				**No significant data between control and experimental group**. Slight statistical differences between male and female Ps with explanations.
Oleson, Billings, Kocsis, Chen & Hancock (2011) (Oleson, et al., 2011, February)	Meta-Analysis	To support the continuing body of research on influencing trust factors between humans and their robotic teammates.	N/A	N/A	Meta-Analysis	Factors identified can be categorized as human influence, machine influence and environmental influences. These factors play an important role in influencing operator mental models in addition to design element recommendations to facilitate human-robot teammate trust. However, the meta-analysis outlines a primarily human based perspective of trust and team cognition.
Olson, Fisk & Rogers (2009) (Olson, et al., 2009, October)	The experiment compared two levels of automation reliability (70% and 100%) for an automated navigation aid with the computer based stimuli (simulated city map for task environment) and participants were grouped into three mental model accuracy groups (low, moderate, and high) post-hoc.	To understand the implications of interactions such as system knowledge and reliance between on forming appropriate HCI trust mental models between older adults and automation.	N=19 (47%;53%) ;60-80	Georgia Institute for Technology (North America)	N/A	Small sample size did not allow for the data to has adequate statistical power to provide significant data results. Therefore only trends are discussed. A trend for both disuse and misuse were observed in the 70% accuracy variable for low-acquisition participants. An interesting trend however was with Ps with low accuracy mental models committed more misuse errors, trusted the automation when the navigational aid was wrong, more often than participants with a moderate or highly accurate mental model. This additionally supports other studies literature regarding age and trust in automation.

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
Sander, Oleson, Billings, Chen & Hancock (2011) (Sanders, et al., 2011, September)	Meta-Analysis	To develop a theoretical model for human-robot trust	N/A	N/A	Systematic Review, Meta-Analysis	The meta-analysis has produced an outline for a theoretical model which includes human, robot and environmental factors. The updated model also includes training and design implications. This paper is in conjunction with Hancock, et al (2011) (Hancock, et al., 2011).
Shaefer et al (2014) (Schaefer, et al., 2014)	Meta-Analysis	To expand upon current literature on facilitating human-machine teaming.	N/A	N/A	Systematic Review, Meta-Analysis. 30 articles included in the meta-analysis.	Experimental effect of trust was +0.48 with a correlational effect of $\bar{r}$ = +0.34. Moderator effects of human-related (g = +0.49; $\bar{r}$ = +0.16) and automation-related ($\bar{g}$= +0.53; $\bar{r}$ = +0.41) factors were examined. Environmental effects could not be moderated due to lack of studies.
Talone, Phillips, Ososky & Jentsch (2015) (Talone, et al., 2015, September)	Between subject design with two conditions (individual perception/expectation of a robotic agent, and a 'Husky' agent – image of UGV provided to Ps) in a C2 tasks of a robotic agent.	To investigate human mental models and expectations of autonomous ground systems under various combinations of commands, constraints and environmental factors.	100	Undergraduate students from the University of Central Florida (North America)	Between subject and Two-Way ANOVA	Mental models of four commands, four command constraints and five environmental features were evaluated. This experiment is exploratory, so final results have not been published. However, study design and discussion are relevant to this literature review.

13 REFERENCES

Abe, G. & Richardson, J., 2006. Alarm timing, trust and driver expectation for forward collision warning systems.. *Applied Ergonomics,* 37(5), pp. 577-586.

Ainsworth, M. D. S. & Bell, S. M., 1970. Attachment, exploration, and separation: Illustrated by the behavior of one-year-olds in a strange situation. *Child development,* pp. 49-67.

Ajiferuke, M. & Boddewyn, J., 1970. Socioeconomic Indicators in Comparative Management. *Administrative Sciences Quarterly,* pp. 153-158.

Akerlof, G. A., 1997. Social distance and social decisions. *Econometrica,* Volume 6, p. 1005–1027.

Allport, G., 1937. *Personality: A psychological interpretation..* Newark: Holt.

Alvesson, M. & Sveningsson, S., 2003. Managers doing leadership: The extra-ordinarization of the mundane. *Human relations,* 56(12), pp. 1435-1459.

Andersson, J., 2010. *A conceptual model for analysis of automation usability problems in control room settings.,* Göteborg, Sweden: Department of Product and Production Development (Division Design & Human Factors) Chalmers University of Technology.

Archer, J., 2012. *Effects of Automation in the Aircraft Cockpit Environment: Skill Degradation, Situation Awareness, Workload,* West Lafayette: School of Industrial Engineering, Purdue University.

Arkin, R., Ulam, P. & Wagner, A., 2012. Moral decision making in autonomous systems: enforcement, moral emotions, dignity, trust, and deception.. *Proceedings of the IEEE,* 100(3), pp. 571-589.

Arksey, H. & Knight, P., 1999. *Interviewing for social scientists: An introductory resource with examples..* s.l.:Sage Publishing.

Arksey, H. & O'Malley, L., 2005. Scoping studies: towards a methodological framework. *International journal of social research methodology,* 8(1), pp. 19-32.

Aronson, E., Willerman, B. & Floyd, J., 1966. The effect of a pratfall on increasing interpersonal attractiveness. *Psychonomic Science,* Volume 4, pp. 227-228.

Asch, S. E., 1946. Forming impressions of personality. *The Journal of Abnormal and Social Psychology,* 41(3), p. 258.

Atkinson, D., Friedland, P. & Lyons, J., 2012. *Human-Machine Trust for Robust Autonomous Systems.* s.l., s.n.

Atran, S., Sheikh, H. & Gómez, Á., 2014. For cause and comrade: Devoted actors and willingness to fight. *Cliodynamics,* 5(1), pp. 41-57.

Aydoğan, R., Sharpanskykh, A. & Lo, J., 2014. A Trust-Based Situation Awareness Model.. *Multi-Agent Systems,* pp. 19-34.

Baier, A., 1986. Trust and antitrust.. *Ethics,* 96(2), pp. 231-260.

Baker, A., Phillips, E., Ullman, D. & Keebler, J., 2018. Toward an Understanding of Trust Repair in Human-Robot Interaction: Current Research and Future Directions. *ACM Transactions on Interactive Intelligent Systems (TiiS),* 8(4), pp. 1-30.

Baker, S. E. & Edwards, R., 2012. How many qualitative interviews is enough?. *National Centre for Research Methods.*

Barak, M., 2018. Are digital natives open to change? Examining flexible thinking and resistance to change. *Computers & Education,* Volume 121, pp. 115-123.

Barber, B., 1983. *The logic and limits of trust.* New Brunswick, NJ: Rutgers University Press.

Bar-Cohen, Y. & Hanson, D. T., 2009. *The coming robot revolution: Expectations and fears about emerging intelligent, humanlike machines.*. New York, NY: Springer.

Barg-Walkow, L. H., 2013. *Understanding the role of expectations on human responses to an automated system.,* s.l.: s.n.

Barg-Walkow, L. H. & Rogers, W. A., 2016. The effect of incorrect reliability information on expectations, perceptions, and use of automation. *Human factors,* 58(2), pp. 242-260.

Barr, P., Stimpert, J. & Huff, A., 1992. Cognitive change, strategic action, and organizational renewal.. *Strategic management journal,* 13(1), pp. 15-36.

Bartholomew, K. & Horowitz, L. M., 1991. Attachment styles among young adults: a test of a four-category model. *Journal of Personality and Social Psychology,* 61(2), pp. 226-244.

Bartone, P. & Kirkland, F., 1991. Optimal Leadership in Small Army Units. In: *Handbook of military psychology.*. s.l.:John Wiley & Sons, pp. 391-409.

Basit, T., 2003. Manual or electronic? The role of coding in qualitative data analysis. *Educational research,* 45(2), pp. 143-154.

Bastian, B., Jetten, J., Thai, H. & Steffens, N., 2018. Shared adversity increases team creativity through fostering supportive interaction. *Frontiers in psychology,* 9(2309), pp. 1-10.

Battistelli, A., Montani, F. & Odoardi, C., 2013. The impact of feedback from job and task autonomy in the relationship between dispositional resistance to change and innovative work behaviour. *European Journal of Work and Organizational Psychology,* 22(1), pp. 26-41.

Bauman, Z., 1993. *Postmodern ethics.* Oxford: Blackwell.

Becker, H., 1998. *Tricks of the trade: How to think about your research whilst doing it.* Chicago: University of Chicago Press.

Beede, D. N. et al., 2011. *Women in STEM: A gender gap to innovation,* s.l.: Economics and Statistics Administration Issue Brief,.

Beggiato, M. & Krems, J. F., 2013. The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information.. *Transportation research part F: traffic psychology and behaviour,* Volume 18, pp. 47-57.

Beggiato, M., Pereira, M., Petzoldt, T. & Krems, J., 2015. Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study.. *Transportation research part F: traffic psychology and behaviour,* Volume 35, pp. 75-84.

Bernard, H., 2000. *Social research methods.* Thousand Oaks: Sage.

Bertaux, D., 1981. From the life-history approach to the transformation of sociological practice. In: *Biography and society: The life history approach in the social sciences.* London: Sage, pp. 29-45.

Besharat, M., 2011. Development and validation of adult attachment inventory. *Procedia-Social and Behavioral Sciences,* Volume 30, pp. 475-479.

Billings, D. R. et al., 2012. *Human-animal trust as an analog for human-robot trust: A review of current evidence (No. ARL-TR-5949),* Orlando: University of Central Florida .

Binks, E. & Cambridge, S., 2018. The transition experiences of British military veterans.. *Political psychology,* 39(1), pp. 125-142.

Birkmeier, B., Korn, B. & Flemisch, F. O., 2011, October. *First findings on the controller's mental model in sectorless air traffic management.* s.l., s.n., pp. 6C1-1.

Bluhm, L., 1987. Trust, terrorism, and technology.. *Journal of Business Ethics,* 6(5), pp. 333-341.

Bolton, M. L., Bass, E. J. & Siminiceanu, R. I., 2013. Using formal verification to evaluate human-automation interaction: A review.. *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* 43(3), pp. 488-503.

Booher, D., 2015. *What More Can I Say?: Why Communication Fails and what to Do about it.* s.l.:Prentice Hall Press.

Bousquet, A., 2008. Chaoplexic warfare or the future of military organization.. *International Affairs,* 84(5), pp. 915-929.

Bowen, G. A., 2009. Document analysis as a qualitative research method. *Qualitative Research Journal,* 9(2), pp. 27-40.

Bowen, G. A., 2009. Document analysis as a qualitative research method. *Qualitative Research Journal,* 9(2), pp. 27-40.

Bowlby, J., 1973. *Attachment and loss: Volume II: Separation, anxiety and anger.* London: The Hogarth Press and the Institute of Psycho-Analysis.

Braun, V., 2013. *Successful Qualitative Research: A practical guide for beginners.* Los Angeles: Sage Publications.

Braun, V. & Clarke, V., 2019. Reflecting on reflexive thematic analysis.. *Qualitative Research in Sport, Exercise and Health.,* 11(4), pp. 589-597.

Braun, V., Clarke, V. & Terry, G., 2014. Thematic analysis. In: *Qualitative Research in Clinical and Health Psychology.* s.l.:Palgrave Macmillian, pp. 95-114.

Brewer, M. B., 2008. Depersonalized Trust and Ingroup Cooperation. In: J. I. Krueger, ed. *Rationality and Social Responsibility.* New York: Psychology Press, pp. 215-232.

Brinkmann, S., 2014. Interview.. In: *Encyclopedia of critical psychology.* New York: Springer, pp. 1008-1010.

Brinkmann, S. & Kvale, S., 2015. *Interviews: Learning the craft of qualitative research interviewing.* 3rd ed. Thousand Oaks, CA: Sage.

Broadbent, E. et al., 2011. Mental schemas of robots as more human-like are associated with higher blood pressure and negative emotions in a human-robot interaction. *International Journal of Social Robotics,* 3(3), pp. 291-297.

Bruemmer, D. J., Gertman, D. I. & Nielsen, C. W., 2007. Metaphors to Drive By: Exploring New Ways to Guide Human- Robot Interaction.. *Open Cybernetics & Systemics Journal,* Volume 1, pp. 5-12.

Bullock, A. & Trombley, S. eds., 1999. The Other. In: *The New Fontana Dictionary of Modern Thought.* 3rd ed. London: Harper Collins, p. 620.

Bunt, A., Lount, M. & Lauzon, C., 2012, February. *Are explanations always important?: A study of deployed, low-cost intelligent interactive systems..* s.l., s.n., pp. 169-178.

Carayon, P., 2006. Human Factors of Complex Sociotechnical Systems. *Applied Ergonomics,* 37(4), pp. 525-535.

Casebeer, W. D., 2006. *Military Force and Culture Change: Systems, Narratives, and the Social Transmission of Behavior in Counter-Terrorism Strategy,* Monterey CA.: Naval Postgraduate School.

Caspi-Berkowitz, N. et al., 2019. To die for a cause but not for a companion: Attachment-related variations in the terror management function of self-sacrifice. *Journal of Personality and Social Psychology,* 117(6), p. 1105–1126.

Cassidy, A. M., 2009. *Mental models, trust, and reliance: Exploring the effect of human perceptions on automation use.,* Monterey, CA: Naval Postgraduate School.

Castelvecchi, D., 2016. Can we open the black box of AI?. *Nature News,* Volume 538, pp. 20-23.

Castillo-Montoya, M., 2016. Preparing for interview research: The interview protocol refinement framework. *The Qualitative Report,* 21(5), pp. 811-831.

Charmaz, K., 2006. *Constructing grounded theory: A practical guide through qualitative analysis..* Thousand Oaks, CA: Sage.

Chauvin, C., 2011. Human factors and maritime safety.. *Journal of Navigation,* 64(4), pp. 625-632..

Chengqi, X., Cen, Q. & Yan, Z., 2009, November. *Design and research of human-computer interaction interface in autopilot system of aircrafts..* IEEE 10th International Conference, Design

and research of human-computer interaction interface in autopilot system of aircrafts. In Computer-Aided Industrial Design & Conceptual Design, 2009. CAID & CD 2009. IEEE 10th International Conference on (pp. 1498-1501). IEEE., pp. 1498-1501.

Chen, J. Y. & Barnes, M. J., 2013. *Human-Agent Teaming for Multi-Robot Control: A Literature Review (No. ARL-TR-6328).* , Maryland: Army Research Lab Aberdeen Proving Ground.

Chen, J. Y. & Barnes, M. J., 2014. Human–agent teaming for multirobot control: a review of human factors issues.. *IEEE Transactions on Human-Machine Systems,* 44(1), pp. 13-29.

Chien, S., Sycara, K., Liu, J. & Kumru, A., 2016. *Relation between Trust Attitudes Toward Automation, Hofstede's Cultural Dimensions, and Big Five Personality Traits..* s.l., SAGE Publications, pp. 841-845.

Chirkov, V., Ryan, R. M., Kim, Y. & Kaplan, U., 2003. Differentiating autonomy from individualism and independence: A self-determination perspective on internalisation of cultural orientations, gender and well being. *Journal of Personality and Social Psychology,* 84(1), pp. 97-110.

Chiu, H. & Chiang, P., 2019. A trickle-down effect of subordinates' felt trust. *Personnel Review,* 48(4), pp. 957-976.

Christensen, J. & Lyons, J., 2017. Trust between humans and learning machines: developing the gray box. *Mechanical Engineering Magazine Select Articles,* 139(6), pp. 9-13.

Chua, C. E. H., Storey, V. C. & Chiang, R. H., 2012. Deriving knowledge representation guidelines by analyzing knowledge engineer behavior.. *Decision Support Systems,* 54(1), pp. 304-315.

Clancey, W. J., Linde, C., Seah, C. & Shafto, M., 2013. *Work practice simulation of complex human–automation systems in safety critical situations: the Brahms Generalized Überlingen Model (Tech. Publ. 2013-216508).* , Washington.: NASA.

Clayton, T., 2011. *Sea Wolves: The Extraordinary Story of Britain's WW2 Submarines.* Hachette: Little, Brown Book Group.

Clegg, B. A., Heggestad, E. D. & Blalock, L. D., 2010, September. *The influences of automation and trainee aptitude on training effectiveness..* SA, s.n., pp. 2329-2332.

Cofta, P., 2007. *Trust, Complexity and Control: Confidence in a Convergent World..* s.l.:John Wiley and Sons.

Cohen, L., Manion, L. & Morrison, K., 2011. *Research Methods in Education.* 7th ed. s.l.:Routledge.

Cohen, W., 2010. *Heroic leadership: Leading with integrity and honor.* s.l.:John Wiley & Sons.

Coleman, J., 1990. *Foundations of Social Theory. Cambridge, MA: Harvard University.* Cambridge, MA: Harvard University Press.

Collins, E., 2019. Drawing parallels in human–other interactions: a trans-disciplinary approach to developing human–robot interaction methodologies. *Philosophical Transactions of the Royal Society B.*

Colquitt, J. A., Scott, B. A. & LePine, J. A., 2007. Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance.. *Journal of applied psychology,* 92(4), p. 909.

Combéfis, S., Giannakopoulou, D., Pecheur, C. & Feary, M., 2011, October. *A formal framework for design and analysis of human-machine interaction.*. IEEE International Conference, s.n., pp. 1801-1808.

Cooper, L. et al., 2018. Transition from the military into civilian life: An exploration of cultural competence. *Armed Forces & Society,* 44(1), pp. 156-177.

Creswell, J., 1998. *Qualitative inquiry and research design: Choosing among five traditions.* Thousand Oaks: Sage.

Creswell, J., 2015. *30 essential skills for the qualitative researcher.* Los Angeles, CA: Sage Publications.

Creswell, J. W., 2014. *Research design: Qualitative, quantitative, and mixed methods approaches.*. s.l.:Sage publications.

Creswell, J. W. & Clark, V. L. P., 2004. *Principles of qualitative research: Designing a qualitative study.*, Lincoln.: University of Nebraska.

Cronon, W., 1992. A place for stories: Nature, history, and narrative. *Journal of American history,* 78(4), pp. 1347-1376.

Crosston, M., 2017. The Millenials' war: dilemmas of network dependency in today's military. *Defense & Security Analysis,* 33(3), pp. 94-105.

Crotty, M., 1998. *The foundations of social research: Meaning and perspective in the research process.* s.l.:Sage.

Cynkar, A., 2007. The changing gender composition of psychology.. *Monitor on Psychology,* 38(6), p. 46.

Dahlberg, K., 2006. The essence of essences–the search for meaning structures in phenomenological analysis of lifeworld phenomena.. *International journal of qualitative studies on health and well-being,* 1(1), pp. 11-19.

Davis, F., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly,* 13(3), pp. 318-340.

Dawson, S., Crawford, C., Dillon, E. & Anderson, M., 2015, May. *Affecting operator trust in intelligent multirobot surveillance systems.*. Washington, IEEE, pp. 3298-3304.

De Visser, E. et al., 2016. Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied,* 22(3), pp. 331-349.

de Visser, E., Pak, R. & Shaw, T., 2018. From ‘automation’to ‘autonomy’: the importance of trust repair in human–machine interaction. *Ergonomics,* 61(10), pp. 1409-1427.

Deci, E. L. & Vansteenkiste, M., 2004. Self-determination theory and basic need satisfaction: Understanding human development in positive psychology. *Ricerche di psicologia.,* Volume 27, pp. 17-34.

Dehais, F., Causse, M., Vachon, F. & Tremblay, S., 2012. Cognitive conflict in human–automation interactions: a psychophysiological study. *Applied ergonomics,* 43(3), pp. 588-595.

DeNeve, K. M. & Cooper, H., 1998. The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychological bulletin,* 124(2), p. 197.

Denzin, Y. S. & Lincoln, N. K., 2005. *The Sage handbook of qualitative research.* 3rd ed. Thousand Oaks: Sage Publications.

Derrico, P. et al., 2011. Clinical Engineering. In: *Applied Biomedical Engineering.* London: IntechOpen, pp. 169-196.

Deshmukh, A. V., McComb, S. A. & Wernz, C., 2008. Agents as collaborating team members.. In: *Macrocognition in Teams: Theories and Methodologies.* Burlington, VT.: Ashgate Publishing Company, pp. 107-122.

Deutsch, M., 1958. Trust and suspicion. *Journal of conflict resolution,* 2(4), pp. 265-279.

Dey, A. K., 2008, July. *Modeling and Adding Intelligibility to Human Activity.* s.l., IEEE, pp. 5-8.

Dey, A. K., 2009 . *Explanations in Context-Aware Systems.* s.l., s.n., pp. 84-93.

Dillon, A. & Watson, C., 1996. User analysis HCI: The historical lessons from individual differences research. *International Journal of Human-Computer Studies,* 45(6), pp. 619-638.

Doulani, A., 2018. An Assessment of Effective Factors in Technology Acceptance Model: A Meta-Analysis Study.. *Journal of Scientometric Research,* 7(3), pp. 153-166.

Drnec, K., Marathe, A., Lukos, J. & Metcalfe, J., 2016. Trust in automation to decision neuroscience: applying cognitive neuroscience methods to understand and improve interaction decisions involved in human automation interaction. *Frontiers in human neuroscience,* 10(290), pp. 1-14.

Dzindolet, M. et al., 2003. The role of trust in automation reliance. *International journal of human-computer studies,* 58(6), pp. 697-718.

Dzindolet, M. T. et al., 2001. Predicting misuse and disuse of combat identification systems. *Military Psychology,* 13(3), pp. 147-164.

Eagle, M., 1999. Why don't people change? A psychoanalytic perspective. *Journal of Psychotherapy Integration,* 9(1), pp. 3-32.

Eatough, V. & Smith, J. A., 2017. Interpretative phenomenological analysis. In: C. Willig & W. Stainton-Rogers, eds. *Handbook of Qualitative Psychology.* London, UK: Sage, pp. 193-211.

Ein-Dor, T., Mikulincer, M. & Shaver, P. R., 2011. Attachment insecurities and the processing of threat-related information: Studying the schemas involved in insecure people's coping strategies. *Journal of Personality and Social Psychology,* Volume 101, pp. 78-93.

Ein-Dor, T. & Tal, O., 2012. Scared saviors: Evidence that people high in attachment anxiety are more effective in alerting others to threat. *European Journal of Social Psychology,* 42(6), pp. 667-671.

Elliott, V., 2018. Thinking about the coding process in qualitative data analysis. *The Qualitative Report,* 23(11), pp. 2850-2861.

Endsley, M. R., 2015. Final reflections situation awareness models and measures.. *Journal of Cognitive Engineering and Decision Making,* 9(1), pp. 101-111.

Endsley, M. R., Hoffman, R., Kaber, D. & & Roth, E., 2007. Cognitive engineering and decision making: An overview and future course.. *Journal of Cognitive Engineering and Decision Making,* 1(1), pp. 1-21.

Eriksen, J., 2010. Should soldiers think before they shoot?. *Journal of Military Ethics,* 9(3), pp. 195-218.

Erikson, E., 1950. The Eight Ages of Man. In: *Childhood and society.* New York: Norton.

Eriksson, P. & Kovalainen, A., 2008. *Qualitative methods in business research.* London: Sage Publications.

Etherington, K., 2004. *Becoming a reflexive researcher: Using our selves in research.* s.l.: Jessica Kingsley Publishers.

Evans, A. S., 1986. *Beneath the Waves - A History of HM Submarine Losses.* London: William Kimber.

Ezer, N., Fisk, A. D. & Rogers, W. A., 2007. *Reliance on automation as a function of expectation of reliability, cost of verification, and age..* s.l., s.n., pp. 6-10.

Fallon, C. K., Murphy, A. K., Zimmerman, L. & Mueller, S. T., 2010, May. *The calibration of trust in an automated system: A sensemaking process..* s.l., IEEE, pp. 390-395.

Fan, X., Chen, P. C. & Yen, J., 2010. Learning HMM-based cognitive load models for supporting human-agent teamwork.. *Cognitive Systems Research,* 11(1), pp. 108-119.

Fan, X. et al., 2010. Human–agent collaboration for time-stressed multicontext decision making.. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,* 40(2), pp. 306-320.

Fan, X., McNeese, M. & Yen, J., 2010. NDM-Based cognitive agents for supporting decision-making teams. *Human–Computer Interaction,* 25(3), pp. 195-234.

Feigh, K. M., Dorneich, M. C. & Hayes, C. C., 2012. Toward a characterization of adaptive systems a framework for researchers and system designers.. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 54(6), pp. 1008-1024.

Fenell, D., 2008. A distinct culture: Applying multicultural counseling competencies to work with military personnel. *Counseling Today,* Volume 50, pp. 8-9.

Field, M., 2016. *Systematic Search and Review of Mental Models as a Method of Trust Facilitation for Human System Interaction,* s.l.: Unpublished .

Field, M., 2018. *Review of Mental Models as a Method of Trust Facilitation for Human System Interaction.* Swindon, s.n.

Field, M., 2019. *Building risk matrices from interview transcripts utilising HCA and IPA.* Stratford-Upon-Avon, Defence and Security Doctoral Symposium 2018.

Field, M., 2019. *Trust in Automation: How this is shaped by the human operator and the underwater domain.* Stockholm: Sweden, s.n.

Finlayson, M. A. & Corman, S. R., 2013. The military interest in narrative.. *Sprache und Datenverarbeitung,* 37(1-2), pp. 173-191.

Fiske, S. & Taylor, S., 2013. *Social cognition: From brains to culture.* 2nd ed. s.l.:Sage.

Fiske, S. T., Cuddy, A. J. & Glick, P., 2007. Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences,* 11(2), pp. 77-83.

Fisk, S. & Neuberg, S., 1990. A continuum of impression formation, from category based to individuating processes: Influences of information and motivation on attention and interpretation. , 23, pp.1-74.. *Advances in Experimental Social Psychology,* Volume 23, pp. 1-74.

Flemisch, F. et al., 2008. Automation spectrum, inner/outer compatibility and other potentially useful human factors concepts for assistance and automation. *Human Factors for assistance and automation,* pp. 1-16.

Forster, E. M., 1909. *The Machine Stops.* 1st ed. s.l.:Oxford and Cambridge Review.

Fraley, R. & Shaver, P., 2000. Adult romantic attachment: Theoretical developments, emerging controversies, and unanswered questions. *Review of general psychology,* 4(2), pp. 132-154.

French, S. E. & McCain, J., 2004. *The Code of the Warrior: Exploring Warrior Values Past and Present.* s.l.:Rowman & Littlefield Publishers.

French, S. & McCain, J., 2004. *The code of the warrior: Exploring warrior values past and present.* s.l.:Rowman & Littlefield Publishers.

Freud, S., 1919. Papers on Applied Psychoanalysis. The Uncanny.. *Collected papers,* pp. 368-407.

Frost, N., 2011. *Qualitative research methods in psychology: Combining core approaches.* Berkshire: McGraw-Hill Education.

Furukawa, H. P. R., 2003, October. *Supporting system-centered view of operators through ecological interface design: Two experiments on human-centered automation..* s.l., SAGE Publications, pp. 567-571.

Garelick, A. et al., 2017. Beyond the binary: Exploring the role of ambiguity in biphobia and transphobia. *Journal of Bisexuality,* 17(2), pp. 172-189.

Gbrich, C., 2007. *Qualitative Data Analysis: An Introduction.* London: Sage Publications.

Gefen, D., Karahanna, E. & Straub, D., 2003. Inexperience and experience with online stores: The importance of TAM and trust. *IEEE Transactions on engineering management,* 50(3), pp. 307-321.

Gentles, S. J., Charles, C., Ploeg, J. & McKibbon, K., 2015. Sampling in qualitative research: Insights from an overview of the methods literature.. *The Qualitative Report,* 20(11), pp. 1772-1789.

Gerck, E., 2002. Trust as qualified reliance on information. *The COOK Report on Internet,* Volume 10, pp. 19-24.

Glaser, B. G. & Strauss, A. L., 2017. *Discovery of grounded theory: Strategies for qualitative research.* s.l.:Routledge.

Glaser, B. S. & Strauss, A., 1967. *The discovery of grounded theory.* London: Sage.

Goodrich, M. A. & Schultz, A. C., 2007. Human-robot Interaction: A Survey. *Foundations and trends in human-computer interaction,* 1(3), pp. 203-275.

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

Green, J. & Thorogood, N., 2013. *Qualitative methods for health research.* s.l.:Sage.

Grimell, J., 2017. A Service Member's Self in Transition: A Longitudinal Case Study Analysis. *Journal of Constructivist Psychology,* 30(3), pp. 255-269.

Grint, K., 2010. The sacred in leadership: Separation, sacrifice and silence.. *Organization Studies,* 31(1), pp. 89-107.

Gronn, P., 2003. *The new work of educational leaders.* London: Sage.

Groom, V. & Nass, C., 2007. Can robots be teammates?: Benchmarks in human–robot teams.. *Interaction Studies,* 8(3), pp. 483-500.

Guba, E. & Lincoln, Y., 1994. Competing paradigms in qualitative research. In: *Handbook of qualitative research.* California: Sage Publications, pp. 105-117.

Gubrium, J. & Holstein, J., 2009. *Analysing Narrative Reality.* Thousand Oaks, CA: Sage.

Guest, G., Bunce, A. & Johnson, L., 2006. How many interviews are enough? An experiment with data saturation and variability. *Field Methods,* 18(1), pp. 59-82.

Gutzwiller, R. S., Clegg, B. A. & Blitch, J. G., 2013. Part-task training in the context of automation: Current and future directions.. *The American journal of psychology,* 126(4), pp. 417-432.

Hall, S. & McQuay, W., 2010. *Review of trust research from an interdisciplinary perspective-psychology, sociology, economics, and cyberspace.* s.l., IEEE 2010 National Aerospace & Electronics Conference, pp. 18-25.

Hall, S. & McQuay, W., 2010. *Review of trust research from an interdisciplinary perspective-psychology, sociology, economics, and cyberspace..* s.l., IEEE, pp. 18-25.

Hancock, P. A., Billings, D. R. & Schaefer, K. E., 2011. Can you trust your robot?. *Ergonomics in Design: The Quarterly of Human Factors Applications,* 19(3), pp. 24-29.

Hancock, P. A. et al., 2011. A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 53(5), pp. 517-527.

Hancock, P. A. et al., 2011. A meta-analysis of factors affecting trust in human-robot interaction.. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 53(5), pp. 517-527.

Hancock, P. A. et al., 2011. A meta-analysis of factors affecting trust in human-robot interaction.. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* Issue 53, pp. 517-527.

Handy, C., 1995. Trust and the virtual organization. *Harvard Business Review,* 73(3), pp. 40-50.

Hardin, R., 2002. *Trust & Trustworthiness.* New York: Russell Sage Foundation.

Harris, A., 2009. *Distributed Leadership.* London: Springer.

Haven, K., 2007. *Story proof: The science behind the startling power of story.* s.l.:Greenwood Publishing Group.

Hawkley, J. K., Mares, A. L. & Giammanco, C. A., (2005). .. *The human side of automation: Lessons for air defense command and control (No. ARL-TR-3468),* Maryland: s.n.

Hawley, J. K., Mares, A. L. & Giammanco, C. A., 2006. *Training for effective human supervisory control of air and missile defense systems (No. ARL-TR-3765),* Texas: Army Research Labratory.

Heal, J. et al., 2019. *Applying behavioural insights to successful transition.* [Online]
Available at: https://www.fim-trust.org/wp-content/uploads/2019/04/20190411-Applying-behavioural-insights-to-successful-transition-FINAL.pdf
[Accessed 18 January 2020].

Heidegger, M., 1962. *Being and time..* Oxford: Blackwell.

Hendrick, H. W. & Kleiner, B. M., 2001. *Macroergonomics: An Introduction to Work System Design.* Santa Monica, CA: Human Factors and Ergonomics Society.

Heron, J. & Reason, P., 1997. A participatory inquiry paradigm. *Qualitative inquiry,* 3(3), pp. 274-294.

Hess, T., Fuller, M. & Campbell, D., 2009. Designing Interfaces with Social Presence: Using Vividness and Extraversion to Create Social Recommendation Agents. *Journal of the Association for Information Systems,* 10(12), pp. 889-919.

Higgins, J. P. & Green, S., 2008. *Cochrane handbook for systematic reviews of interventions.* Chichester: Wiley-Blackwell.

Higher Education Statistics Agency, 2002. *HESA subject codes.* [Online] Available at: https://www.hesa.ac.uk/support/documentation/jacs/hesa-codes [Accessed 11 September 2019].

Hockey, J., 2016. *Squaddies: Portrait of a subculture..* 3rd ed. Liverpool: Liverpool University Press.

Hoepfl, M., 1997. Choosing qualitative research: A primer for technology education researchers. *Journal of Technology Education,* 9(1), pp. 47-63.

Hoff, K. A. & Bashir, M., 2015. Trust in Automation Integrating Empirical Evidence on Factors That Influence Trust.. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 57(3), pp. 407-434.

Hoffman, R. R., Johnson, M., Bradshaw, J. M. & Underbrink, A., 2013. Trust in automation. *Intelligent Systems,* 28(1), pp. 84-88.

Hoffman, R. R. & Woods, D. D., 2011. Beyond Simon's slice: Five fundamental trade-offs that bound the performance of macrocognitive work systems.. *Intelligent Systems,* 26(6), pp. 67-71.

Hofstede, G., 1991. *Cultures Consequences: Software of the Mind.* London: McGrawHill.

Hogg, M. & Reid, S., 2006. Social identity, self-categorization, and the communication of group norms. *Communication theory,* 16(1), pp. 7-30.

Ho, G., Wheatley, D. & Scialfa, C., 2005. Age Differences in Trust and Reliance of a Medication Management System. *Interacting With Computers,* 17(6), p. 690–710.

Houppert, K., 2007. *Home fires burning: Married to the military-for better or worse.* s.l.:Ballantine Books.

Howard-Payne, L., 2016. Glaser or Strauss? Considerations for selecting a grounded theory study.. *South African Journal of Psychology,* 46(1), pp. 50-62.

Huang, H. & Bashir, M., 2017. *Users' trust in automation: A cultural perspective..* s.l., Springer, pp. 282-289.

Hughes, J. S., Rice, S., Trafimow, D. & Clayton, K., 2009. The automated cockpit: A comparison of attitudes towards human and automated pilots.. *Transportation research part F: traffic psychology and behaviour,* 12(5), pp. 428-439.

Husserl, E., 1970. *The crisis of European sciences and transcendental phenomenology: An introduction to phenomenological philosophy..* s.l.:Northwestern University Press.

Imranali, S., Jain, S., Lal Patidar, B. & Neeraj, N., 2012. A Framework for Modeling and Designing of Intelligent and Adaptive Interfaces for Human Computer Interaction.. *International Journal of Applied Information Systems,* Volume 1, pp. 20-25.

Jacobs, T., 1991. Introduction to Section 4. In: *Handbook of military psychology.* s.l.:John Wiley & Sons..

Jefferson, G., 2004. Glossary of transcript symbols with an introduction. In: G. H. Lerner, ed. *Conversation Analysis: Studies from the First Generation.* Amsterdam: John Benjamins, pp. 19-31.

Jetten, J. et al., 2015. Having a lot of a good thing: Multiple important group memberships as a source of self-esteem. *PloS one,* 10(5).

Jian, J. Y., Bisantz, A. M. & Drury, C. G., 2000. Foundations for an empirically determined scale of trust in automated systems.. *International Journal of Cognitive Ergonomics ,* 4(1), pp. 53-71.

Joe, J., O'Hara, J., Medema, H. & Oxstrand, J., 2014, June. *Identifying requirements for effective human-automation teamwork.* s.l., s.n.

Johns, J. et al., 1984. *Cohesion in the US Military,* Washington, DC: National Defence University Press.

Jones, S. & Shah, P., 2016. Diagnosing the locus of trust: A temporal perspective for trustor, trustee, and dyadic influences on perceived trustworthiness. *Journal of Applied Psychology,* 101(3), pp. 392-414.

Jung, C., 1976. Phenomenology of the Self. In: J. Campbell, ed. *The Portable Jung.* New York: Penguin Books, p. 147.

Jupp, V., 2006. *The Sage Dictionary of Social Research Methods.* Thousand Oaks, CA: SAGE Publications.

Kaplan, F., 2004. Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots. *International journal of humanoid robotics,* 1(3), pp. 465-480.

Kaplan, F., 2004. Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots.. *International journal of humanoid robotics,* 1(3), pp. 465-480.

Kappas, A., 1997. The fascination with faces: Are they windows to our soul?. *Journal of Nonverbal Behaviour,* Volume 21, pp. 157-161.

Karakatsanis, N. & Swarts, J., 2015. *Political and Military Sociology: Military Perceptions and Perceptions of the Military.* 1st ed. New York: Routledge.

Katz, D. & Kahn, R., 1966. *The Social Psychology of Organizations.* New York: Wiley.

Keel, P., Cooke, N. & Sither, M., (2014).. Improving cognitive interaction through computational collaboration agents. Improving cognitive interaction through computational collaboration agents. *Theoretical Issues in Ergonomics Science,* 15(5), pp. 415-431.

Kennedy, W. G., Sibley, C. & Coyne, J., 2015. *Self-Confidence of Autonomous Systems in a Military Environment.* Arlington, VA., 2015 AAAI Fall Symposium Series..

Kiesler, S., (2005, August). . In 2005. ROMAN 2005. (pp. 729-734). IEEE.. *Fostering common ground in human-robot interaction.* s.l., IEEE.

King, W. R. & He, J., 2006. A meta-analysis of the technology acceptance model.. *Information & management,* 43(6), pp. 740-755.

Kirke, C., 2010. Military cohesion, culture and social psychology.. *Defence & Security Analysis,* 26(2), pp. 143-159.

Kirke, C., 2010. Military cohesion, culture and social psychology.. *Defence & Security Analysis,* 26(2), pp. 143-159.

Kjeldsen, J., 2004. *Retorikk i vår tid: en innføring i moderne retorisk teori.* Oslo: Spartacus.

Klein, G., 2008. Naturalistic decision making. *Human Factors,* 50(3), pp. 456-460.

Knight, A. & Eisenkraft, N., 2015. Positive is usually good, negative is not always bad: The effects of group affect on social integration and task performance. *Journal of Applied Psychology,* 100(4), pp. 1214-1227.

Koltko-Rivera, M. E., 2004. The psychology of worldviews. *Review of General Psychology,* Volume 8, pp. 3-58.

Krämer, N., von der Pütten, A. & Eimler, S., 2012. Human-agent and human-robot interaction theory: Similarities to and differences from human-human interaction. In: Z. M. & d. O. J.V., eds. *In Human-computer interaction: The agency perspective.* Heidelberg, Berlin: Springer, pp. 215-240.

Kramer, R., 1999. Trust and distrust in organizations: Emerging perspectives, enduring questions.. *Annual review of psychology,* 50(1), pp. 569-598.

Kramer, R. & Tyler, T., 1996. *Trust in Organizations: Frontiers of Theory and Research.* California: Sage Publications.

Krueger, G. P. & Banderet, L. E., 2007. Implications for studying team cognition and team performance in network-centric warfare paradigms.. *Aviation, space, and environmental medicine,* 78(1), pp. B58-B62.

Kulesza, T., 2012, February. *An explanation-centric approach for personalizing intelligent agents.* Lisbon, Portugal, ACM, pp. 375-378.

Kulms, P. & Kopp, S., 2018. A social cognition perspective on human–computer trust: The effect of perceived warmth and competence on trust in decision-making with computers. *Frontiers in Digital Humanities,* 5(14), pp. 1-11.

Lael, R. L., 1982. *The Yamashita Precedent: War Crimes and Command Responsibility.* Wilmington: Scholarly Resources.

Langan-Fox, J., Canty, J. M. & Sankey, M. J., 2009. Human–automation teams and adaptable control for future air traffic management.. *International Journal of Industrial Ergonomics,* 39(5), pp. 894-903.

Larkin, M., Eatough, V. & Osborn, M., 2011. Interpretative phenomenological analysis and embodied, active, situated cognition.. *Theory & Psychology,* 21(3), pp. 318-337.

Latour, B., 1991. Technology is Society Made Durable. In: J. Law, ed. *A Sociology of Monsters Essays on Power, Technology and Domination.* s.l.:Sociological Review Monograph, pp. 103-132.

Latour, B., 2005. *Reassembling the social: An introduction to actor-network-theory..* Oxford: Oxford University Press.

Latour, B., 2013. *Bruno Latour,* Bergen: University of Bergen.

Layder, D., 2012. *Doing excellent small-scale research.* s.l.:Sage.

Lebel, U. & Ben-Shalom, U., 2018. Military Leadership in Heroic and Post-heroic Conditions.. In: G. Caforio & M. Nuciari, eds. *Handbook of the Sociology of the Military.* s.l.:Springer, pp. 463-475.

Lee, J. D. & See, K. A., 2004. Trust in automation: Designing for appropriate reliance.. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 46(1), pp. 50-80.

Lee, J. D. & See, K. A., 2004. Trust in Automation: Designing for Appropriate Reliance.. *Human Factors: The Journal of the Human Factors and Ergonomics Society.,* 46(1), pp. 50-80.

Lee, S. L., Lau, I. Y. M., Kiesler, S. & Chiu, C. Y., 2005, April. *Human mental models of humanoid robots. In . Proceedings of the (pp. 2767-2772). IEEE.*. s.l., s.n., pp. 2767-2772.

Lewicki, R. J. & Bunker, B. B., 1996. Developing and maintaining trust in work relationships. In: R. M. Kramer & T. R. Tyler, eds. *Trust in organizations: Frontiers of theory and research.* s.l.:Sage, pp. 114-139.

Lewicki, R. J., Tomlinson, E. C. & Gillespie, N., 2006. Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. *Journal of management,,* 32(6), pp. 991-1022..

Lewis, J. D. & Weigert, A., 1985. Trust as a social reality. *Social forces,* 63(4), pp. 967-985.

Lim, B. Y., Dey, A. K. & Avrahami, D., 2009, April. *Why and why not explanations improve the intelligibility of context-aware intelligent systems.*. Boston, Massachusetts, USA., ACM., pp. 2119-2128.

Lincoln, Y. S. & Guba, E. G., 1985. *Naturalistic Inquiry.* Beverly Hills, CA.: Sage.

Lo, J. C., Pluyter, K. R. & Meijer, S. A., 2015. Individual Markers of Resilience in Train Traffic Control The Role of Operators' Goals and Strategic Mental Models and Implications for Variation, Expertise, and Performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* p. 1–12.

Looser, C. & Wheatley, T., 2010. The tipping point of animacy: How, when, and where we perceive life in a face.. *Psychological Science,* Volume 21, p. 1854–1862.

Loy, D., 2002. *A Buddhist History of the West: Studies in Lack.* Albany: State University of New York Press.

Luhmann, N., 1979. *Trust and power.* Chichester: Wiley.

Lyon, F. M. G. &. S. M. N., 2015. Introduction. Researching trust: the ongoing challenge of matching objectives and method. In: *Handbook of research methods on trust.* s.l.:Edward Elgar Publishing, pp. 1-24.

Lyon, F., Möllering, G. & Saunders, M. N., 2015. Introduction. Researching trust: the ongoing challenge of matching objectives and method. In: *Handbook of research methods on trust.* s.l.:Edward Elgar Publishing, pp. 1-24.

Lyons, J. B. & Stokes, C. K., 2012. Human–human reliance in the context of automation. *Human factors,* 54(1), pp. 112-121.

Lyons, J. B., Stokes, C. K. & Schneider, T. R., 2018. Predictors and outcomes of trust in teams. In: *Trust in military teams.* s.l.:CRC Press, pp. 55-72.

Lyons, J. et al., 2018. Exploring trust barriers to future autonomy: A qualitative look. In: *Advances in Human Factors in Simulation and Modeling.* s.l.:Springer, pp. 3-11.

MacDorman, K. F., Vasudevan, S. K. & Ho, C. C., 2009. Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures.. *AI & society,* 23(4), pp. 485-510.

Madni, A. M., 2010. Integrating humans with software and systems: Technical challenges and a research agenda.. *Systems Engineering,* 13(3), pp. 232-245.

Madni, A. M., 2011. Integrating humans with and within complex systems.. *CrossTalk*, May/June, Volume 5.

Malterud, K., 2001. Qualitative research: Standards, challenges and guidelines. *The Lancet,* Volume 358, pp. 483-488.

Marek, R. et al., 2018. Hippocampus-driven feed-forward inhibition of the prefrontal cortex mediates relapse of extinguished fear. *Nature Neuroscience,* 21(3), pp. 384-392.

Marques, J. M. & Paez, D., 1994. The 'black sheep effect': Social categorization, rejection of ingroup deviates, and perception of group variability. *European review of social psychology,* 5(1), pp. 37-68.

Martin, J., 1992. *Cultures Organizations: Three Perspectives.* London: Oxford University Press.

Mason, M., 2010. Sample size and saturation in PhD studies using qualitative interviews.. 11(3).

Matthews, G., Wohleber, R., L. J. & Panganiban, A., 2019. Fatigue, Automation, and Autonomy: Challenges for Operator Attention, Effort, and Trust. In: *Human Performance in Automated and Autonomous Systems: Current Theory and Methods.* s.l.:CRC Press, p. 127.

Mayer, A., 2008. *The manipulation of user expectancies: Effects on reliance, compliance, and trust using an automated system.,* Atlanta: Georgia Institute of Technology,.

Mayer, R., Davis, J. & Schoorman, F., 1995. An integrative model of organizational trust. *Academy of management review,* 20(3), pp. 709-734.

Mayer, R., Davis, J. & Schoorman, F., 2007. An Integrative Model of Organizational Trust: Past, Present, and Future. *Academy of management review,* 32(2), p. 344–354.

McAllister, D., 1995. Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of management journal,* 38(1), pp. 24-59.

McCrae, R. R. & John, O. P., 1992. An introduction to the five-factor model and its applications.. *Journal of personality,* 60(2), pp. 175-215.

McDorman, B., Clabaugh, C. & Mataric, M., 2016. *Attachment Theory in Long-Term Human-Robot Interaction..* [Online]

Available at:
https://robotics.usc.edu/publications/media/uploads/pubs/pubdb_933_5438c02a77f74409bafeacbfe291363f.pdf
[Accessed 17 03 2020].

McEvily, B. & Tortoriello, M., 2011. Measuring trust in organisational research: Review and recommendations. *Journal of Trust Research,* 1(1), pp. 23-63.

McLeod, C., 2000. Our attitude towards the motivation of those we trust.. *The Southern journal of philosophy,* 38(3), pp. 465-479.

Merritt, S. M., Heimbaugh, H., LaChapell, J. & Lee, D., 2013. I trust it, but I don't know why: effects of implicit attitudes toward automation on trust in an automated system.. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 55(3), pp. 520-534.

Messick, D. M. & Kramer, R. M., 2001. Trust as a Form of Shallow Morality. In: R. Kramer & K. S. Cook, eds. *Trust in Society.* New York: Sage, pp. 89-117.

Mikulincer, M. & Florian, V., 1998. The relationship between adult attachment styles and emotional and cognitive reactions to stressful events. In: J. A. Simpson & W. S. Rholes, eds. *Attachment theory and close relationships.* s.l.:Guildford Press, pp. 143-165.

Mikulincer, M. & Shaver, P. R., 2007. *Attachment in adulthood: Structure, dynamics, and change.* s.l.:Guilford Press.

Miles, M. & Huberman, A., 1994. *Qualitative data analysis: A sourcebook..* Beverly Hills: Sage Publications.

Milgram, S., 1963. Behavioral study of obedience. *Journal of Abnormal and Social Psychology,* Volume 67, pp. 371-378.

Miller, A. S. & Mitamura, T., 2003. Are surveys on trust trustworthy?. *Social Psychology Quarterly,* 66(6), pp. 62-70.

Miller, C. A. & Parasuraman, R., 2007. Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control.. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 49(1), pp. 57-75.

Mindock, J. & Klaus, D., 2012. *Development and application of spaceflight performance shaping factors for human reliability analysis.* s.l., Mindock, J., & Klaus, D. (2012). Development and application of spaceflight performance shaping factors for human reliability analysis. In 41st International Conference on Environmental Systems (p. 5158)., p. 5158.

Ministry of Defence, 2018. *Human-Machine Teaming (JCN 1/18).* [Online]
Available at: www.gov.uk/government/publications/human-machine-teaming-jcn-118
[Accessed 20 February 2019].

Ministry of Defence, 2019. *Deaths in the UK Regular Armed Forces: Annual Summary and Trends over Time, 1 January 2009 to 31 December 2018,* s.l.: National Statistics.

Ministry of Defence, 2019. *Ministry of Defence Annual Report and Accounts 2018 to 19.* s.l.:s.n.

Misztal, B., 2013. *Trust in modern societies: The search for the bases of social order*. s.l.:John Wiley & Sons.

MoD U.K, 2014. *Global strategic trends out 2045.*. [Online]
Available at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/348164/20140821_DCDC_GST_5_Web_Secured.pdf

Mogalakwe, M., 2006. The use of documentary research methods in social research. *African sociological review,* 10(1), pp. 221-230.

Moher, D., Liberati, A., Tetzlaff, J. & Altman, D., 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *BMJ,* 339(2535).

Moher, D. & Liberati, A. T. J. A. D., 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement.. *BMJ,* 339(2535).

Möllering, G., 2001. The nature of trust: From Georg Simmel to a theory of expectation, interpretation and suspension. *Sociology,* 35(2), pp. 403-420.

Möllering, G., 2017. Cultivating the field of trust research. *Journal of Trust Research,* 7(2), pp. 107-114.

Molm, L. S. D. a. C. J., 2007. The value of reciprocity. *Social Psychology Quarterly,* 70(2), pp. 199-217.

Moore, G. E., 1965. Cramming more components onto integrated circuits. *Electronic,* 38(8), p. 114.

Moore, G. E., 1965. Cramming more components onto integrated circuits. *Electronics,* 38(8), p. 114.

Moravec, H., 1988. *Mind Children.* s.l.:Harvard University Press.

Morgner, C., 2018. Trust and Society: Suggestions for Further Development of Niklas Luhmann's Theory of Trust.. *Canadian Review of Sociology,* 55(2), pp. 232-256.

Mori, M., MacDorman, K. & Kageki, N., 2012. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine,* 19(2), pp. 98-100.

Morita, P. P. & Burns, C. M., 2014. Understanding ‘interpersonal trust’ from a human factors perspective: insights from situation awareness and the lens model. *Theoretical Issues in Ergonomics Science,* 15(1), pp. 88-110.

Morrison, E. & Milliken, F., 2000. Organizational silence: A barrier to change and development in a pluralistic world. *Academy of Management review,* 25(4), pp. 706-725.

Morrison, Z., Gregory, D. & Thibodeau, S., 2012. “Thanks for using me”: an exploration of exit strategy in qualitative research. *International Journal of Qualitative Methods,* 11(4), pp. 416-427.

Morse, J., 1994. Designing funded qualitative research.. In: *Handbook of qualitative research.* 2nd ed. Thousand Oaks, CA: Sage, pp. 220-235.

Morse, J. M., 1994. *Critical issues in qualitative research methods.* s.l.:Sage.

Mosier, K. L. et al., 2013. utomation, Task, and Context Features Impacts on Pilots' Judgments of Human–Automation Interaction.. *Journal of Cognitive Engineering and Decision Making,* 7(4), pp. 377-399.

Myers, M., 1997. Qualitative research in information systems. *MIS Quarterly,* Volume 21, pp. 241-242.

Nachtwei, J., 2011. The many faces of human operators in process control: a framework of analogies.. *Theoretical Issues in Ergonomics Science,* 12(4), pp. 297-317.

Nass, C. & Moon, Y., 2000. Machines and mindlessness: Social responses to computers. *Journal of social issues,* 56(1), pp. 81-103.

NATO, 2010. *STANAG 2116: NATO codes for grades of military personnel (6th Edition),* Brussels: NATO.

Neerincx, M. A. et al., 2008, January. *Situated cognitive engineering: Developing adaptive track handling support for naval command and control centers..* s.l., s.n., pp. 3-20.

Neven, H. et al., 2019. Quantum supremacy using a programmable superconducting processor. *Nature,* 574(7779), pp. 505-510.

Newell, A. & Rosenbloom, P. S., 1981. Mechanisms of skill acquisition and the law of practice. In: *Cognitive skills and their acquisition.* s.l.:s.n., pp. 1-55.

Nilsson, M., Riveiro, M. & Ziemke, T., 2008. *Investigating human-computer interaction issues in information-fusion-based decision support.,* Sweden: s.n.

Nimmo, R., 2011. Actor-network theory and methodology: social research in a more-than-human world. *Methodological Innovations Online,* 6(3), pp. 108-119.

Nisbett, R. E. & Wilson, T. D., 1977. Telling more than we can know: Verbal reports on mental processes.. *Psychological Review,* 84(3), p. 231–259.

Nothdurft, F., Lang, H., Klepsch, M. & Minker, W., 2013, April. *Bases of human-computer trust and explanations..* Paris, France., ACM, pp. 1737-1742.

Nothdurft, F. & Minker, W., 2014, January. *Justification and transparency explanations in dialogue systems to maintain human-computer trust.* Napa, Springer.

Oakes, M., 2011. Formation of identity in wartime El Salvador: Protective and harmful factors. *Journal of Human Behavior in the Social Environment,* 21(8), pp. 926-947.

Oleson, K. E. et al., 2011, February. *Antecedents of trust in human-robot collaborations.* s.l., IEEE, pp. 175-178.

Olson, K. E., Fisk, A. D. & Rogers, W. A., 2009, October . *Collaborative Automated Systems: Older Adults' Mental Model Acquisition and Trust in Automation..* s.l., SAGE Publications, pp. 1704-1708.

Onwuegbuzie, A. J., Johnson, R. B. & Collins, K. M., 2009. Call for mixed analysis: A philosophical framework for combining qualitative and quantitative approaches. *International journal of multiple research approaches,* 3(2), pp. 114-139.

Osanloo, A. & Grant, C., 2016. Understanding, selecting, and integrating a theoretical framework in dissertation research: Creating the blueprint for your "house".. *Administrative issues journal: connecting education, practice, and research.,* 4(2), p. 7.

Ososky, S. J., 2013. *Influence of Task-Role Mental Models on Human Interpretation of Robot Motion Behavior,* Florida: University of Central Florida Orlando.

Ostrom, E., 2005. *Understanding Institutional Diversity.* s.l.:Princeton University Press.

Paharia, N., Keinan, A., Avery, J. & Schor, J., 2011. The underdog effect: The marketing of disadvantage and determination through brand biography. *Journal of Consumer Research,* 37(5), pp. 775-790.

Pak, R., Rovira, E., McLaughlin, A. C. & Baldwin, N., 2017. Does the domain of technology impact user trust? Investigating trust in automation across different consumer-oriented domains in young adults, military, and older adults. *Theoretical Issues in Ergonomics Science,* 18(3), pp. 199-220.

Pak, R., Rovira, E., McLaughlin, A. C. & Leidheiser, W., 2017. Evaluating attitudes and experience with emerging technology in cadets and civilian undergraduates.. *Military Psychology,* 29(5), pp. 448-455.

Pandit, N. R., 1996. The creation of theory: A recent application of the grounded theory method.. *The qualitative report,* 2(4), pp. 1-15.

Parasuraman, R., Sheridan, T. B. & Wickens, C. D., 2000. A Model for Types and Levels of Human Interaction with Automation. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,* 30(3), pp. 286-297.

Pardo del Val, M. & Martínez Fuentes, C., 2003. Resistance to change: a literature review and empirical study. *Management Decision,* 41(2), pp. 148-155.

Parks, W. H., 1973. Command Responsibility for War Crimes. *Military Law Review,* 52(1).

Parush, A., Ahuvia, S. & Erev, I., 2007. Degradation in spatial knowledge acquisition when using automatic navigation systems. In: *Spatial information theory.* Berlin, Heidelberg: Springer, pp. 238-254.

Patton, M. Q., 2005. *Qualitative research.* s.l.:John Wiley & Sons, Ltd.

Payne, G. & Payne, J., 2004. *Key Concepts in Social Research.* London: Sage Publications.

Peoples, J. & Bailey, G., 2011. *Humanity: An introduction to cultural anthropology.* 9th ed. s.l.:Cengage Learning.

Petriglieri, J., 2011. Under threat: Responses to and the consequences of threats to individuals' identities. *Academy of Management Review,* 36(4), pp. 641-662.

Phillips, E., Ososky, S., Grove, J. & Jentsch, F., 2011, September. *From Tools to Teammates Toward the Development of Appropriate Mental Models for Intelligent Robots.* s.l., SAGE Publications, pp. 1491-1495.

Phillips, E., Ososky, S. & Jentsch, F., 2014, September. *Leveraging Features of Human—Technology Teams to Support Mental Models in Future Soldier—Robot Teams..* s.l., SAGE Publications, pp. 1179-1183.

Piccinini, G. F. B., Rodrigues, C. M., Leitão, M. & Simões, A., 2015. Reaction to a critical situation during driving with adaptive cruise control for users and non-users of the system. *Safety Science,* Volume 72, pp. 116-126.

Pidgeon, N., 1998. Safety culture: key theoretical issues.. *Work & Stress,* 12(3), pp. 202-216.

Pinel, E. et al., 2006. Seeing I to I: a pathway to interpersonal connectedness. *Journal of personality and social psychology,* 90(2), pp. 243-257.

Popper, K. R., 1968. *The Logic of Scientific Discovery.* s.l.:Hutchinson.

Powell, J. & Menendian, S., 2016. The Problem of Othering: Towards Inclusiveness and Belonging.. *Othering & Belonging: ,* Issue 1, pp. 14-40.

Prewett, M. S. et al., 2010. Managing workload in human–robot interaction: A review of empirical studies.. *Computers in Human Behavior,* 26(5), pp. 840-856.

PytlikZillig, L. et al., 2016. The dimensionality of trust-relevant constructs in four institutional domains: Results from confirmatory factor analyses. *Journal of Trust Research,* 6(2), pp. 111-150..

QSR International Pty Ltd, 2012. *NVivo qualitative data analysis Software.* Version 10 ed. s.l.:s.n.

Quinn, K., Lewis, D., O'Sullivan, D. & Wade, V. P., 2009. An analysis of accuracy experiments carried out over of a multi-faceted model of trust. *International Journal of Information Security,* 8(2), pp. 103-119.

Ramiller, N. C., 2020. Making the Case: The Systems Project Case Study as Storytelling. *Journal of Information Systems Education,* 14(5).

Rapley, T., 2004. Interviews. In: C. Seale, G. Gobo, J. Gubrium & D. Silverman, eds. *Qualitative Research Practice.* London: Sage, pp. 15-33.

Rasmussen, J., 1997. Risk management in a dynamic society: a modelling problem.. *Safety science,* 27(2-3), pp. 183-213.

Reason, J., 1990. *Human Error.* s.l.:Cambridge University Press.

Redmond, S. et al., 2015. A brief introduction to the military workplace culture. *Work,* 50(1), pp. 9-20.

Reeves, B. & Nass, C., 1996. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places..* s.l.:Cambridge University Press.

Richards, K., 2003. *Qualitative inquiry in TESOL.* Basingstoke: Palgrave Macmillan.

Richards, L., 2015. *Handling qualitative data: A practical guide.* 3rd ed. London, UK: Sage Publication.

Riessman, C., 1993. *Narrative analysis (Qualitative research methods, volume 30).* Newbury Park: Sage Publications.

Ritchie, J., Lewis, J. & Elam, G., 2003. Designing and selecting samples. In: *Qualitative research practice. A guide for social science students and researchers.* Thousand Oaks: Sage, pp. 77-108.

Robbins, B., 2014. *On the Origins of Trust,* s.l.: University of Washington.

Robbins, B., 2016. What is trust? A multidisciplinary review, critique, and synthesis. *Sociology compass,* 10(10), pp. 972-986..

Robson, C., 2002. *Real World Research: A Resource for Social Scientists and Practitioner-Researchers.* 2nd ed. Oxford: Blackwell Publishers Ltd.

Rom, E. & Mikulincer, M., 2003. Attachment theory and group processes: The association between attachment style and group-related representations, goals, memories, and functioning. *Journal of Personality and Social Psychology,* 84(6), pp. 1220-1235.

Rossman, G. & Rallis, S., 2016. *An introduction to qualitative research: Learning in the field.* 4th ed. s.l.:Sage Publications.

Rotter, J., 1966. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs,* 80(1), pp. 1-28.

Rotter, J., 1967. A new scale for the measurement of interpersonal trust. *Journal of personality,,* 35(4), pp. 651-665.

Rotter, J., 1980. Interpersonal trust, trustworthiness, and gullibility. *American Psychologist,* 35(1), pp. 1-7.

Rousseau, D., Sitkin, S., Burt, R. & Camerer, C., 1998. Not so different after all: A cross-discipline view of trust. *Academy of management review,* 23(3), pp. 393-404.

Rousseau, D., Sitkin, S., Burt, R. & Camerer, C., 1998. Not so different after all: A cross-discipline view of trust.. *Academy of management review.,* 23(3), pp. 393-404.

Rowland, R., 2003. *The Yamashita standard: Sugamo and the River Kwai.* s.l., Princeton University, pp. 1-32.

Royakkers, L. & van Est, R., 2015. A literature review on new robotics: automation from love to war. *International journal of social robotics,* 7(5), pp. 549-570.

RTO/NATO, 2007. *Virtual Environments for Intuitive Human-System Interaction – Human Factors Considerations in the Design, Use, and Evaluation of AMVE-Technology,* s.l.: s.n.

Rubin, H. J. & Rubin, I. S., 2011. *Qualitative interviewing: The art of hearing data..* s.l.:Sage Publishing.

Sætrevik, B. &. E. J., 2013. The “Similarity Index” as an Indicator of Shared Mental Models and Situation Awareness in Field Studies.. *Journal of Cognitive Engineering and Decision Making,* 8(2), p. 119–136.

Saldaña, J., 2009. *The coding manual for qualitative researchers.* 1st ed. Thousand Oaks, CA: Sage Publications.

Saldana, J., 2011. *Fundamentals of Qualitative Research.* Oxford: Oxford University Press..

Saldaña, J., 2016. *The coding manual for qualitative researchers.* London, UK: Sage Publications.

Salmon, P., 2018. *"'Human error' is not the cause of anything, it is a consequence of badly designed and poorly functioning systems. When planes, trains, trams etc crash, it is never the fault of one person, there is always a web of interacting contributory factors",* s.l.: s.n.

Salmon, P. M. et al., 2008. What really is going on? Review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science,* 9(4), pp. 297-323.

Sanders, T. et al., 2011, September. *A model of human-robot trust theoretical model development.* s.l., SAGE Publications, pp. 1432-1436.

Sarantakos, S., 2013. *Social Research.* 4th ed. Basingstoke: Palgrave Macmillan.

Savela, N., Turja, T. & Oksanen, A., 2018. Social acceptance of robots in different occupational fields: A systematic literature review. *International Journal of Social Robotics,* 10(4), pp. 493-502.

Scanzoni, J., 1979. Social exchange and behavioral interdependence. In: *Social exchange in developing relationships.* s.l.:Academic Press, pp. 61-98.

Schaefer, K. E. et al., 2014. *A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Human-Robot Interaction* , s.l.: s.n.

Schaefer, K. E., Evans III, A. W. & Hill, S. G., 2015. *Command and Control in Network-Centric Operations: Trust and Robot Autonomy.,* s.l.: U. S. Army Research Laboratory.

Schaffernicht, M. & Groesser, S. N., 2011. A comprehensive method for comparing mental models of dynamic systems. *European Journal of Operational Research,* 210(1), pp. 57-67.

Schein, E., 2010. *Organizational culture and leadership.* 4th ed. San Francisco: John Wiley & Sons.

Schöbel, M., 2009. Trust in high-reliability organizations.. *Social Science Information,* 48(2), pp. 315-333.

Schoenherr, J. & Hamstra, S., 2016. Psychometrics and its discontents: An historical perspective on the discourse of the measurement tradition. *Advances in Health Sciences Education,* 21(3), pp. 719-729.

Schuman, H. & Harding, J., 1963. Sympathetic identification with the underdog. *Public Opinion Quarterly,* 27(2), pp. 230-241.

Scott, J., 2014. *A matter of record: Documentary sources in social research.* s.l.:John Wiley & Sons.

Shaw, R., 2001. Why use interpretative phenomenological analysis in health psychology?. *Health Psychology Update,* 10(4), pp. 48-52.

Sheridan, T. B. & Nadler, E. D., 2006. *A review of human-automation interaction failures and lessons learned,* Cambridge, MA: Volpe National Transportation System Center.

Shin, I. J., Busby, J. S., Hibberd, R. E. & McMahon, C. A., 2005. *A theory-based ontology of design induced error.* Melbourne, Australia, s.n.

Siemens, G., 2008. *Description of connectivism. Connectivism: A learning theory for today's learner.* [Online]
Available at: http://www.connectivism.ca/about.html

Silva, S. S. & Hansman, R. J., 2015. Divergence Between Flight Crew Mental Model and Aircraft System State in Auto-Throttle Mode Confusion Accident and Incident Cases. *Journal of Cognitive Engineering and Decision Making,* 9(4), p. 312–328.

Silverman, D., 2006. *Interpreting Qualitative Data.* 3rd ed. London: Sage.

Singh, R. & Teoh, J. B. P., 2000. Impression formation from intellectual and social traits: Evidence for behavioural adaptation and cognitive processing. *British Journal of Social Psychology,* 39(4), pp. 537-554.

Smith, B. & McGannon, K., 2018. Developing rigor in qualitative research: Problems and opportunities within sport and exercise psychology. *International review of sport and exercise psychology,* 11(1), pp. 101-121.

Smith, D. et al., 2017. Cooperation and the evolution of hunter-gatherer storytelling.. *Nature communications,* 8(1), pp. 1-9.

Smith, E., Borgvall, J. & Lif, P., 2007. *Team and collective performance measurement.,* Bedfordshire, UK: Defence Science and Technology Lab.

Smith, J., 1964. Group selection and kin selection. *Nature,* 201(4924), pp. 1145-1147.

Smith, J. A., 2004. Reflecting on the development of Interpretative Phenomenological Analysis and its contribution to qualitative research in psychology. *Qualitative Research in Psychology,* Volume 1, pp. 39-54.

Smith, J., Flowers, P. & Larkin, M., 2009. *Interpretative phenomenological analysis: Theory, method and research.* s.l.:Sage.

Smith, R., 1983. Why soldiers fight: Part I. Leadership, cohesion and fighter spirit.. *Quality and Quantity,* 18(1), pp. 1-32.

Sobieszek, R. A., 1999. *Ghost in the shell: photography and the human soul, 1850-2000: essays on camera portraiture..* s.l.:Los Angeles County Museum of Art..

Social Research Association, 2017. *Code of Safety for Socail Researchers.* [Online]
Available at: http://the-sra.org.uk/sra_resources/safety-code/
[Accessed 31 January 2017].

Soeters, J. L., Winslow, D. J. & Weibull, A., 2006. Military culture. In: *Handbook of the Sociology of the Military.* Boston, MA: Springer, pp. 237-254.

Sparkes, J. & Huf, S., 2003. *Mental Models Theory and Military Decision-Marking: A Pilot Experimental Model,* Edinburgh, Australia: Defence Science and Technology Organisation.

Spencer-Oatey, H., 2004. *Culturally speaking: Managing rapport through talk across cultures.* 3rd ed. London: Continuum.

Stanton, N., 2011. *Trust in military teams..* s.l.:Ashgate Publishing, Ltd..

Stanton, N. A. & Young, M. S., 2005. Driver behaviour with adaptive cruise control. *Ergonomics,* 48(10), pp. 1294-1313,.

Stanton, N. et al., 2006. Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. *Ergonomics,* 49(12-13), pp. 288-1311.

Star, S. L., 1992. The skin, the skull and the self: Towards a sociology of the brain.. In: A. Harrington, ed. *So Human a Brain: Knowledge and Values in the Neurosciences.* Boston, MA: Birkhäuser, pp. 204-228.

Steinfeld, A. et al., 2006, March. *Common metrics for human-robot interaction..* Salt Lake City. Utah, USA, ACM, pp. 33-40.

Stokes, C., 2017. *Developing and managing the unique relationships envisioned for future quantified warriors and their machine partners,* s.l.: Air Force Research Laboratory/HMSS.

Stokes, C. K. et al., 2010. *Accounting for the human in cyberspace: Effects of mood on trust in automation.* s.l., 2010 International Symposium on Collaborative Technologies and Systems, pp. 180-187.

Strait, M. et al., 2015. *Too much humanness for human-robot interaction: exposure to highly humanlike robots elicits aversive responding in observers.* Seoul, Republic of Korea, Association for Computing Machinery, pp. 3593-3602.

Strauss, A. & Corbin, J. M., 1990. *Basics of qualitative research: Grounded theory procedures and techniques..* Newbury Park, CA.: Sage.

Strauss, A. & Corbin, J. M., 1990. *Basics of qualitative research: Grounded theory procedures and techniques..* s.l.:Sage Publications.

Stroup, D. et al., 2000. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA,* 283(15), pp. 2008-2012.

Sturgis, P. a. S. P., 2010. Assessing the validity of generalized trust questions: what kind of trust are we measuring?. *International journal of public opinion research,* 22(1), pp. 74-92.

Swann, W. B. J. et al., 2014. What makes a group worth dying for? Identity fusion fosters perception of familial ties, promoting self-sacrifice. *Journal of Personality and Social Psychology,* 106(6), pp. 912-926.

Szczuka, J., Hartmann, T. & Krämer, N., 2019. Negative and Positive Influences on the Sensations Evoked by Artificial Sex Partners: A Review of Relevant Theories, Recent Findings, and Introduction of the Sexual Interaction Illusion Model. In: Y. Zhou & M. H. Fischer, eds. *AI Love You.* s.l.:Springer, pp. 3-19.

Sztompka, P., 1999. *Trust: A Sociological Theory.* Cambridge, UK: Cambridge University Press.

Taherdoost, H., 2018. A review of technology acceptance and adoption models and theories. *Procedia Manufacturing,* Volume 22, pp. 960-967.

Tajfel, H., 1979. Individuals and groups in social psychology. *British Journal of Social and Clinical Psychology,* 18(2), pp. 183-190.

Talone, A. B., Phillips, E., Ososky, S. & Jentsch, F., 2015, September. *An Evaluation of Human Mental Models of Tactical Robot Movement.* s.l., SAGE Publications, pp. 1558-1562.

Tang, Y. et al., 2012. Using argumentation to reason about trust and belief. *Journal of Logic and Computation,* 22(5), pp. 979-1018.

Tanis, M. & Postmes, T., 2005. A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour. *European Journal of Social Psychology,* 35(3), pp. 413-424..

Tashakkori, A. & Teddlie, C., 2010. *Sage handbook of mixed methods in social & behavioral research..* s.l.:Sage.

Techopedia.com, 2015. *What is a Digital Immigrant? - Definition from Techopedia.* [Online] Available at: https://www.techopedia.com/definition/28139/digital-immigrant

Teddlie, C. &. Y. F., 2007. Mixed methods sampling: A typology with examples.. *Journal of Mixed Methods Research,* 1(1), pp. 77-100.

Tesch, R., 1990. *Qualitative research: Analysis types and software tools.* New York, NY: Falmer.

Tetlock, P. et al., 2000. The psychology of the unthinkable: taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of personality and social psychology,* 78(5), pp. 853-870.

The Quality Assurance Agency for Higher Education, 2014. *The Frameworks for HE Qualifications of UK Degree-Awarding Bodies.* [Online]
Available at: https://www.qaa.ac.uk/en/quality-code/qualifications-and-credit-frameworks
[Accessed 22 December 2019].

Thomas, E., Amiot, C., Louis, W. & Goddard, A., 2017. Collective self-determination: How the agent of help promotes pride, well-being, and support for intergroup helping. *Personality and Social Psychology Bulletin,* 43(5), pp. 662-677.

Tinwell, A., Nabi, D. & Charlton, J., 2013. Perception of psychopathy and the Uncanny Valley in virtual characters. *Computers in Human Behavior,* 29(4), pp. 1617-1625.

Toulmin, S., 2003. *The Uses of Argument.* 2nd ed. Cambridge, UK: Cambridge University Press.

Tracy, S. J., 2010. Qualitative quality: Eight “big-tent” criteria for excellent qualitative research. *Qualitative inquiry,* 16(10), pp. 837-851.

Trochim, W. M., 2006. *Research Methods Knowledge Base: Qualitative measures,* s.l.: Atomic Dog Publishing,

Tullio, J., Dey, A. K., Chalecki, J. & Fogarty, J., 2007, April. *How it works: a field study of non-technical users interacting with an intelligent system..* s.l., ACM, pp. 31-40.

Turkle, S., 2010. In good company?. In: *Close engagements with artificial companions.* s.l.:John Benjamins, pp. 3-10.

Turner, J. et al., 1987. *Rediscovering the social group: a self-categorization theory.* Oxford, UK: Blackwell.

Tversky, A. & Kahneman, D., 1974. Judgment under Uncertainty: Heuristics and Biases. *Science,* 185(4157), pp. 1124-1131.

Tylor, E., 1871. *Primitive culture: researches into the development of mythology, philosophy, religion, art, and custom.* London: John Murray.

Uhl-Bien, M., Marion, R. & McKelvey, B., 2007. Complexity leadership theory: Shifting leadership from the industrial age to the knowledge era. *The leadership quarterly,* 18(4), pp. 298-318.

Venkatesh, V. & Morris, M., 2000. Why don't men ever stop to ask for directions? Gender, social influence and their role in technology acceptance and usage behavior. *MIS quarterly. 2000,* 24(1), pp. 115-139.

Wang, L., Jamieson, G. A. & Hollands, J. G., 2009. Trust and Reliance on an Automated Combat Identification System.. *Human Factors,* 51(3), p. 281–291.

Waterson, P. et al., 2015. Defining the methodological challenges and opportunities for an effective science of sociotechnical systems and safety.. *Ergonomics,* 58(4), pp. 565-599.

Watt, M., Maitland, R. & Gallagher, C., 2017. A case of the "heeby jeebies": An examination of intuitive judgements of "creepiness. *Canadian Journal of Behavioural Science,* 49(1), pp. 58-69.

Webb, C., 2013. *History of Graphic Design: Infographic.* Albuquerque(New Mexico): Behance.

Weber, L., Weber, L. & Carter, A., 2003. *The social construction of trust..* s.l.:Springer Science & Business Media.

Wegner, D. M. & Gray, K., 2017. *The mind club: Who thinks, what feels, and why it matters..* s.l.:Penguin.

Westin, C., Borst, C. & Hilburn, B., 2016. Strategic Conformance: Overcoming Acceptance Issues of Decision Aiding. *IEEE Transaction on Human-Machine Systems,* 46(1), pp. 41--52.

Weyers, B., 2012. *Reconfiguration of User Interface Models for Monitoring and Control of Human-Computer Systems,* Essen, Germany: s.n.

Whitehouse, H., McQuinn, B., Buhrmester, M. & Swann, W., 2014. Brothers in Arms: Libyan revolutionaries bond like family. *Proceedings of the National Academy of Sciences,* 111(50), pp. 1783-1785..

Wickens, C., Mavor, A., Parasuraman, R. & McGee, J., 1998. *The future of Air Traffic Control: Human operators and automation, . National Research Council, Commission on Behavioural and Social Sciences and Education. Washington, D.C. N.* Washington, D.C.: National Academy Press..

Wiese, E. & Weis, P., 2020. It matters to me if you are human: Examining categorical perception in human and nonhuman agents. *International Journal of Human-Computer Studies,* Volume 133, pp. 1-12.

Wilkinson, S., 2011. Focus group research. In: D. Silverman, ed. *Qualitative Research.* 3rd ed. London: Sage, pp. 168-184.

Wilkison, B. D., Fisk, A. D. & Rogers, W. A., 2007, October. *Effects of mental model quality on collaborative system performance.* s.l., Sage Publications, pp. 1506-1510.

Wilson, H. S. & Hutchinson, S. A., 1991. Triangulation of qualitative methods: Heideggerian hermeneutics and grounded theory. *Qualitative Health Research,* 1(2), pp. 263-276.

Wogalter, M. S., 2006. Communication-human information processing (C-HIP) model.. In: *Handbook of warnings.* s.l.:CRC Press, pp. 51-61.

Woodward, R. & Jenkings, N. K., 2011. Military identities in the situated accounts of British military personnel. *Sociology,* 45(2), pp. 252-268.

Wu, K. et al., 2011. A meta-analysis of the impact of trust on technology acceptance model: Investigation of moderating influence of subject and context type. *International Journal of Information Management,* 31(6), pp. 572-581.

Xu, Y. J., 2008. Gender disparity in STEM disciplines: A study of faculty attrition and turnover intentions. *Research in Higher Education,* 49(7), pp. 607-624.

Yamagishi, T., 2001. Trust as a form of social intelligence.. In: K. S. Cook, ed. *Trust in Society.* New York: Russell Sage Foundation, pp. 121-147.

Yanow, D., 2007. Qualitative-interpretive methods in policy research. In: F. Fischer, G. Miller & M. S. Sidney, eds. *Handbook of public policy analysis: theory, politics, and methods.* Boca Raton, FL: Taylor & Francis, pp. 405-416.

Yilmaz, K., 2013. Comparison of quantitative and qualitative research traditions: Epistemological, theoretical, and methodological differences. *European Journal of Education,* 48(2), pp. 311-325.

Yip, J., Ehrhardt, K., Black, H. & Walker, D., 2018. Attachment theory at work: A review and directions for future research. *Journal of Organizational Behavior,* 39(2), pp. 185-198.

Zak, P. J., 2017. *Trust Engenders Trust: A Q&A with Paul J. Zak: The neuroscientist's research shows that a culture of trust leads to better business outcomes* [Interview] (27 February 2017).

Zhang, T., Kaber, D. & Hsiang, S., 2010. Characterisation of mental models in a virtual reality-based multitasking scenario using measures of situation awareness. *Theoretical Issues in Ergonomics Science,* 11(1-2), pp. 99-118.

Zimbardo, P. G., 1969. (1969). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. *Nebraska Symposium on Motivation*, Volume 17, pp. 237-307.

PAGE LEFT INTENTIONALLY BLANK

Table 34 - Summary Table of Included Sources (System Related // Performance Based)

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
Barg-Walkow, (2013). (Barg-Walkow, 2013)	Written description of AWMS with verbal description of system performance framing of Levels of reliability (>90%, <60% and ~75% (actual reliability)). Dual-task scenario, 8 experiment blocks, with 20 trials per block.	Assess how UX (such as expectation and level of system reliability) influences human responses (perception, compliance, reliance, dependence).	N= 60 (38%;62%19.8 (0.21 SD)) 54 Native English speakers, 6 fluent. Multicultural (40 Caucasian, 8 Black, 10 Asian, 1 Mixed race, 1 other/unspecified)	Automated warehouse management system (AWMS) Georgia Institute of Technology (North America)	ANOVA Visual acuity (Snellen), Perception Speed (Digit Symbol Substitution test (Wechsler)), Memory span (Reverse Digit Span test, Wechsler), Verbal acuity (Shipley Vocabulary test, Shipley). Automation Experience Questionnaire, Automation Attitudes Questionnaire, Automation-Induced Complacency Potential Questionnaire, Demographic and Health Questionnaire, Expectancy Questionnaire, Perceptions Questionnaire.	Perceived reliability did not converge with actual reliability. Explicit statement groups had more stable perceptions of system reliability than initial exposure groups, **NOTE:** contains metanalysis of Studies Involving Automation Reliability Expectations for Explicit Statements (ES) and Initial Exposures (IE)
Cassidy (2009) (Cassidy, 2009)	Three phases of 40 tasks requiring threat DRI (120 total per P), G1 received minimal info, G2 accurate info, G3	Examine the relationship between an operator's mental model of how an automated aid	N=42 (26%;74%)	Naval Postgraduate School of	N/A	H1: Trusts automation less but relied on it more for task DRI.

	inaccurate info. Following presentation of each image with a threat, Ps decided whether to utilise assistive automation.	works and the appropriateness of their reliance on the aid.	Ages between 21-45. Military service experience between 0-4 to 20+ (~45% between 5-10). Majority of Ps have some experience with computer-based games.	California (North America)		H2: Dichotomy between automation performance and operator perception of trustworthiness. Although G3 HMM of the TDD performance because more accurate, trust still declined. H3: Prior information affect HMM and reliance on automation.
Dawson, Crawford, Dillon, & Anderson (2015) (Dawson, et al., 2015, May)	Mixed design with between-subject variable (algorithm training and STEC (*spatial and temporal environmental cues*) vs none) and within-subject variables (surveillance approach) Operators provided with UI to monitor UAVs using a 3D physics based sim enviro for a multi-agent surveillance task. No training vs algorithm training presentation to explain system behaviours.	Whether operator trust can be affected so that a single operator can more effectively team with multiple system agents while interacting with high LOAs.	40 68% with software development experience. 28% with military experience.	North America	3D SART and NASA-TLX questionnaires. Trust in Automation questionnaire (Bisantz and Drury, 2000) Two-way ANOVA IV: surveillance approach and algorithm training. DV: trust, workload, situation awareness, performance and usability.	Results regarding increase in workload and lowered perceived SA are consistent with the literature. Training was not significantly helpful and, in some cases, detrimental to trust. Therefore the results show that operators may have trouble relating the training to the interface. Results showed that STEC is key for operators to understand autonomous robot/agent behaviour.
Mosier et al (2013) (Mosier, et al., 2013)	Discussion piece	Automation failure and errors impact on pilot judgements of Human Automation Interaction (HAI).	N/A	N/A	Literature Review	The paper discusses the negative impact of poor observability of the system and mental models on the automation's functionality. Furthermore, errors can compound on existing inadequate mental models with

						monitoring and awareness issues complicating faulty or incomplete MMs.
Shaefer, Evans & Hill (2015) (Schaefer, et al., 2015)	Discussion/Final military report.	The impact of Autonomous Intelligent Military Robots integration and the changing independencies of human-technology teams (embodied and non-embodied system agents).	N/A	N/A	Literature Review	The change in C2 away from directly controlled operations towards automation and supervisory roles will have an impact on performance if development and implementation of these systems are not carefully considered. Consideration must be noted for the paradigm of traditional C2 displays (head-down visual display) to multimodal feedback and integration. The relationship between communication and Shared MM/Individual MM is key for the development and calibration of trust.
Westin, Borst & Hilburn (2015) (Westin, et al., 2016)	Discussion	To introduce 'strategic conformance' – or the degree to which system problem solving styles align with the operator – and its influence on the acceptance of the automation.	N/A	N/A	Literature Review Modelling	This paper sought to expand upon the already existing Technology Acceptance Model, with a specific focus on automation and human system interaction. The paper explores design heuristics limitations and compatibility limitations of both system and human agents and that strategic conformance can be used as a mediator to overcome resistance to accept decision aiding automation.

Table 35 - Summary Table of Included Sources (System Related // Attribute Based)

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
Andersson (2010) (Andersson, 2010)	Field Observational – Nuclear Power Plant Control Room	To propose a conceptual model that capture aspects of human-machine interaction that are relevant for analysing automation usability problems To use the conceptual model for analysis of automation usability related problems and explain the emergence of different types of problems in a single framework.	N/A	Sweden (EU)	Model Hierarchy of Control loop and HSI. Abstraction/Decomposition Model of Work Domain Analysis. Qualitative - commentary	NOTE: Figure 22 Model-reality mismatches that may lead to degraded operator performance
Sheirdan & Nadler (2006) (Sheridan & Nadler, 2006)	Review	To review human-automation failures and misunderstanding as a result of mismatched mental models in pilots and the behaviour of automation systems.	N/A	N/A	Accident Analysis	Human-automation and system failure is discussed at length in this report. Human related issues (such as allocation of function, situational awareness and behaviour characteristics are discussed. Furthermore, system relation factors such as level of automation, trust and over-reliance are also explored. These are all factors influencing mental models which are covered at length within this report. Recommendations, such as including behavioural and cognitive paradigms in the design requirements, are also given.
Shin, Busby, Hibberd & McMahon	Discussion	To provide recommendation in the identification of design	N/A	N/A	Literature and Pilot Study (results not published).	Rationales for user performance are compared across designers and operators. These include risk factors, reliability of

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
(2005) (Shin, et al., 2005)		induced errors (DiE) to aid in providing design solutions with HCI/HSI.			Accident analysis and modelling.	automation, trust in automation and automation failures. The mental model frameworks of each rational differs and it is the inappropriate frameworks which have an impact on design and system interaction by operators. Overall, the complexity of increased level of automation create new mechanisms for user error. Psychological research (informing and changing their mental models towards the operator) into the information technology is required by designers to facilitate appropriate interaction.
Silva & Hansman (2015) (Silva & Hansman, 2015)	Discussion/Accident report review	To formally define the divergence of flight crew mental models and actual system state in auto-throttle mode confusion in aviation accidents and incidents.	N/A	N/A	Accident Analysis	This article reports breakdowns in perception, comprehension, and projection based on the observed SA issues in 10 aviation incidents where mental model divergence with automation system states occurred. The three main factors outlined were attentional limitation, masked feedback and shared mental models/team communication. The article serves to introduce mitigation strategies to target causes of mental model divergence or to aid in facilitating reconvergence.

Table 36 - Summary Table of Included Sources (Environmental // Team Collaboration)

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
Chen (2013) (Chen & Barnes, 2013)	Literature Review	Explore the literature around Human-Agent Teaming for Multi-Robot control. Namely: Efficient human supervision of multiple robots. Appropriate human trust in the automated systems. Maintenance of human operator's SA. Individual differences in human-agent (H-A) interaction. Retention of human decision authority.	N/A	N/A	Systematic Review	**NB:** 'Robots' include unmanned systems and autonomous/intelligent systems. Term agent for the broader class of intelligent systems without physical embodiment. Fan (2010) (Fan, et al., 2010): H-A out-performed H-H in C2 sims when shared mental models were acknowledge (using R-CAST agent architecture to enable appropriate information exchange). Ecological Interface Design help to achieve better mental models of systems (Furukawa & Parasuraman. 2003 (Furukawa, 2003, October))
Hawley, Mares & Giammanco (2006) (Hawley, et al., 2006)	Recommendation report/Literature Review	Review of appropriate mental model and formation support in C2 and the effects of training.	N/A	N/A	Literature Review	*"When the learner has a proper conceptual framework or mental model, accretion is easy, painless, and efficient. Without a good conceptual framework, accretion*

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
						is slow, tedious, and error-prone."
Joe, O'Hara, Medema & Oxstrand (2014) (Joe, et al., 2014, June)	Review	To identify requirements and recommendations to facilitate appropriate multi-agent teamwork with automation.	N/A	N/A	Literature Review	Supports other research into multi-agent teams which move away from past research into human-system teams which seek to mimic H-H teams/behaviour and that this is a core dynamic within mental models of trust.
Morita & Burns (2014) (Morita & Burns, 2014)	Observation study of scenario		200	University of Waterloo (Canada)	SA and Brunswick Lens Model Framework to influence information synthesis into a mental model.	Breaks down trust into the more socially influenced 'Intuitive trust' and analytical 'calculated confidence' – these echelons are used to breakdown the perception mechanisms within the mental models. The paper seeks to organize the antecedents and regulating factors that affect type of trust and influences on trust formation.
Ososky (2013) (Ososky, 2013)	Two-way between subjects MANOVA. The study manipulated the preparatory information provided to participants with either congruent or incongruent information to the task-role of the robot.	The influence of task-role mental models of robots on interpreting robot behaviour	189 in initial recruitment Total N=120 (55%;45%; 18.78 (SD 1.61) Undergraduate psychology student population.	University of Central Florida. (North America)	ANOVA	Support the hypothesis that with a congruent mental model Ps were had greater accuracy interpreting behaviour than with an unsupported mental model in addition to statistically significant hindrance to accuracy in Ps with incongruent MMs. The strength of effect of MM was moderated by ease of explanation through 'mental model applicability'.

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
						Positive associations found in observers' interpretation accuracy and differences in subjective ratings of robot intelligence, safety, and trustworthiness.
Phillips, Ososky & Jentsch (2014) (Phillips, et al., 2014, September)	Literature Review	To explore the use of metaphors to aid in the facilitation of accurate mental models in human-technology teams in military conditions.	N/A	N/A	Literature Review	Further research to support and expand on from past literature review (Phillips, et al., 2011, September). Explores human-technology teams and human teaming with automation systems and mental models surrounding transference of human behaviour on to non-human teammates in addition to human cognitive capacity. Provides recommendations for future research and applications.
Phillips, Ososky, Grove & Jentsch (2011) (Phillips, et al., 2011, September)	Literature Review	To investigate the factors that affect developing appropriate mental models to robotic teammates specifically in military conditions.	N/A	N/A	Literature Review	In depth analysis of antecedents to trust, misuse and disuse of automation/robotics using a mental model framework for military specific applications. The paper explores primitive mental models, social interaction (social robots) and viewing

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
						robotics/systems as equipment or teammates.
Sætrevik & Eid (2013) (Sætrevik, 2013)	EXP1: Field experiment Daily work scenario (between 2-3hr tasking) with the inclusion of a critical situation (fire/gas leak on facility) with dedicated 'freeze points' for probing questions – to monitor frequency and detail of communication. EXP2: Controlled experiment.	To investigate a methodology where the extent of sharing information among teammates as an indicator of shared mental models and situational awareness.	N/A	Second-line emergency preparedness centre in a large hydrocarbon energy company. University of Bergen, Norway (EU)	SME Observer/TADMUS One-way ANOVA	The objective in EXP1 was to examine the extent to which individual belief coincided with team beliefs. EXP2 sought to obtain a similarity index between individual beliefs and top-performing individuals' belief. The overall assumption of the research is that shared cognitive states are beneficial to team processes. Higher degrees of shared information between team members reflected higher degrees of SMM and SA and lead to improved team performance. Additionally, a misinformed leader (or a well-informed leader who has failed to communicate appropriately) negatively affect team members' similarity index scores.
Schaffernicht & Groesser, (2011) (Schaffernicht & Groesser, 2011)	Provide a meta-analysis to compare MMDS. They explore the distance approach (Markóczy and Goldberg, 1995), and thereafter, enhance the distance approach by accounting for delays between cause and	The paper proposes to explore the existing approaches to compare Mental Models in Dynamic Systems (MMDS)	N/A	N/A	Meta-analysis Distance Ratio, Element distance ration, loop distance ratio and model distance ratio	The comprehensive method proposed is to be used to compare several versions of mental models both within and between subjects. Utilised with solely statistical Quantitative data and explicit variables and highly controlled experimental environment.

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
	effects, feedback loops and polarities.					
Smith, Borgvall & Lif(2007) (Smith, et al., 2007)	Exploratory report of the literature with a focus on measures and techniques in assessing team SA and shared mental models.	VR and operator performance (individual and team level) in a military context.	N/A	N/A	Literature Review SALIENT (Situational Awareness Linked Instances Adapted to Novel Tasks) IOC (Integration Organisation and Cohesion) – used to assess how well a collective is working together and used to quantify the extent to which collective training has an impact. Bedford rating scale (has correlations with mental workload, situational awareness and heart rate)	The literature outlines the metrics of exploring experimental and dynamic individual and shared mental models. The core focus of this report is the importance of appropriate team MMs for effective team communication and achieving goals.

Table 37 - Summary Table of Included Sources (Environmental // Tasking)

Author	Study Protocol	Hypothesis	No. of participants. N(%female;%male;mean age (y))	Setting	Method of analysis	Results
Clancey, Linde, Seah, & Shafto (2013) (Clancey, et al., 2013)	To use Brahms simulation model of actors and events identified using Events and Causal Factors Analysis to identify human/automation error in authority and autonomy conflicts.	To provide a technique to evaluate and test Authority vs Autonomy conflicts	N/A	Part of the "Authority and Autonomy" Task within the Aviation Safety Program (AvSP) to design a verification and validation tool (V&V) in aid of assessing safety critical situations involving autonomous systems in aviation. NASA Ames Research Centre, California (North America)	Events and Causal Factors Analysis	Mental models are touched upon in developing the Brahms framework as an understanding of how people monitor and understand system behaviour, in addition to the operational role determines what model of their environment is useful. Exploration of mental models with the Brahms GU Model is noted in future research recommendations. Brahms GÜM can be augmented to represent agent and object ontologies.

Table 38 - Summary Table of Excluded Sources

Author	Reason for Exclusion
Archer, J. (2012). (Archer, 2012)	Although discusses the use of mental models, it is primarily an overview scope of problems in aviation, in association with situational awareness and performance workload. After review, there is no useful information which can be gleamed from this report.
Billings et al (2012) (Billings, et al., 2012)	Although discusses mental models in the discussion, the scope of this literature review only explores current research into comparing human-animal relationships to human-robot interaction rather than the underlying mental models which may cause individuals to act similarly towards animals/robots.
Bolton, Bass & Siminiceanu, (Bolton, et al., 2013)	Discussed mental models but as a secondary aspect into producing a formal verification algorithm used to mathematically assist in appropriately scaling models of a system to return desirable or undesirable properties.
Chengqi., Cen, & Yan. (2009, November). (Chengqi, et al., 2009, November)	Although Mental Models were discussed in the article, it would be more accurately described as a heuristic design characteristic and was only utilised in a minute HCI change for a concept proposal.
Combéfis, Giannakopoulou, Pecheur, & Feary (2011) (Combéfis, et al., 2011, October)	The context of a mental model in this article is "*not meant to capture a human cognitive model; rather, it is meant to capture the implicit and intended model of operation according to which the system developer designs the system".*
Imranali, Jain, Lal Patidar & Neeraj (2012). (Imranali, et al., 2012)	Discussed literature around the design of ubiquitous computing and intelligent systems however primarily discussed computer logic and the framework outlined was standard user experience protocol.
Chauvin (2011) (Chauvin, 2011)	The discussion mentions mental models and shared mental models as a consideration into reasons for human error – however this paper is primarily about the types of accident models utilised in psychology and human factors assessments of accidents.
Chen (2014) (Chen & Barnes, 2014)	Journal article following conference proceeding of Chen (2013) (Chen & Barnes, 2013) – Repeat of information already noted in this literature review.
Clegg, Heggestad & Blalock (2010) (Clegg, et al., 2010, September)	Explored the influences of automation and trainee aptitude on training effectiveness, which showed that automation in training varies with types of automation, task but also aptitude of operators. However, mental models were only mentioned in the discussion as future research.
Deshmukh, McComb & Wernz (2008) (Deshmukh, et al., 2008)	Mental models are only mentioned as a theoretical framework to approaching issues with shared and individual situational awareness.

Author	Reason for Exclusion
Dey (2008) (Dey, 2008, July)	The approach to mental models was both inappropriate for this project (basic learning concept methodology), in addition to mainly focussed on elderly and cognitively impaired participants.
Dey (2009) (Dey, 2009)	Similarly, to (Dey, 2008, July), the approach to mental models was not appropriate for this literature review. Furthermore, the focus was towards design elements and system architecture rather than psychosocial/sociotechnical behaviour.
Endsley, Hoffman, Kaber & Roth (2007) (Endsley, et al., 2007)	Although a substantial overview of the subject of cognitive engineering and decision making, this article is a proposal for a new journal to incorporate research specific to these domains.
Endsley (2015) (Endsley, 2015)	Similar to (Endsley, et al., 2007), this article has insightful discussion pieces, it is primarily a communication/response to misinterpretations of her well known situational awareness model and therefore not relevant to this literature at present.
Fan. Chen & Yen (2010) (Fan, et al., 2010)	The experimental study is regarding shared mental models in expert multi-agent teams; however, the focus is on measuring (and simulating) team cognitive load for future influence into Hidden Markov Model-based cognitive load models. Automation is not a factor in this study.
Fan, McNeese and Yen (2010) (Fan, et al., 2010)	Similarly, to (Fan, et al., 2010), the discussion of mental models is computational simulation framed rather than psychosocial/cognitively framed.
Feigh, Dorneich & Hayes (2012) (Feigh, et al., 2012)	Mental models were touched upon in discussion, however the majority of the paper was on characteristics of developing adaptive automation.
Goodrich and Schultz (2007) (Goodrich & Schultz, 2007)	A comprehensive lay-man report into themes, trends and theories within the domain of Human-Robot Interaction; however, mental models are only mentioned in passing in sections 4.4 (Adaption, Learning and Training) and 7.2 (Human Factors and Automation Science).
Gutzwiller, Clegg & Blitch (2013) (Gutzwiller, et al., 2013)	Mental models are acknowledged in the development of usability of automation; however, they are not assessed within the confines of the experiment.
Hancock, Billings & Shaefer (2011) (Hancock, et al., 2011)	Overview of trust with HR interaction. Mental models acknowledge but not explored.
Hughes, Rice, Trafimow & Clayton (2009) (Hughes, et al., 2009)	Study into the comparison of attitudes of novice pilots with human-pilot VS autopilot. Valid for exploring trust in automation with age/expertise as a variable. However, mental models are not discussed.

Author	Reason for Exclusion
Keel, Cooke & Sither (2014) (Keel, et al., (2014).)	Mental models are discusses and appropriately considered in this article, however, the context (computational collaboration, such as tele-conferencing) is not appropriate to this review.
Kiesler (2005) (Kiesler, (2005, August). . In 2005. ROMAN 2005. (pp. 729-734). IEEE.)	Mental models are covered throughout this paper in fostering common-ground with robotic agents, however, the context of anthropomorphism and robotics, rather than systems and automation.
Krueger & Banderet (2007) (Krueger & Banderet, 2007)	The focus of this paper is exploring the implications of studying team cognition and team performance in future system warfare paradigms, utilising interconnected comms and technology. Although mental models are briefly mentioned, they are not expanded upon and therefore not appropriate for this review.
Kulesza (2012) (Kulesza, 2012, February)	Relevant research question exploring the mental models which support interaction with intelligent systems, however this paper only outlines the research questions – related article as yet unfound.
Langan, Canty & Sankey (2009) (Langan-Fox, et al., 2009)	The research acknowledges the importance of mental models in human-automation teams; however, this is not the focus of the paper which is automation and system adaptability.
Madni (2010) (Madni, 2010)	This paper focuses on the technical and interaction limitations of human and system integration. Human performance (and to a point, mental models) is discussed but not expanded in any detail.
Madni (2011) (Madni, 2011)	This magazine article (featured in The Journal of Defence Software Engineering) is a summary of points earlier remarked in Madni (2010) (Madni, 2010)
Miller & Parasuraman (2007) (Miller & Parasuraman, 2007)	The paper explores the recommendations for designing for flexible integration between human-automation teams (such as adaptive automation) and delegation interfaces to facilitate supervisory control with various levels of automation. However, neither trust nor mental models are explored within this paper and is therefore excluded from this study.
Mindock & Klaus (2012) (Mindock & Klaus, 2012)	Spaceflight specific context and explores the macro-cognition and –ergonomics of the organisational structure. Mental models are outlined but are not an integral part of the thesis but act more as a consideration for shared communication and human performance capabilities.
Nilsson, Riveiro & Ziemke (2008) (Nilsson, et al., 2008)	The report focuses on information fusion in military environment. Mental models are discussed as a cognitive variable with human operators but not expanded upon throughout the document.
Nothdurft & Minker (2014) (Nothdurft & Minker, 2014, January)	A continuation of research initialised in (Nothdurft, et al., 2013, April) however, continues onto explore companion computing which is not relevant to this document.

Author	Reason for Exclusion
Parush, Ahuvia & Erev (2007) (Parush, et al., 2007)	The paper does not explore the use of mental models. However, this experiment does look into the effects of automation and out-of-the-loop behaviour.
Prewett, Johnson, Saboe, Elliott & Coovert (2010) (Prewett, et al., 2010)	Literature review of issues surrounding mental workload in human and synthetic agent interaction. Mental models is referred to in a single reference which is already covered in this document (Cassidy, 2009 (Cassidy, 2009))
Salmon et al (2008) (Salmon, et al., 2008)	Although mental models are discussed with regards to teamwork and individuals, it is primarily used as a framework for SA or interchangeably with situational awareness.
Schöbel, M. (2009) (Schöbel, 2009)	The impact of trust with HCI in high-reliability organisations is the main focus of this article, however mental models as a framework for developing or facilitating trust is not covered, only that mental models are a cognitive element to individuals' interactions with technology.
Stanton & Young (2005) (Stanton & Young, 2005)	Situational awareness and mental models are used interchangeably regarding the cognitive framework of participants.
Stanton et al (2006) (Stanton, et al., 2006)	Situational awareness and mental models are used interchangeably regarding the cognitive framework of participants.
Steinfeld, Fong, Kaber, Lewis, Scholtz, Schultz & Goodrich (2006) (Steinfeld, et al., 2006, March)	This paper outlines common human factor metrics in task-oriented robotic agents – the performance and operations of the robotic agent and quantitative human metrics are discussed. Mental models are touched upon in relation to interaction with household appliances and GUIs.
Tullio, Dey, Chalecki & Fogarty (2007) (Tullio, et al., 2007, April)	This study was deemed inappropriate for this study due to the type of HSI – The intelligent system was a 'interruptability sensor system' that used existing mental models of subordinate colleagues and social etiquette of an office environment to observe behaviour changes if a digital system was introduced to indicate availability of management instead of social or environmental cues.
Weyers (2012) (Weyers, 2012)	Although mental models are discussed in this doctoral dissertation, the main goal is the development of a formal visual language to model human computer interaction in relation to user interface design.