University of Southampton Research Repository

Copyright © and Moral Rights for this thesis and, where applicable, any accompanying data are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis and the accompanying data cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content of the thesis and accompanying research data (where applicable) must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holder/s.

When referring to this thesis and any accompanying data, full bibliographic details must be given, e.g.

Thesis: Author (Year of Submission) "Full thesis title", University of Southampton, name of the University Faculty or School or Department, PhD Thesis, pagination.

Data: Author (Year) Title. URI [dataset]

University of Southampton

Faculty of Engineering and Physical Sciences

School of Engineering

Human-Machine Interfaces for Future Submarine Control Rooms: Graphical Integration of Sonar and Target Motion Analysis

by

Daniel Thomas Fay

Thesis for the degree of Doctor of Philosophy

November 2023

University of Southampton

ABSTRACT

Faculty of Engineering and Physical Sciences

School of Engineering

Doctor of Philosophy

Human-Machine Interfaces for Future Submarine Control Rooms: Graphical Integration of Sonar and Target Motion Analysis

by

Daniel Thomas Fay

Submarine control rooms are an advanced product of evolution, but this does not preclude further improvements. Whilst the traditional process of incremental interface improvement has met requirements, an evolutionary design approach has unnecessarily retained older constraints, limiting contemporary designs. Consequently, future operational requirements may exceed a control room's capacity to perform effectively, despite highly trained operators utilising advanced technology. This thesis explores the utility of Ecological Interface Design (EID) to address the issue, concentrating jointly on the Sonar and Target Motion Analysis (TMA) roles due to their prevalence in the control room.

Two stages of CWA (Cognitive Work Analysis), Work Domain Analysis (WDA) and Worker Competencies Analysis (WCA), were completed for both roles to understand their operation, key design aspects, and shortcomings. This revealed actionable design insights, including investigating merging the interfaces. These insights were used to design a combined interface named Graphically Integrated Sonar and TMA (GIST), which is a novel application of EID and the roles (due to their merged application) in the literature.

GIST was created using a proposed agile software process designed to help bridge the gap between CWA and EID implementation, adapting prevailing methods to integrate better with contemporary software engineering. Subsequent evaluation of GIST against contemporary Sonar and TMA interfaces in a repeated-measures Human in the Loop study indicated that EID was a suitable design choice for GIST,

with statistically significant improvements observed in objective performance, subjective usability, and subjective workload. GIST was also evaluated against a User-Centred Design Mashup interface as part of this study. This comparison revealed that subjective usability, subjective workload, and a tracker assignment (aliasing data) task were better in the Mashup. However, GIST had better solution positioning (where other vessels are in relation to the submarine). It was inferred that each interface had excelled at items linked to its design process (e.g., GIST was better at a task involving the submarine's environment). Consequently, it was proposed that while the application of EID was beneficial over contemporary designs, a blend of EID and User-Centred Design might yield even better results than individually in future work.

By demonstrating the benefit of EID applied to the problem of future submarine control room design, including merging the Sonar and TMA interfaces, and documenting how this could be achieved, it is hoped that this thesis will serve to inspire future change in submarine control room interface design.

Table of Contents

Table of Tables

Table of Figures

Academic Thesis: Declaration of Authorship

I, **Daniel Thomas Fay**, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Human-Machine Interfaces for Future Submarine Control Rooms: Graphical Integration of Sonar and Target Motion Analysis

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University.
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
3. Where I have consulted the published work of others, this is always clearly attributed.
4. Where I have quoted from the work of others, the source is always given. Except for such quotations, this thesis is entirely my own work.
5. I have acknowledged all main sources of help.
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.
7. Parts of this work have been published as:

Fay, D., Roberts, A.P. and Stanton, N.A. (2019) 'All at Sea with User Interfaces: From Evolutionary to Ecological Design for Submarine Combat Systems', *Theoretical Issues in Ergonomics Science*, 20(5), pp. 632-658.

Fay, D., Roberts, A.P.J. and Stanton, N.A. (2020a) 'Designing Future Submarine Control Room HMIs', *Ergonomics & Human Factors 2020*. Online. Chartered Institute of Ergonomics & Human Factors. Available at: https://publications.ergonomics.org.uk/publications/designing-future-submarine-control-room-hmis.html.

Fay, D., Roberts, A.P.J. and Stanton, N.A. (2020b) 'Interfaces with Legs? Documenting the Design Sprint of Prototype Future Submarine Control Room User Interfaces', *AHFE 2019*. Washington, D.C, USA. Springer International Publishing, pp. 669-680.

Fay, D., Roberts, A.P.J. and Stanton, N.A. (2020c) 'A Tactical Picture Accuracy Assessment Tool for ComTET', *Ergonomics & Human Factors 2021*. Online. Chartered Institute of Ergonomics & Human Factors. Available at: https://publications.ergonomics.org.uk/publications/a-tactical-picture-accuracy-assessment-tool-for-comtet.html.

Fay, D., Stanton, N. and Roberts, A. (2018a) 'Streamlining experimental processes using bespoke software', *Ergonomics & Human Factors 2018*. Birmingham. Available at: https://eprints.soton.ac.uk/429470/.

Fay, D., Stanton, N. and Roberts, A.P.J. (2019) 'Exploring ecological interface design for future ROV capabilities in maritime command and control', *AHFE 2018*. Orlando, FL, USA. Springer, pp. 264-273. Available at: https://eprints.soton.ac.uk/422563/.

Fay, D., Stanton, N.A. and Roberts, A. (2017) 'Designing New Interfaces for Submarines: From Cognitive Work Analysis to Ecological Interface Design', *AHFE 2016*. Orlando, FL, USA. Springer International Publishing, pp. 413-425. doi: 10.1007/978-3-319-41682-3_35. Available at: http://dx.doi.org/10.1007/978-3-319-41682-3_35 (Accessed: 13/11/2023).

Fay, D., Stanton, N.A. and Roberts, A.P.J. (2018b) 'Assessing Sonar and Target Motion Analysis Stations in a Submarine Control Room Using Cognitive Work Analysis', *AHFE 2017*. Los Angeles, CA, USA. Springer International Publishing, pp. 191-198. doi: 10.1007/978-3-319-60441-1_19. Available at: https://doi.org/10.1007/978-3-319-60441-1_19.

Acknowledgements

This work was funded and conducted as part of the Command Team Experimental Testbed (ComTET) project by the Defence Science and Technology Laboratory (DSTL) on behalf of the Ministry of Defence (MoD), contract numbers TIN3.113 (ComTET) and TIN 3.228 (ComTET2). All views expressed are those of the author.

Definitions and Abbreviations

A Class	Astute Class
ADS	Abstraction Decomposition Space
Aft	Back of a boat, inside
AH	Abstraction Hierarchy
Amidships	Middle of the boat
ANOVA	ANalysis Of VAriance
ARCI	Acoustic Rapid COTS Insertion
AWS	AEGIS Weapon System
BBB	Bow Broadband
Boat	A ship or submarine
Bow	Front of a boat
C#	C Sharp
C++	C Plus Plus
C2	Command and Control. May also be stylised as "C^2"
CASD	Continuous at Sea Deterrent
CCCS	Common Core Combat System
CO	Commanding Officer
ComTET	Command Team Experimental Testbed
COTS	Consumer off the Shelf
CPA	Closest Point of Approach
CS2	Common Submarine Combat System
CWA	Cognitive Work Analysis
DCLT	Detection, Classification, Localisation, Tracking
DE&S	Defence Equipment and Support
Deck	Floors/ceilings separating levels of a boat. These levels are also called decks.
DEMON	Detection Envelope Modulation on Noise, Demodulation of Noise

DoD	Department of Defence
DSA	Distributed Situation Awareness
DSTL	Defence Science Technology Laboratory
DT	Dived Tracking
ECDIS	Electronic Chart Display and Information System
EID	Ecological Interface Design
ESM	Electronic Support Measures
FKBB	Flank Broadband
GIST	Graphically Integrated Sonar and TMA
HFI-DTC	Human Factors Integration Defence Technology Centre
HMI	Human-Machine Interface
HMI	User Interface
HMNB	Her/his Majesties Naval Base
HMS	Her/his Majesties Ship
Hull	The exterior body of a boat
IO	Inshore Operations
KBB	Knowledge-Based Behaviour
L	Left
Lofagram	Low-Frequency Analysis and Recording GRAM
LOP	Local Operations Plot
M	Master, or Mike
MAIB	Marine Accident Investigation Branch
MANOVA	Multivariate ANalysis Of VAriance
MFC	Multi-Function Console
MFO	Medium Frequency Oscillations
MILSPEC	Military Specification
MoD	Ministry of Defence
MoDREC	Ministry of Defence Research Ethics Committee

NASA-TLX	National Aeronautics and Space Agency Task Load IndeX
NG	Next Generation
NTSB	National Transportation Safety Board
OOW	Officer of the Watch
OOW	Officer of the Watch
OPSO	Operations Officer
Peri	Periscope Operator
PIS	Participant Information Sheet
Port	The left side of a boat when facing the bow
R	Right
RAN	Royal Australian Navy
RBB	Rule-Based Behaviour
RN	British Royal Navy
RTPD	Return(ing) to Periscope Depth
S	Sierra, or Sonar
SA	Situation Awareness
SBB	Skill-Based Behaviour
SCE	Shared Computing Environment
SD	Standard Deviation
ShC	Ship Control
Ship	A surface boat
SMCS	Submarine Combat System
SNR	Signal-to-Noise Ratio
SoC	Sonar Controller
Solution	Bearing, Course, Range, Speed of Contact
SONAR	SOund NAvigation and Ranging
SoP	Sonar Operator
SRK Taxonomy	Skills, Rules, and Knowledge Taxonomy

SSBN	Nuclear warhead-equipped submarine
SSGN	Cruise missile-equipped submarine
SSN	Nuclear-powered attack submarine
Starboard	The right side of a boat when facing the bow
Stern	Back of a boat, outside
STOPS	Submarine Tactical Operating Procedures
Submarine	A sub-surface boat
SUS	System Usability Scale
SWFTS	Submarine Warfare Federated Tactical Systems
T Class	Trafalgar Class
TANB	Towed Array Narrowband
TMA	Target Motion Analysis
TPK	Turns per Knot
UK RN	United Kingdom's Royal Navy
UoS	University of Southampton
USN	United States Navy
USS	United States Ship
V	Victor, or Visual
V Class	Vanguard Class
WCA	Worker Competencies Analysis
WDA	Work Domain Analysis
WECDIS	Warship Electronic Chart Display and Information System
WPF	Windows Presentation Foundation
WT	Warner Transmissions

Chapter 1 **Introduction**

1.1 Submarines in the United Kingdom and beyond

As an island nation, the United Kingdom (UK) has a proud history of seafaring. It has traded, travelled, and defended on the seas for centuries, and this shows no signs of abating as the world continues to become more globalised. An integral part of the UK's naval capability is the Royal Navy (RN), operating globally to project and protect the values that the UK strives for. The RN website identifies six main objectives that form their operational remit: maintaining international partnerships, preventing conflicts, protecting the economy, providing security at sea, providing humanitarian assistance, and maintaining fighting readiness (Royal Navy, 2019). These objectives are wide-ranging and are not constrained to defence activities, but are a spectrum that directly contributes to society, freedom, prosperity, and security, both in the UK and worldwide.

The submarine service is a key aspect of the RN; the advent of submarines offered new capabilities, which could also be used to complement surface fleet activities. For example, they could escort ships during a humanitarian aid mission to ensure their safe arrival. However, they are most often used where secrecy is tantamount due to their clandestine nature, such as intelligence gathering in non-sovereign waters or maintaining the Continuous at Sea Deterrent (CASD; the UK's nuclear deterrent). This need for secrecy limits the amount of information known about their missions, but the 'Silent Service' appellation speaks volumes; whilst always silent, it continuously stands at the service of the nation and its citizens.

The work for this research was conducted in the UK, and as such there is a heavy RN focus. However, there are other navies with submarines, each demonstrating the same operational commitment, that this research could be utilised by. The work has broader global applicability arising from commonalities between allied navies, and the use of abstracted generic processes and interfaces. For the former, research, technology, and training are shared in partnerships between allied navies, such as the United States Navy (USN; Royal Navy, 2020;2022;2023b) or the Royal Australian Navy (RAN; Ministry of Defence, 2022; Royal Australian Navy, 2023; Royal Navy, 2023a). This means that any outcome from this research is likely to be refined, adapted, and shared where appropriate across one of the many partnerships. Furthermore, existing commonalities, such as their layout and ways of working means that this research could be applied outside of these partnerships, and across other navies that exhibit the same baselines. For the latter, this research was designed to be unclassified and generalisable, irrespective of the specifics of a given navy. Consequently, the work is applicable to submarine control rooms across the globe, where the principles discussed within this thesis are employed.

1.1.1 Control Rooms

The control room is a key aspect of submarine operation, utilising trained operators of different specialisms as well as advanced technology to understand the environment and how operational or strategic goals should be met (Hautamaki, Bagnall and Small, 2005; Stanton and Bessell, 2014; Stanton, Roberts and Fay, 2017; Stanton and Roberts, 2018). It is the predominant location for Command and Control (C2) in a submarine, housing the duty command team and control (sensor and systems) technology, which work in tandem to ensure mission success, whilst maintaining safety and covertness. A comprehensive environmental understanding is important for all submarine operations, from both safety (sea depth, commercial vessels) and tactical (covertness, achieving objectives) perspectives. Submarine control room capabilities are highly advanced, having progressed over several decades (Dominguez *et al.*, 2006; Smith *et al.*, 2013), but this does not mean improvements cannot be made (Stanton, 2014).

Future operational requirements may require changes to ensure that submarines remain at the forefront of capability and maintain the safety of operations, especially in the control room. These challenges include: the need to remain covert to avoid detection in disputed waters, as this could lead to the deterioration of relationships between countries, increased tensions in the region, or conflict (Bateman, 2011); increased operation in littoral (close to the shore/coast) waters (Roberts, Stanton and Fay, 2017b); and becoming an information-orientated navy, where information is readily used as a tool and a weapon to meet given objectives (Stanhope, 2012). To achieve these aims, submarine platforms of the future will employ new sensors with improved capabilities (Duryea, Lindstrom and Sayegh, 2008; Roberts, Stanton and Fay, 2015), process larger volumes of data (Stillion and Clark, 2015), and likely utilise a greater number of displays (Chalmers, Easter and Potter, 2000). However, an increase in the volume of data presented to an operator may not be matched by their capacity to interpret the data effectively (Woods, Patterson and Roth, 2002), and information that is not appropriately representative of the work environment can place additional processing requirements on users (McIlroy and Stanton, 2015b). Simultaneously, there is a drive to reduce crew sizes (Masakowski, 2000; Ly, Huf and Henley, 2007; Stanton and Roberts, 2018), which could further increase the amount of data each operator has to process, and increase operator workload (Carrigan, 2009; Henley, Schmitt and Huf, 2013). These factors mean that operator ways of working must be a key consideration for any changes implemented, ensuring that operators benefit from capability advances, and are not burdened with unwarranted cognitive requirements.

1.2 Command Team Experimental Testbed (ComTET)

The Command Team Experimental Test-Bed (ComTET) project aimed to understand current ways of working in submarine control rooms and provide evidence-based recommendations for future platforms that could be exploited to meet these future challenges (Roberts, Stanton and Fay, 2015). This would ensure that they remain at the vanguard of capability to enact the objectives of the RN. The project was designed as a lower fidelity facility than pre-existing facilities, such as trainers, enabling low-cost, agile, evaluation of new ways of working. A key underlying principle was the focus on recommendations over solutions; the aim was not to provide fully-fledged solutions to identified issues, but rather to furnish recommendations for further and more concrete (i.e., less abstract, and closer to real-world) evaluation in higher fidelity contexts.

Recommendations were passed to the funding body (Defence Science Technology Laboratory; Dstl), which has subsequently evaluated them for suitability of inclusion within the fleet. This addressed any potential issues arising from the difference between the simulator and the real-world environment as potential issues can be remedied, or the idea could be assessed as unsuitable. The ComTET project strived for appropriate fidelity to ensure that the work was as transferable as possible, and a representative submarine control simulation facility was developed at the University of Southampton (UoS; Roberts, Stanton and Fay, 2015). Using this facility, data from novice (non-submariners, trained for studies) and expert (Royal Navy submariners) teams was recorded and analysed to propose new and updated ways of working to stakeholders. Various avenues were explored, including different layouts (Stanton and Roberts, 2019), Submarine Tactical Operating Procedures (STOPS), and new Human-Machine Interfaces (HMIs; Fay, Stanton and Roberts, 2017; Fay, Roberts and Stanton, 2019).

The ComTET simulator used a version of Dangerous Waters as the simulation software. Dangerous Waters is a naval simulator that is sold as a game. It was customised by Sonalysts to be more aligned with the operation of Royal Navy submarines, and this version was used in this thesis. The simulator was split into two main areas, the control room and the experimentation room (see Section 4.1.1.2 and Figure 14 for a full description). The control room was a representative layout of a Trafalgar Class (T Class) submarine. However, exact measurements were not used to ensure the work remained unclassified. This was important as it allowed for a greater number of participants in studies, as novices, who were more numerous and readily available, could participate in an unclassified environment. The use of novices to understand relative differences in performance has been demonstrated to be appropriate (Walker *et al.*, 2010c; Stanton and Roberts, 2019). A variety of recording mechanisms (webcams, camcorders, microphones) were placed throughout the facility to enable high-quality data capturing for subsequent analysis. The simulator was remotely controlled from the experimentation room by researchers. This maintained immersion and ensured data integrity as much as possible.

The work in this thesis was conducted by the author (and acknowledged sources of help) within the bounds of the ComTET project. While the project and thesis are interlinked due to shared funding and resources, they are distinct. The ComTET project focused on ways of working for command teams, conducting several studies to identify and act upon opportunities for improvement. This thesis concentrated on the development of new submarine control room HMIs to understand if new designs could be exploited to maintain their effectiveness for future requirements. This contrasts with other ComTET work, which largely focused on non-HMI changes. New designs could contribute to achieving the goals of future submarine control rooms (more sensors, more data, and reduced crew sizes) without overloading operators, potentially improving control room effectiveness and safety. Unless explicitly stated otherwise, work in this thesis was performed for the ComTET project using its resources.

This research uses a sociotechnical systems approach, from its context in the wider ComTET (see Section 2.3.4), to appreciate the interactions between highly trained operators interacting with advanced technological systems in the control room. The term sociotechnical system was originally coined to describe systems incorporating complex interactions between humans and the environmental aspect of the work system (Baxter and Sommerville, 2010), citing Emery *et al.* (1960). A sociotechnical system can be further defined as the interaction of multiple social agents utilising technology for the completion of purposeful goal-directed behaviours (Walker *et al.*, 2008). A submarine control room is an excellent example of a sociotechnical system, with multiple operators interacting as a team, utilising a variety of sensors to generate knowledge of the environment to safely complete mission objectives (Ly, Huf and Henley, 2007; Stanton, 2014; Stanton and Bessell, 2014). The social aspects of the system are formed by the command team and its structure, while the technical systems are comprised of sensor, information, and control technology. As each subsystem contributes to overall system goals, a poorly performing subsystem could reduce the entire control room's effectiveness (Meshkati, 1991). This includes HMIs, which can be critical to success as they facilitate interaction between the human and technological aspects of the system (Walker *et al.*, 2010b).

1.3 Sonar and Target Motion Analysis Interface Shortcomings

This thesis concentrates on Sonar and Target Motion Analysis (TMA) due to their prevalence in a submarine control room, and their direct impact on ensuring safety. A detailed exploration of their work is presented in Chapter 4. However, brief explanations of Sonar and TMA are provided below.

Sonar plays a vital role in submarine operations, allowing a command team to understand ownship's (an alias for the boat a submariner is on) environment and proximity to other vessels without counter detection. This is achieved by detecting passive ambient sounds and performing analysis to understand what is emitting the sound and the behaviour of the emitter. Biological, natural, or mechanical movement with a body of water can create noise, such as the call of a whale, waves crashing, or a propeller rotating. This noise propagates through surrounding water and can be detected by a submarine's Sonar arrays. A Sonar array, or sensor, is a collection of hydrophones (underwater microphones) each receiving aural data from a specific direction (or a known direction from beamforming). This direction is called a bearing. As a hydrophone's direction is known (Solomon and Knight, 2002), the bearing a sound was heard from will also be known. The Sonar sensors (arrays) a submarine has installed, and their detection parameters differ, although they will generally have a bow (or spherical) array, a flank (or fin/sail) array, and a towed array. The bow array sits at the bow (front) of the submarine, the flank amidships (middle), and the towed deployed from the stern (back) from a wire spool (Baggeroer, 2005). The detected sound is shown to an operator for analysis, whereby they process signals of interest using a Sonar HMI to understand them.

TMA contributes to the creation and maintenance of the tactical picture, processing information to build a clear picture of the submarine's surroundings. This is vital to avoid collisions (Danczyk *et al.*, 2015), which can have catastrophic consequences (National Transportation Safety Board, 2001; Bateman, 2011; Marine Accident Investigation Branch, 2016). The data intensity arises as contacts must be sufficiently managed, ensuring each one has a valid and accurate solution using various available sources of data. When a sensor, such as Sonar or optronics, detects a vessel, it will send the bearing at which the detection occurred to TMA. This information is a '*cut*', which manifests itself as a straight line on a TMA map. Cuts are repeatedly sent by sensors across time, allowing operators to understand where a vessel has been, and create a solution for it (Ince *et al.*, 2009; Genç, 2010). A solution consists of bearing, course, speed, and range (Coll, 1994; Genç, 2010), allowing the command team to understand both a vessel's location and behaviour. This information allows the command team to navigate safely and track contacts of interest (Murphy, 2000c; Mack, 2003). A contact is an entity, such as a fishing vessel or carrier, that a command team has detected, requiring processing (Wang, 2016) to understand its location, disposition, and behaviour (Beevis, Vicente and Dinadis, 1998; Dominguez *et al.*, 2006; Carrigan, 2009).

The importance of Sonar and TMA is evidenced by the communications related to each forming a significant proportion of command team activities (Stanton and Roberts, 2018). As integral components of the submarine control room, the interfaces for these roles have been iteratively improved over time to meet the requirements of modern submarine operation, such as increased frequency of operating in littoral waters (Schank *et al.*, 2011). While successful in providing adequate functionality, it has been argued (Hall, 2012) that a redesign of these interfaces, instead of evolution from previous interface designs, could offer improved usability by capitalising on contemporary design paradigms.

This is because the ability of HMI subsystems to affect entire control rooms' operation has manifested itself several times; in several submarine accidents, HMIs or associated processes have been identified as a significant causal factor. For example, the National Transportation Safety Board (NTSB; 2001) investigation into the USS Greeneville incident highlighted the lack of a working Analogue-Video Signal Display Unit (AVDSU) as a critical failure which resulted in the submarine colliding with a surface vessel (Roberts and Tadmor, 2002). The purpose of the AVDSU unit was to provide repeated sonar (Sound Navigation and Ranging; Bj, 2011) sensor data to operators in the control room, and its failure reduced the availability of Sonar data. This contributed to the USS Greeneville being unaware of the fishing vessel Ehime Maru, causing catastrophic damage as it surfaced underneath.

Similarly, the Marine Accident Investigation Branch (MAIB; 2016) identified an incorrect classification of the fishing vessel 'Karen' as a causal factor in an incident involving a Royal Navy submarine. This wrong classification meant that the submarine operated close enough to Karen to cause damage to her trawling nets and associated equipment on board. Sonar operators had inappropriately classified Karen as a merchant vessel. Subsequently, it was assessed to be further away than it was by Target Motion Analysis operators maintaining the tactical picture. This led to a loss of separation between the two vessels, during which time the submarine briefly interacted with Karen's fishing nets. The accident report stated that whilst all systems were functioning properly, the command team (including Sonar) were cognitively overloaded and did not have sufficient time to assess contacts. Consequently, a potential opportunity to reassess Karen's distance, avoiding the incident, was missed (Marine Accident Investigation Branch, 2016).

These incidents demonstrate that HMIs are critical components of the control room system and their impact on safe operations and mission outcomes is not just theoretical. Despite decades of evolution to successfully meet the demands of today's maritime environment, they highlight ongoing challenges and the need to ensure these are continuously met to maintain effective control room performance. Specifically, these incidents highlighted potential shortfalls in Sonar and Target Motion Analysis (TMA) HMIs, demonstrating that control room capacity could be maximised via the development of HMIs that reduce operator workload when gaining an understanding of and operating in complex environments.

1.4 Ecological Interface Design as a Step Change

It is hypothesised that Ecological Interface Design (EID) is a suitable design paradigm to make such a step change. As will be comprehensively explained in Chapter 2, EID is a theoretical framework for designing Human-Machine Interfaces (HMIs) of a complex nature (Vicente and Rasmussen, 1992) that aims to make the affordances and constraints of a system and its environment apparent to operators (Van Dam, Mulder and van Paassen, 2008). In doing so, operators' innate perceptual and cognitive capabilities can be capitalised on, reducing cognitive demand, as well as workload. A simple example of this would be using colour to represent sea depth or using shapes to represent friend or foe status. These representations would be immediately more perceptible than representing that data in other formats, such as tables, or purely in text. This is not unique to EID, and HMI designers naturally seek to make their interfaces as accessible as possible for their intended users. However, what is unique to EID, setting it apart from other design methodologies, is its explicit focus on addressing the innate complexities of modern sociotechnical systems (Rasmussen and Vicente, 1989; Vicente and Rasmussen, 1992).

This is achieved using Cognitive Work Analysis (CWA; Rasmussen, Pejtersen and Goodstein, 1994; Vicente, 1999b), a framework designed for the analysis of complex sociotechnical systems, with an emphasis on how work could be conducted (formative modelling; Naikar, 2013; Stanton *et al.*, 2013; Stanton *et al.*, 2017a). CWA differs from approaches that analyse what a system currently does (descriptive modelling) or what it should do (normative modelling; Stanton and McIlroy, 2012). It does so by creating an understanding of constraints that define and bound a system, using different stages to represent different aspects of a system. The outputs from these stages describe how work could be conducted in a task-agnostic fashion. They can be used to construct an EID that enables users to determine and structure their work within these bounds. A focus on constraints, instead of tasks/processes, ensures that design solutions are not bound to specific courses of action, and mechanisms to respond to novel and/or unexpected situations are appropriately provisioned.

The ever-increasing complexities of modern sociotechnical systems, coupled with the unpredictability of all operational eventualities, can vastly diminish the practical ability of other approaches. This does not devalue these approaches, which have been demonstrated to yield their own benefits. However, EID distinguishes itself by recognising that not all operational eventualities can be planned for, and therefore operators should be facilitated in performing their tasks by situating them in an action state-space and providing them with autonomy to achieve goals (Vicente and Rasmussen, 1992; Borst, Flach and Ellerbroek, 2015). In doing so, operators gain an understanding of how the system works and can utilise this knowledge to reach their goals. Related to this is the provision of problem-solving information during unanticipated situations; if operators understand how the system works, and have operating autonomy, they can address the situation without being pre-emptively bound by predetermined sets of information and action.

1.5 Software Processes for Implementing Ecological Interface Designs

The theoretical principles of CWA and EID and their application are well established. Several seminal works address the design gap between them to create an interface design (Burns and Hajdukiewicz, 2004; Naikar, Hopcroft and Moylan, 2005; Read *et al.*, 2018). However, there is limited literature that addresses implementing the interfaces as software artefacts. There is a distinction between the frontend design and backend model, and they must both be designed for software to work as intended. Therefore, there is an opportunity to add to the literature linking EID to software engineering. This is especially pertinent as modern software processes are cognisant of the complexities of modern work domains, with methods such as Agile (continuous design and implementation) becoming preferred over older methods such as the Waterfall Model (implementation from a design that was completed at the start). Additionally, software is designed to be as modular as possible, creating flexibility that could be exploited when iterating through multiple EID ideas; the underlying code representing a work domain could stay the same, with only the interface and interaction changing. However, for such ideas to be realised, the processes involved must be well-defined. This thesis presents a method for moving between an EID and a software product for testing, which was used to create the evaluated interface.

1.6 Looking Forward

The timing for this research is opportune, owing to planned future submarines, and the implementation of vastly more configurable software systems onboard compatible boats. These factors create a window of opportunity that this research intends to leverage to exact tangible real-world benefits. Owing to submarine building lead times, there are opportunities to influence the design of future control rooms to best accommodate ComTET recommendations, including HMI recommendations made as part of this thesis. This is made achievable, in part, through contemporary software systems that facilitate modern software capabilities (see Chapter 3). This leverage is vastly increased by buy-in from various organisations, including the RN itself, providing an avenue of exploitation for recommendations. Whilst the control room is a relatively small component of an entire submarine, and the HMIs a sub-component of this, they are nonetheless integral components whose improvements will benefit the entire submarine.

1.7 Thesis Overview

1.7.1 Aims and Objectives

This thesis aimed to demonstrate whether EID is suitable for future submarine control room Sonar and TMA HMIs. RN submarines are vastly capable, although this does not mean they cannot be improved upon (Stanton, 2014). Future requirements may require control rooms to depart from contemporary operations to ensure continued efficacy. This includes the HMIs, which serve as a vital interface between highly trained operators and sophisticated technology; should this interface be compromised, adverse situations can occur. This research concentrates on the hypothesis that EID is a suitable design paradigm for ensuring control HMIs continue to be suitable for purpose and evaluation of the designs created to assess this. To achieve the overall aim, this thesis has the following objectives:

1.7.1.1 Objective 1: Creating a detailed understanding of Sonar and TMA operation

An understanding of Sonar and TMA was required to ensure feature-completeness of the new designs and understand design directions. This was achieved in two ways: a comprehensive literature review to understand control room operation (descriptive analysis), and the completion of Sonar and TMA CWAs (formative analysis). By reviewing the literature, a base understanding of how the submarine control room was established, creating a context for the operation of Sonar and TMA, as well as their operation. This enabled more in-depth CWAs to be conducted, as a base level of understanding of system operation was present. The CWA outputs were used to create an EID and address a gap in the literature where there is a paucity of publicly available information on how submarine Sonar and TMA operate. By filling this gap, other system designers could use the outputs to facilitate their understanding of Sonar and TMA operation or adapt the analyses to their specific systems. Finally, the CWAs were also used to verify the compatibility of the simulator used for the experiments with a high-fidelity simulator, establishing that the simulator used was appropriate.

1.7.1.2 Objective 2: Creating a documented analysis and design process, oriented towards software engineering

There was a requirement to document the analysis process followed to create the EID HMIs. While there is literature that explores this, it largely does not do so with an appreciation of software engineering, an eventual requirement if EID is to move from designs and prototypes to full-blown implementations. Given that a core aim of EID research is to propose systems that will eventually be implemented, it was pertinent to explore how CWA and EID could be more robustly linked to software engineering practices to ensure optimal implementation. This process is vital to any recommendations made, as it creates reproducibility for future iterations or application to entirely new systems. This also contributes to wider knowledge, as there is currently a dearth of literature addressing how this design gap is bridged with software engineering in mind.

1.7.1.3 **Objective 3: Assessment of a novel Sonar and Target Motion Analysis HMI**

The final objective was to assess how novel Sonar and TMA EID HMIs impacted operator performance regarding usability, task accuracy, and workload. This took the form of a Human in the Loop experiment in the ComTET simulator. The results of the experiment were generally positive, supporting the assertion that EID would be suitable for submarine control rooms. In a wider context, this adds to the ever-growing body of EID literature detailing its application for a variety of purposes across several domains.

1.7.2 Structure

This thesis comprises the following chapters:

Chapter 1 Introduction

This chapter contextualised the research, highlighting how future requirements for submarine control rooms are necessitating changes to keep pace. The ComTET project’s main aim was to conduct statistically robust, repeatable, experiments to generate recommendations for the RN. A line of reasoning is put forth describing the impetus for investigating a step change in HMIs, which is the focus of this thesis. While the research presented in this thesis was conducted within the ComTET project, all main contributions are that of the author, unless otherwise specified.

Chapter 2 Theoretical Underpinnings

CWA and EID are based on theory which must be comprehended for their appropriate application. This chapter introduces these theoretical underpinnings, enabling comprehension of the principles followed in subsequent chapters. It focuses on presenting concepts integral to EID, the Abstraction Hierarchy (AH; an output from CWA) and the Skills Rules Knowledge Taxonomy.

Chapter 3 A Taxonomy of Constraints

The chapter builds on Chapter 2, presenting a systematic review of EID literature to create a taxonomy of constraints for use in CWA and EID. Literature was systematically collated through searches on multiple search engines and all mentions of constraints were coded. This coding activity informed a bottom-up synthesis of twenty categories present in the literature, which are presented with their top three constraints in each. Additionally, Social Network Analysis was applied to reveal the connections between the twenty categories and transition the results from being a checklist to a connection-driven taxonomy that could be used for future systems analyses by enabling the discovery of related connections. Consequently, this chapter found that it was possible to categorise constraints in the EID literature and link these as part of a taxonomy that can be used for systems analysis.

Chapter 4 How submarine control rooms operate, and the Case for Ecological Interface Design

This chapter presents a comprehensive overview of submarine control room operation, including both the operators and technology that form the sociotechnical system. Building on this understanding, a case is presented for using EID to address HMI shortcomings in submarine control rooms, based on the theory presented in Chapter 2. The case is comprised of the hypothesised benefits of EID itself, in addition to recognising the synergy between the method and command team operation. The submarine command team works to understand their operational environment, what limits it imposes, and how it can be leveraged; parallels can be drawn between these aims and the core goals of EID. The chapter provides context to the submarine control room, command team, Sonar, and TMA operation to illustrate these core points, imbuing an understanding for readers, and making the research contained within this thesis more accessible.

Chapter 5 Cognitive Work Analysis of Sonar and Chapter 6 Cognitive Work Analysis of Target Motion Analysis

These chapters present the CWA outputs conducted for the ComTET versions of Sonar and Target Motion Analysis, complimenting the research from Chapter 4. Two stages of Cognitive Work Analysis were performed, Work Domain Analysis and Worker Competencies Analysis. The first part of this chapter presents the method used to conduct the CWA for both, and how the traditional approach to WDA was modified to incorporate items in the interfaces at the Physical Objects level, a change arising from observing that the interfaces made use of representations of real-world Physical Objects that would traditionally be expected. The approach is proposed as a mechanism to manage the initial analyses of complex work domains, creating a sufficient, yet not inappropriate, scope and level of abstraction from advanced underpinning concepts (e.g., sound propagation, sensors, and algorithms).

The operation of both roles is detailed at each level of the Abstraction Hierarchy, before concentrating on a key component of each interface where improvements were likely to be made, known as a leverage point (Read *et al.*, 2018). For Sonar, this was the waterfalls, which display sonar data in a rectangular area on-screen, and are a pervasive representation of data in Sonar interfaces. They are named as such because new data appears at the top and pushes other data down, giving a waterfall-like visual effect (Asplin and Christensson, 1988; Matthews *et al.*, 2006), see Figure 15d and Figure 15e in Section 4.1.1.4. Which is a prevalent mechanism for displaying and interacting with aural data, owing to the amount of data that can be presented in a compact format. While a common representation of Sonar data, it is pertinent to evaluate whether data could be represented in a different, better, format. The Local Operations Plot was selected for TMA, as this is one of the key mechanisms for entering solutions (data about where another entity is and its behaviour). The skills required to operate each station are presented as the output of a Worker Competencies Analysis. Combined, these outputs enabled the identification of design directions for both roles, feeding into the design process in Chapter 8. For Sonar, the recommendations focused on reducing operator workload and making information more readily available. The TMA recommendations were aimed at suitably introducing automation capabilities to assist operators with managing an increase in data, without degrading their ability to perform their role by themselves should the automation require their intervention of fail.

Chapter 7 Validation Against the Talisman Trainer

This chapter addresses the need to ensure that the simulator is fit for purpose. This is important to ensure that experimental results would accurately translate to a real-world context. A validation activity of the simulator versions of Sonar and Target Motion Analysis against their high-fidelity simulator equivalents is presented, showing the similarities and differences between the two. This established that the simulated versions are suitable for purpose, exhibiting a high degree of functional fidelity with a medium degree of physical fidelity.

Chapter 8 Creation of Graphically Integrated Sonar and Target Motion Analysis

This chapter presents a method for creating a software EID from the CWA outputs. It is practically applied to the outputs from previous chapters, detailing how they were translated into initial designs, which were then developed in a simulation engine. It explains how multiple aspects of the design were created. Following on from the creation description, an overview of the resultant interface, Graphically Integrated Sonar and TMA (GIST) is provided. This serves to familiarize the reader with the interface, but also to illustrate how the design directions, and underlying challenges, have been addressed. This chapter found that it was possible to link EID to software engineering processes, with a proof-of-concept application in the form of GIST.

Chapter 9 Evaluation of Graphically Integrated Sonar and Target Motion Analysis

GIST required testing against a contemporary baseline to establish how it affected operator performance. 45 participants took part in a study as individuals (as opposed to a team study), aimed at assessing how their objective performance and perceived subjective usability differed between interfaces. Low- and high- difficulty scenarios were used to vary participant workload. It was discovered that there were statistically significant differences in subjective workload, measured using the System Usability Scale and NASA Task Load Index, in favour of GIST. Significant improvements were also found in the objective performance of operators in favour of GIST, with tasks being completed with a higher degree of accuracy. This was especially important for the solution accuracy task, where participants placed markers where they believed entities were, as it can directly affect submarine safety. However, usability scores showed that there was room for improvement in future iterations. Overall, this chapter proved that implementing EID for Sonar and TMA can be beneficial over contemporary designs.

Chapter 10 Comparing Graphically Integrated Sonar and Target Motion Analysis to a User-Centred Design

The experiment in Chapter 9 also included a Mashup display that was created using a User-Centred Design approach, to evaluate the performance of GIST as an EID against. The experiments were combined due to restrictions of the Coronavirus-19 pandemic. It was discovered that the Mashup display was rated better for usability, although this was not significant. Task performance was equally split between both of the interfaces, with each being rated as significantly better than the other for one task. This provided insights into how the design of each one could have affected this, specifically regarding their focus on either the operator or the operator’s environment. Finally, the Mashup was shown to have significantly improved subjective workload, which is discussed to posit why this was, as EID has been shown to have better subjective workload than User-Centred Design in other literature.

Chapter 11 Conclusions

This final chapter details how the objectives detailed in Section 1.7.1 have been met by providing a summary of work relevant to each and describes how the contributions to literature were made. It then presents key points on the evaluation of this research project as a whole, and ideas for future work, before the concluding remarks.

Chapter 2 **Theoretical Underpinnings**

2.1 Introduction

Chapter 1 introduced the case for Ecological Interface Design (EID) being applied to Sonar and Target Motion Analysis (TMA) as the central thread of this thesis. This chapter is to familiarize the reader with the required theory for EID, ahead of future chapters which will build on this foundation. It starts with an overview of the Skills, Rules, and Knowledge (SRK) Taxonomy and Cognitive Work Analysis (CWA) as underlying theory, before introducing the theory of Ecological Interface Design (EID) and how it utilises the SRK Taxonomy and CWA. The case for using EID over User Centred Design, which differs from EID as it focuses on user tasks, is also presented.

The impetus for exploring EID arises from the challenges of modern sociotechnical systems, which are rapidly becoming more complex as the capability of technology continues to evolve. Across multiple domains, vast amounts of data can be collected and processed in real-time, a feat that would simply not be possible for humans alone. However, humans are still a vital aspect of these systems, as they can provide adaptability and creativity, enhancing the system as a whole (Borst, Flach and Ellerbroek, 2015). Whilst technology can automate routine aspects, humans are still required to act as knowledge workers, or to respond to non-routine and unanticipated situations (Vicente, 2002). This dyad is largely successful, with complex sociotechnical systems operating as expected most of the time, such as rail operating centres, air traffic control, nuclear power plant control rooms, or submarine control rooms. However, as their capability grows, so does their effect on society. For example, power plants can generate power for larger areas, and transport control rooms can control larger areas of traffic. Due to this growing impact on society, it is imperative that these systems are designed in the best possible way to ensure continuous successful operation. It is inevitable that mistakes or unanticipated situations will occur during operation, and the system should support operators in their successful handling. This is particularly vital, as whilst most can be recovered from successfully, such as an operator misspeaking or entering incorrect data, not recovering could lead to catastrophic outcomes.

A vital aspect of effective system design to address this issue is creating appropriate Human-Machine Interfaces (HMIs) for operators to use (Stanton *et al.*, 2017a). To address this in complex sociotechnical systems, EID was proposed by Vicente and Rasmussen (1992). It aims to optimise the design of HMIs to reduce operator workload and support operator cognition at multiple, appropriate, levels, by making the constraints and affordances of a work domain apparent (Van Dam, Mulder and van Paassen, 2008). It has been applied in multiple domains, such as aviation, medicine, and power generation (Vicente, 2002; McIlroy and Stanton, 2015b). The benefits of EID include improved performance compared to traditional (non-EID) systems (Vicente, 2002), increased work domain transparency (Van Dam, Mulder and van Paassen, 2008), reduced workload (Nielsen, Goodrich and Ricks, 2007), and reduced memory requirements (Lau and Jamieson, 2006).

2.2 Skills, Rules, Knowledge (SRK) Taxonomy

To understand how these benefits are achieved, it is important to first understand the levels of control that users could employ when interacting with a system and the demands involved. The Skills, Rules, and Knowledge (SRK) Taxonomy (Rasmussen, 1983; Rasmussen and Vicente, 1989) classifies behaviour into three basic types and describes cognitive behaviour for the completion of tasks (Drivalou and Marmaras, 2009). These levels interact, as tasks rarely belong to only one class of behaviour (Stanton *et al.*, 2017a, p. 40). For example, novices may use Knowledge-Based Behaviour (see Section 2.2.3) for novel tasks, gradually shifting to Rule-Based Behaviour (see Section 2.2.1). Multiple levels of cognitive control can be active at once, and thus there are significant interactions between levels (Rasmussen, 1983; Naikar, 2013). Rasmussen (1983) further notes that levels are not alternatives to each other; rather, they are processing responses to different information categories present in the environment. Additionally, it describes how information at each level is perceived, in the form of signals, signs, and symbols (Rasmussen, 1983; Vicente and Rasmussen, 1992). It is not designed to be a single quantitative model of human performance, but rather an overall qualitative model to match categories of performance to situation types (Rasmussen, 1983). The three taxa, descriptions of associated behaviour, and associated information perception method are explained in the following sections (Rasmussen, 1983; Kilgore and St-Cyr, 2006; Naikar, 2013). The SRK Taxonomy is represented in Figure 1, showing a simplified view of each taxon and their interrelation. It illustrates the explanation above of how the perception of signals, signs, and symbols trigger a response at the corresponding behavioural level.

2.2.1 Skill-Based Behaviour (SBB)

Automated and integrated actions involving little to no conscious attention, coupled with the environment as a perception-action loop. This level of control is indicative of an expert performer utilising tacit knowledge. Information for this level of control is perceived as time-space signals, which are continuous and quantitative representations of time-space patterns.

2.2.2 Rule-Based Behaviour (RBB)

If-then mappings between familiar environmental perceptual cues and courses of action. Behavioural rules can be sourced from procedures, experience, instruction, or prior problem-solving (Vicente, 1999b). This level of control is indicative of an intermediate performer, capable of utilising rules to determine a course of action. Information for this level of control is perceived as signs, which are arbitrary, but familiar, cues to initiate RBB.

2.2.3 Knowledge-Based Behaviour (KBB)

Full conscious control is dedicated to completing a task whereby no prior experience exists. This process is goal-controlled, with the goal formulated based on an analysis of the environment and the person's aims. Courses of action are synthesised and evaluated against the goal to determine a course of action, making KBB slow and effortful. This level of control is indicative of a novice performer or performers faced with unfamiliar situations, requiring internalised models of the system, mental models, to problem solve. There are many different definitions of mental models (Revell, 2015). However, the definition adopted in this work is that of Johnson-Laird (1983; 1889), who defined mental models as dynamic representations or simulations of the world (Stanton and Young, 2000). Stanton and Young (2000) note that one's picture of their working environment has a high degree of overlap with mental models. Brewer (1987) differentiates mental models from schema by defining schemas as "generic mental structures underlying knowledge and skill, whilst mental models are inferred representations of a specific state of affairs" (Stanton and Young, 2000; Revell, 2015). In the context of the SRK Taxonomy, this is interpreted to mean that the operator must construct, maintain, and utilise their mental model of the current environment to support their problem solving. Information for this level of control is perceived as symbols, which are formal structures, mental models, representative of the environment's functional properties. These can be utilised to determine an appropriate course of action.

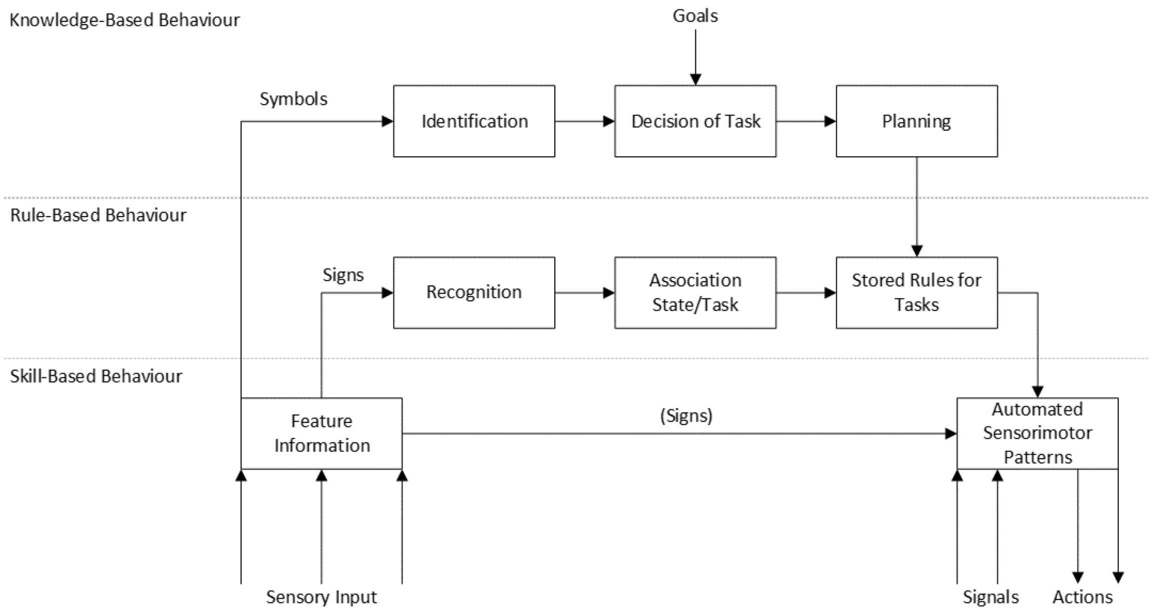


Figure 1 – The Skills, Rules, and Knowledge Taxonomy, adapted from Rasmussen (1983) and Waterson, Le Coze and Andersen (2017).

2.3 Cognitive Work Analysis (CWA)

CWA (Rasmussen, Pejtersen and Goodstein, 1994; Vicente, 1999b) is a framework designed for the analysis of complex sociotechnical systems. It focuses on how work could be conducted, known as formative modelling (Naikar, 2013; Stanton *et al.*, 2013; Stanton *et al.*, 2017a). By mapping constraints to define how the system could perform, it differs from approaches that analyse what a system currently does (descriptive modelling) or what it should do (normative modelling; Stanton and McIlroy, 2012). It has been applied across a variety of domains for a multitude of purposes, such as nuclear power, or medicine, (McIlroy and Stanton, 2015b; Stanton *et al.*, 2017a), including specifically in the maritime command and control domain (Bisantz *et al.*, 2001; Bisantz *et al.*, 2003; Burns, Bisantz and Roth, 2004; Burns, Bryant and Chalmers, 2005). There are five phases to CWA, each designed to elicit different domain constraints: Work Domain Analysis (WDA), Control Task Analysis (ConTA), Strategies Analysis (StrA), Social Organisation and Cooperation Analysis (SOCA), and Worker Competencies Analysis (WCA). Depending upon the purpose of the analysis the appropriate phrases are selected (McIlroy and Stanton, 2015b). For HMI design, including EID, WDA and WCA can be used (Burns and Hajdukiewicz, 2004; Jenkins *et al.*, 2009). Thus, these two stages were selected to use for the creation of the EID for this thesis and are described below.

2.3.1 Work Domain Analysis (WDA)

WDA (Rasmussen, 1985; Vicente, 1999b) assesses a system on multiple levels of abstraction to understand its constituent aspects and purposes for existing (Lintern, 2006). It describes constraints governing the work domain (Jenkins *et al.*, 2008a), which define reasons and resources for agent behaviour (Naikar, 2013). The levels of abstraction are described in Table 1.

Table 1 – Description of abstraction levels present in Work Domain Analysis (Naikar, 2006; Stanton *et al.*, 2017a)

Level of Abstraction	**Description**
Functional Purpose	A system's reason for existing, detailing its high-level aims or objectives. These exist for as long as the system exists.
Values & Priority Measures	Criteria for measuring if Functional Purposes are being addressed.
Purpose-Related Functions	Object-independent functionality necessary to achieve the Functional Purposes. These can be viewed as describing what Physical Objects and their Object-Related Processes are used for in a system (Miller and Vicente, 1998). Functions can affect Value & Priority Measures.
Object-Related Processes	The purpose-independent affordances and limitations of the Physical Objects to perform Purpose-Related Functions. They are immutable (cannot change) due to their dependence on Physical Objects; the properties of an object cannot change unless the object itself does.
Physical Objects	Objects within the system, which can either be corporal or ethereal in nature.

The five levels of abstraction are represented as an Abstraction Hierarchy (AH), a diagram composed of horizontal rows of textboxes (nodes) to represent each level, with connectors linking nodes on adjacent levels. Connectors are called means-end links, revealing a node's place within the system. A means-end link represents a '*Why, What, How*' triad. Any given node is the '*What*'. Following connections upwards reveals '*why*' the '*what*' exists, and following connections downwards reveals '*how*' the '*what*' is implemented (Stanton *et al.*, 2017a). This permits an understanding of why and how a given component affects the system, facilitating a better understanding of its operation. Figure 2 illustrates the concept of means-end links showing that the triad can move to any node on the AH to understand why it exists and how it achieves its aims.

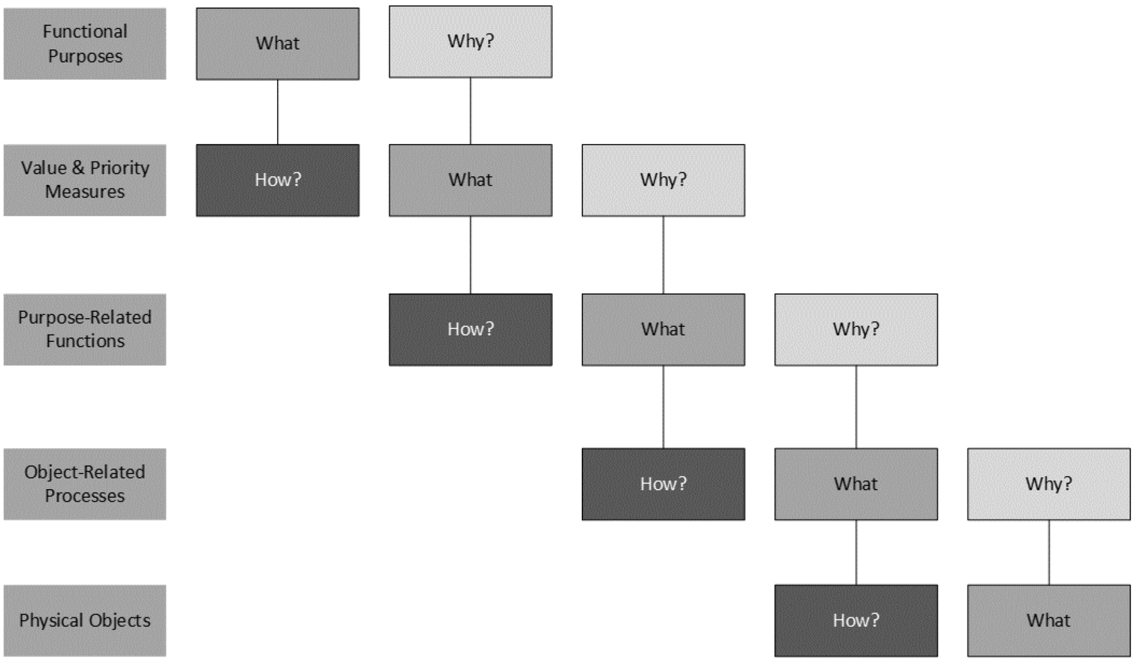


Figure 2 – A representation of means-end links

A sample AH for a kettle is shown in Figure 3, with each level labelled and the means-end links between levels illustrated. For each node, it is possible to determine why it exists and how it achieves this. A kettle's main purpose is to provide hot water, which is represented at the Functional Purpose level. One level below, the Value & Priority Measures are outlined. Each of these are criteria that would be considered when determining if the kettle is providing hot water: Is the temperature hot enough and has it been heated in a reasonable time? Moving to the Physical Objects level at the bottom, each Physical Object is outlined. Some scoping has been performed, so as not to unnecessarily include items. For example, the infrastructure to provide water is not detailed. The affordances of these items are detailed on the Object-Related Processes level and are related to each object. An example of these properties can be seen in the kettle unit not providing power. To change this, the kettle object would have to be changed to one that does not use a separate base stand. Finally, the functionality that can be carried out to contribute towards the Value & Priority Measures is displayed at the Purpose-Related Functions level. The kettle can bring water to a boil, but this is independent of the Physical Objects and their Object-Related Processes; boiling water could be attained on the stovetop, in a microwave, or from a boiling-water tap system. Through presenting each level of abstraction, the kettle's constituent aspects and purpose for existing can be understood. Furthermore, the constraints of the system can be understood; using means-ends links, it can be understood what effect the removal of a node would have on a system (Salmon, Carden and Stevens, 2018). For example, power is required to bring water to the boil, and that the base stand is used to do so. Thus, it is required for the kettle's operation and its removal would negatively affect the system.

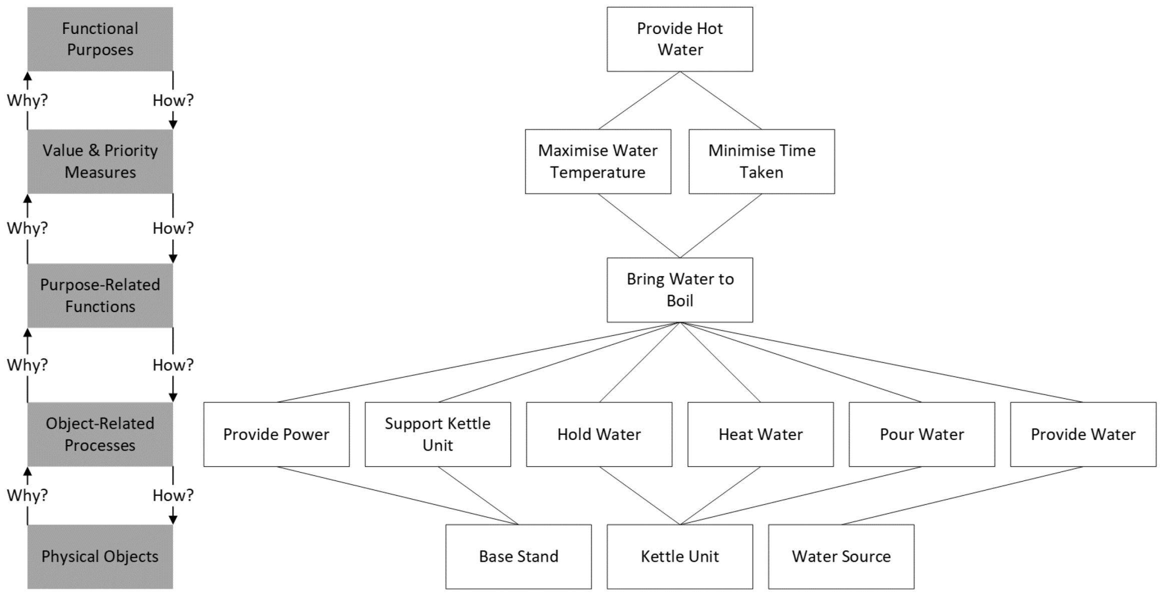


Figure 3 – A sample Abstraction Hierarchy for a kettle, showing levels and means-end links

2.3.2 Worker Competencies Analysis (WCA)

WCA categorises worker capabilities into different levels of cognition for the completion of available tasks within a system (Naikar and Elix, 2016). The modes of cognition are populated from the Skills, Rules, and Knowledge taxonomy. This can inform task distribution by facilitating an understanding of which tasks could be assigned to actors or could be automated. It is represented as a matrix of activities (rows) and cognition modes (columns), populated by how each activity would be completed using each cognition mode.

Traditionally, activities are populated from information processing steps in a CWA decision ladder. Decision ladders (Rasmussen, 1974b) represent the flow of information processing steps taken by operators to complete tasks (Stanton *et al.*, 2017a). However, it is possible to use Object-Related Processes from a WDA as well (McIlroy and Stanton, 2011; Stanton and Bessell, 2014). Stanton *et al.* (2017a) argue that this is more suitable when designing a system, as decision ladder construction requires knowledge of workflow, which would not be available before the system has been designed. Table 2 shows an example subset of a WCA for the kettle AH using the Object-Related Processes for the base stand Physical Object. It uses the headers and adapted associated definitions from the work of Stanton and Bessell (2014), who completed a WCA as part of a full CWA on a how a submarine returns to periscope depth. The SBB is typical of people with experience using a kettle and solving associated issues, and the RBB behaviour is steps that might be taken in response to certain situations whilst doing so. Both are likely to not require excessive cognitive effort. However, more cognitive effort might be expended for KBB to understand potential causal factors for the kettle not being powered as they must reason without the SBB experience.

Table 2 – An example WCA matrix for the affordances of the base stand in Figure 2, with headers and adapted associated definitions from Stanton and Bessell (2014)

	Skill-Based Behaviour	**Rule-Based Behaviour**	**Knowledge-Based Behaviour**
	Observed behaviour of experts	If-Then rules to identify if a task should be done	Prerequisite knowledge / capability for novices to complete tasks
Provide Power	Understand other causes that may cause the kettle to not turn on, such as a blown fuse or power cut.	If the kettle does not start, then check the power is on.	Placing kettle on the base stand. Flicking the on button.
Support Kettle Unit	Inspection of both units to determine how to seat kettle. Inspection of both units to determine why kettle is not seated correctly.	If the kettle is not sitting properly, then try adjusting the kettle's position, or reseating it.	Placing kettle on the base stand.

WCA can be used to inform interface design, by informing design that supports novice, intermediate, and expert users (McIlroy and Stanton, 2011). Furthermore, it can inform shifting KBB into SBB or RBB, reducing cognitive workload during interface use (Morineau *et al.*, 2009). Cognitive workload can be defined as the relationship between task demands placed on an operator, and the capacity of the operator to meet them (Michailovs *et al.*, 2022), citing Parasuraman, Sheridan and Wickens (2008). KBB can also indicate if a task could be automated (Stanton and Bessell, 2014). For example, automation could be designed by removing routine SBB and RBB responsibilities from an operator, or similar tasks could be consolidated for an operator to perform.

2.3.3 Justification for not using other stages

EID was initially conducted and prescribes (Burns and Hajdukiewicz, 2004; King, Read and Salmon, 2022) using only WDA (Rasmussen, 1985) and WCA (Rasmussen, 1983). It has since expanded to use all subsequent stages of CWA, although this is not prevalent within the literature, with only 29.31% of reviewed applications utilising a combination or subset of WDA, WCA, and the SRK only (McIlroy and Stanton, 2015b). They, along with others (Jenkins *et al.*, 2008a; Stanton *et al.*, 2017a; Simon *et al.*, 2022), advocate for the use of the remaining stages to analyse the system in more detail, which should lead to a more considered interface from using the additional information. A brief description of each remaining stage and the constraints that they identify is described in Table 3, adapted from (Vicente, 1999b; Jenkins *et al.*, 2008a; Stanton and Bessell, 2014; Rauffet *et al.*, 2015).

Table 3 – Description of remaining stages of CWA

Stage	Analyses	Representation
Control Task Analysis (ConTA) Consideration of recurring activities within a system, independent of the actor and strategy. Represents activity as intersections between functions and situations. Intersections are situational constraints, determining whether work is mandatory, possible, or impossible (McIlroy and Stanton, 2015b).	Decisional and Situational Constraints	Decision Ladder (DL) Contextual Activity Template (CAT)
Strategies Analysis (StrA) Details different methods for executing the same task (Stanton and Bessell, 2014).	Strategy Constraints to Achieve System Goals	Ahlstrom (2005) flow diagrams (McIlroy and Stanton, 2015b).
Social Organization & Cooperation Analysis (SOCA) Actor attribution for who can perform activities, and in which situations (McIlroy and Stanton, 2015b).	Functional Allocation Constraints	Actor colour coding for artefacts from other stages, such as the WDA (SOCA-WDA), CAT (SOCA-CAT), and DL (SOCA-DL)

The inclusion of each stage is not absolute and should be reflective of the constraints present within the system (Burns and Hajdukiewicz, 2004; McIlroy and Stanton, 2015b). These types of constraints are present within the submarine control room, and so were considered as part of the design process in this thesis. However, the stages of CWA above were not utilised owing to other analyses being conducted within ComTET that could provide the requisite information, either alone or combined with others. This is congruent with the literature, where supplemental methods for CWA are used (McIlroy and Stanton, 2015b) and encouraged (Burns and Hajdukiewicz, 2004). From a practical standpoint, it is well-known that CWA is a time-consuming and resource-intensive method (Rehak, Lamoureux and Bos, 2006; Stanton *et al.*, 2013; Stanton and Bessell, 2014; Read *et al.*, 2018), and analysis replication would have been detrimental to project completion. Thus, a decision was made to utilise the other methods to derive the information required in lieu of completing these stages.

The first is Hierarchical Task Analysis (HTA; Annett and Duncan, 1967), which is a method for describing a system in terms of its goals in a hierarchical manner (Stanton, 2006; Salmon *et al.*, 2010). The approach is more directed than CWA, although there is still variance in how it is conducted. Generally, however, it takes the appearance of a top-down tree of sub-goals originating from a singular top-level goal. Leaves in the tree are not a procedural list of operations, rather they are sub-goals (Annett and Stanton, 2000). They are structured by their contribution to a common super-ordinate goal (i.e., contribute to its completion), and are assigned plans that detail execution strategies.

HTA differs from CWA in that it focuses on system goals over the constraints present within a system (Salmon *et al.*, 2010). Constraints are defined by Vicente (1999b) as relationships between, or limits on, behaviour. A full treatment of this definition is provided in Section 3.2.3. Goals are defined as dynamic state targets to be met by actors at a given time-point (Vicente, 1999b). This difference is explored in detail by Salmon *et al.* (2010), who argued that while they have differing approaches, their division is not clear-cut. The DL, an artefact of ConTA, presents activity completion strategies, with some literature focusing on goals. While there are clear differences in theoretical underpinning, it could be argued that this permeation between the methods could be exploited for utilising HTA output through a CWA theoretical lens. For example, it could be possible to use HTA as source material for ConTA and/or StrA. While there is a dearth of literature on such a possibility, McIlroy and Stanton (2015b) indicated that the methods can be combined for use in EID by highlighting such examples in the literature. For example, Segall *et al.* (2013) used a HTA to identify plans and strategies in their study examining anaesthesia crisis management. Upton and Doherty (2008) also used HTA to supplement WDA, taking additional steps to decouple task descriptions from the existing system they were redesigning. Therefore, while not widespread in the literature, there is evidence to support using HTA as a mechanism for understanding decisions and strategies is viable.

The second utilised was Event Analysis of Systemic Teamwork (EAST; Walker *et al.*, 2006; Stanton, Baber and Harris, 2008; Walker *et al.*, 2010a), which is a multi-network method for distributed cognition representation in sociotechnical systems (Young *et al.*, 2007; Walker *et al.*, 2010a; Stanton, 2014). The three networks are task, information, and Social [Network Analysis] (SNA; Driskell and Mullen, 2004). It incorporates normative methods such as HTA, but differs in that it specifies what happens (over what 'should' happen) (Walker *et al.*, 2010a); this is achieved by observations, such as watching scenarios being completed in trainers. Rather than representing tasks as part of a linear process, they are instead represented as a non-sequential representation of the strategy being executed (Stanton, 2014). This bears a similarity to CWA's StrA in that it would be possible to visualise the different strategies one could employ to reach a designated state. This builds upon HTA's similarity by also providing an awareness of the affective constraints, such as close-quarters contacts in a return to periscope depth scenario, as described by Stanton (2014). The social network is constructed by analysing information flow (namely communications) between actors (both social and technological) to construct a matrix of their communications (actors as headers, with intersections populated by their directed communication counts). This can then be used to visually represent communications between each different actor in graph format. The information network is constructed from the transcripts used for the social network analysis. Concepts are identified in the communications, and these are linked with other concepts through proximity, such as appearing in the same sentence together. Stanton (2014) argues that this accounts for distributed cognition as the network represents information relationships to both agents and tasks, accounting for holistic system phenomena.

Individually, these networks reveal substantial information about a network. However, it is their combination where the value of EAST is demonstrated. This is demonstrated in the submarine domain by Stanton (2014) who combined various models in their description of a submarine returning to periscope depth safely. Of specific interest is the assignment of actors from the social network to the nodes of the information and task networks. If these networks can be related back to CWA, then the application of agents could be utilised as SOCA. By examining which agents can complete the tasks identified, it becomes possible to understand functional allocation constraints, which are required for CWA. This proposition is supported by the work of Baber, Stanton and Houghton (2017), citing Pfautz and Pfautz (2008), who examine how SNA could be conducted using a SOCA-CAT.

To summarise, while the artefacts from the use of HTA and EAST as part of ComTET and their exact purposes are outside the scope of this thesis, their outcomes were used as supplementary analyses to CWA. The stages used, WDA and WCA, are theoretically required for EID (Rasmussen, 1983;1985; Burns and Hajdukiewicz, 2004; King, Read and Salmon, 2022), so there is no detriment in adherence to the method. While there are differences between the other stages of CWA, and HTA/EAST used to replace them, the literature supports both their use as 'surrogate' artefacts that can be used. The differing theoretical approaches does create a limitation over using all stages of CWA, although limited resources made it a prudent option.

2.3.4 Accounting for Sociotechnical Systems Theory

Sociotechnical is the interrelatedness between the social and technical, as introduced at the end of Section 1.2 to describe the submarine control room. Its theory consists of two main principles (Walker *et al.*, 2008). The first is that the interactions are instrumental to success or failure. This applies to the system as a whole, as for poorly performing sub-systems, which can reduce holistic performance (Meshkati, 1991). Interactions in the system can either be linear cause-and-effect relations, or more complex non-linear relationships, which can be unpredictable and unexpected. Walker *et al.* (2008) argued that optimisation of one aspect of a sociotechnical system over another can introduce more of the latter relationships, which can be detrimental to performance. Consequently, they defined sociotechnical theory as concerned with methods of joint optimisation, designing for systems exhibiting open systems properties, allowing them to better handle environmental complexity, dynamics, new technology, and competition. von Bertalanffy (1950) defines open systems as those that are not self-contained, having a ‘permeable’ boundary that permits outside interaction and changing of components. Conversely, closed systems are sealed, with all activity taking place internally.

The submarine control room is a complex sociotechnical system (Stanton and Bessell, 2014). This thesis employs CWA and EID, both methods designed for such systems (Chalmers, Easter and Potter, 2000; Ly, Huf and Henley, 2007)/(Kilgore and Voshell, 2014). The argument for EID is made in Sections 2.4 (definition and use over User-Centred Design), and 4.2 (applicability to submarine control rooms), and the method acts as a central focal point for the research. However, this research was an open system and incorporated aspects from ComTET, where sociotechnical systems theory was used extensively. This means that while the theory was not a central tenet of this research, it was used to conceptualise, contextualise, and understand submarine control room operation.

The primary thrust of ComTET was human in the loop experimentation, using a new streamlined version of EAST (Roberts, Stanton and Fay, 2018; Stanton and Roberts, 2019), to examine the results from a sociotechnical systems theory standpoint. This was achieved by applying social network analysis metrics to each network to quantify macro effects of changes (Stanton and Roberts, 2020). Firstly, a baseline was established to examine current ways of working across a variety of scenarios (Roberts, Stanton and Fay, 2017b; Stanton, Roberts and Fay, 2017; Roberts, Stanton and Fay, 2018; Stanton and Roberts, 2018). This established an understanding of how tasks were completed in the control room, and core information concepts, including their passage around the control room, and temporal delays (Pope, Roberts and Stanton, 2019). The experiment revealed that there was a bottleneck in communication between supervisory operators for Sonar and Target Motion Analysis, which was an important bridge for effective control room operation.

A subsequent experiment aimed to address this bottleneck by applying a joint optimisation to the system by co-locating Sonar and Target Motion Analysis operators with the aim of improving their communication and access to each other’s information (both held in their mind, and on their screens; a technical agent) (Roberts *et al.*, 2019; Stanton and Roberts, 2019; Pope *et al.*, 2020; Stanton and Roberts, 2020). This study also examined the effect of removing one operator from each team. It was revealed that the bottleneck was addressed, and afforded greater productivity in terms of task completion, even in conditions with reduced operator counts (Roberts *et al.*, 2019). This led to a final study, where all operators were arranged in a novel, inwards-facing, circular configuration (Stanton *et al.*, 2020b). This revealed that more information was shared, with less communication, indicative of information communication efficiency. As with the previous study, there was an increase in task completion.

All these results, driven by sociotechnical systems theory, influenced this research. The baseline work was vital to understanding current issues with submarine control room design, and the component diagrams of the resultant EAST diagrams established the foundations of the interface designs in this thesis. The identified bottleneck reinforced the notion that Sonar and Target Motion Analysis were suitable roles to investigate for this research, as information was not optimally flowing. While a social issue, identified from the social networks, the operators use technical agents, see Section 4.1.2, to complete their work. Thus, the principle of joint optimisation was utilised to improve the interfaces along other measures, aiming to provide operators with updated technology to exploit the increased capability. Similarly, the principle of co-locating operators used in the second experiment lent credibility to the notion that the interfaces could be utilised by a single operator (i.e., *"Why are operators not provided with their own copy of this interface, instead of relying on their neighbours?"*). This then opened the avenue of exploration for merging the roles, as it was illogical to not enable individuals to take advantage of the tools for use in their own work, instead of being able to view the information, but relying on others to complete tasks designed around old practices.

In summary, sociotechnical systems theory formed the core of ComTET research, driving investigations into how work in the submarine control room could be optimised. The results of the ComTET experiments, especially the baseline, were utilised to drive this research forward. However, this was through using the insights gained, as opposed to direct application of the methods applied here. While this contextual link is not made explicit throughout, it served as confirmation of the problem explored within, and is complimentary to the methods directly employed to design the interfaces to investigate a resolution.

2.4 Ecological Interface Design

EID is a theoretical framework for designing Human-Machine Interfaces (HMIs) of a complex nature (Vicente and Rasmussen, 1992), making the affordances and constraints of a system and its environment apparent to operators (Van Dam, Mulder and van Paassen, 2008). Constraints are defined by Vicente (1999b) as relationships between, or limits on, behaviour (Salmon *et al.*, 2010). The framework is based on Rasmussen's Skills Rules Knowledge (SRK) Taxonomy (Rasmussen, 1983) and Abstraction Hierarchies (Vicente, 1999b; Jenkins *et al.*, 2009). In this context, ecology refers to the reflection of the actual environment, maximising ecological validity (Brunswik, 1956). This is achieved by inferring distal variables from proximal variables (King, Read and Salmon, 2022), therefore making visible the invisible (Vicente and Rasmussen, 1987). Distal variables are objective representations of state, and proximal variables are sensory inputs received from an environment (an organism's ecology) (Vicente and Rasmussen, 1987). Brunswik (1957) believed that distal variables could not be perceived directly, and that they must be perceived through imperfect (i.e., probabilistic) proximal variables (Vicente and Rasmussen, 1987; King, Read and Salmon, 2022). However, Gibson (1979) argued that users could directly perceive the environment without intermediary processing, viewing their environment in a goal-oriented manner (Rasmussen and Vicente, 1989) (Vicente and Rasmussen, 1987). EID capitalises on this direct perception by representing the work environment and possible affordances directly to users, reducing unnecessary cognitive processing (Gibson, 1979; McIlroy, 2016). The aim is to transition cognitive tasks into perceptual tasks, making invisible or imperceptible aspects of the work domain visible in a virtual ecology, such that they can be appropriately acted upon at the lowest possible level of cognitive control (Van Dam, 2014; Cravens, 2021).

EID has two objectives, not forcing cognitive processing to a higher level than tasks require, and supporting all levels of cognitive control described by the SRK Taxonomy (McIlroy and Stanton, 2015b). It has been shown to improve performance compared to traditional (non-EID) systems (Vicente, 2002), increase work domain transparency (Van Dam, Mulder and van Paassen, 2008), reduce workload (Nielsen, Goodrich and Ricks, 2007), and reduce memory requirements (Lau and Jamieson, 2006). Given the synergistic goal for both EID and a submarine command room of understanding environmental constraints, it is posited that the framework would be suitable and could potentially address issues identified with current HMIs.

Reduction of cognitive demand is achieved by displaying physical and functional information from the AH in an ecological manner, allowing the HMI to take advantage of human perception and psychomotor abilities (Dinadis and Vicente, 1996). Traditional HMIs only present physical information, whereas EID presents functional information in addition. In a HMI, physical information represents and describes the status of system components and Functional interface information is representative of system structure as well as its constraints (Pawlak and Vicente, 1996). This can lead to better performance than either alone (Torenvliet, Jamieson and Vicente, 2000; Vicente, 2002), which could combat the issue Woods, Patterson and Roth (2002) identified of increased control room data potentially exceeding an operator's capability to effectively process it; operators would be able to process more data, without compromising their ability to do so.

Each SRK taxa correlates to an EID design principle, supporting behaviour based upon it (Vicente and Rasmussen, 1992): Skill-Based Behaviour (SBB) requires that an operator should be able to manipulate the interface directly, with objects being isomorphic to what they represent, Rule-Based Behaviour (RBB) requires a consistent one to one relationship between signs and constraints, Knowledge-Based Behaviour (KBB) requires that the work domain is represented as an AH, providing an external mental model to support knowledge-based problem solving. Support of different levels of cognitive control is achieved by observing these design principles, and facilitating innate response mechanisms (Vicente and Rasmussen, 1992), such as recognising red as dangerous or understanding the environment from looking at a tactical picture. The design process can be informed by semantic and attentional mapping processes (Watson and Sanderson, 2007). These are defined by Zestic *et al.* (2019) as mapping perceptual forms to information to display information in a meaningful way, and directing attention at appropriate times using appropriate methods, respectively. By implementing design principles in HMIs to provide support for operators of different skill levels and cognitive requirements, there is a potential to make them easier to use, which could have usability as well as performance benefits for the entire control room.

2.4.1 Types of Domains

There are multiple potential domains of application for EID, each with their own idiosyncrasies, processes, and challenges. Rasmussen, Pejtersen and Goodstein (1994) developed a continuum for the categorisation of these domains, based on the nature of the events expected to occur (Bennett, Posey and Shattuck, 2008). Law-driven (also known as causal) domains, based on a system's physical structure and functionality, are one end of the continuum. In these domains, highly trained and frequent users respond to demands originating from the domain itself. An example domain would be process control, which is largely driven by physical laws. At the other end of the continuum are intent-driven domains, in which events arise from a user's intentions, goals, and needs. Users tend to interact with these domains, such as information searching, on a more casual basis, with their skills, training, and knowledge being more heterogenous. Rasmussen (1999) links the domain continuum to constraint types, discussing how the structure of a work domain will in turn affect the types of constraints, and whether they are intentional or casual. Thus, it is vital to understand how a domain is categorised, to understand what constraints will be pertinent to creating an interface that is suitable for its representation. Pertinent to this understanding is that no system is purely causal or intentional, but rather influenced by both ends of the continuum in varying degrees (Burns, Bryant and Chalmers, 2005); one should not pre-emptively omit constraints based on a domains perceived position on the continuum. This is demonstrated in a review by Bennett and Flach (2019), who found that the domains EID had been applied to were predominantly mixed.

Talcott, Martinez and Stansifer (2007) and Bennett and Flach (2011) highlight that for domains on either end of the continuum there are differing optimal design methods. For law-driven (causal) domains, analogous visual displays should be developed, utilising geometric forms representative of domain constraints, such as the work of Vicente (1991b). For intent-driven domains, spatial metaphors should be developed, relating interaction requirements to familiar concepts and activities, such as the work of Pejtersen (1992). However, the optimal design method for intermediate domains is less clear, due to the equal requirement from both ends of the continuum (Bennett, Posey and Shattuck, 2008). Domains in this category include military command and control, and computer network defence (Bennett, 2014), as they are constrained by physical equipment capabilities and also the intentions of other actors within the system. This includes the submarine control room, which is driven by a variety of factors from both ends of the continuum, such as the: laws of physics for sonar, limitations of technological processing, command team structure, application rules and regulations, and goals of the submarine at any given time. Thus, designs for submarine control rooms, Sonar and TMA in this thesis, a mix of design approaches should be utilised, in line with the work domain being mixed.

2.4.2 Why not User-Centred Design?

This thesis uses EID and associated theory (WDA, WCA, and SRK Taxonomy) and explores their utility for updated Sonar and TMA HMIs. However, EID is not the only viable option for achieving the goal, with other methods available, such as User-Centred Design (Norman and Draper, 1986; Norman, 1988), Human-Centred Design (HCD; Norman, 2013; International Organization for Standardization, 2019), and Activity-Centred Design (ACD; Norman, 2013), which all have a focus on the user and their activities over their work environment. Whilst any of these methods could have been selected, their principles and context in which they are applied is vital to their efficacy. Consequently, it was prudent to examine if EID, focusing on the work domain, was appropriate. Perhaps one of the most popular methods above is UCD (Saffer, 2010; Chammas, Quaresma and Mont'Alvão, 2015; Hasani *et al.*, 2020). It originated from Donald Norman's research laboratory, becoming widely used after books introducing the concept (Norman and Draper, 1986; Norman, 1988). This section will explore the reasoning behind selecting EID over UCD in the context of submarine Sonar and TMA. However, it is important to note that EID is not designed to replace UCD, but rather compliment it (Burns and Hajdukiewicz, 2004; Kwok, 2007), an approach which is explored in Chapter 10. In this respect, the discussion in this section is designed to advocate for the use of EID within the context of the research goal – that is, to design novel Sonar and TMA interfaces that are capable of future requirements.

UCD is a prevalent design approach (Vredenburg *et al.*, 2002). It makes the user a core consideration of design (Williams, 2009), focusing on effective usability from their point of view to meet their needs and interests (Abras, Maloney-Krichmar and Preece, 2004), including them at all stages of the design process (Williams, 2009; Kleiner *et al.*, 2015). This benefits the user(s) by focusing on them and the activities that they are seeking to complete, orienting subsequent designs such that they are accessible, understandable, and have predictable outcomes. By contrast, EID is oriented around the assumption that the environment drives user behaviour, and designs for this (Kwok, 2007). While the users are still involved, likely through subject matter expert input to CWA artefact creation and validation (Burns and Hajdukiewicz, 2004; Naikar, Hopcroft and Moylan, 2005), their input is not as significant as with UCD.

However, Burns and Hajdukiewicz (2004) argue that asking users to co-design might not yield the requisite information in the first place for two reasons. The first is that even very experienced users might not be aware of how the entire domain works and would provide information from their own perspective. The second is that users might not be aware of the constraints that underpin their domain. Without explicit elicitation, as with CWA, these might not be accounted for in the redesign, even with user involvement. This is reminiscent of the XY problem in technical support, where questions are about the end-user's attempted solution (Y), over the problem itself (X). Applied to awareness of domain constraints, users might be aware of what actions they need to take, but the underlying constraints that drive them. Consequently, any derived solution would be bound by the remnants of design choices (Y), over the constraints that should actually drive these (X). Rather, they advocate for the constraints to be discovered first, and designing a system to account for this. As will be discussed in detail in Section 4.1.1.4 and Section 4.2, the systems being discussed in this thesis, especially Sonar, are a product of evolution as a measure of risk avoidance. This runs the risk of anchoring user input to pre-existing designs with UCD, over examining how they could be changed using EID, within the work domain’s constraints identified by users.

Another factor to consider is that operators are highly trained and experienced; they know how to use their HMIs, and they know how to complete tasks using them. The application of UCD could yield usability or performance improvements. However, without an explicit focus on the environment and constraints, these vital considerations may be omitted from the analysis and design (Davies, Burns and Pinder, 2006). As the factors are demonstrable key factors in submarine control room operation, it would be remiss to choose a design method that did not explicitly consider them. A UCD approach may improve task efficacy, although focuses on certain well-defined tasks (Burns, Kuo and Ng, 2003). Submarine accident reports have shown that novel and unanticipated events that degrade performance of the control room sociotechnical system are typical causes (see Section 1.3). There is no possibility of managing to account for every situation that a submarine might encounter in the future to design tasks around, and guessing is not appropriate (Burns and Hajdukiewicz, 2004). However, there is a possibility of providing users with the ability to assess their situation and act appropriately, a capability which must be appropriately designed for, especially where a user’s ability to improvise might outperform automation solutions (Chalmers, Easter and Potter, 2000; Van Dam, Mulder and van Paassen, 2008).

The importance of handling non-routine events is exacerbated by the increasing use of automation. In complex sociotechnical systems, such as a submarine control room, technology can autonomously perform routine activities, leaving operators to largely act as knowledge workers, and to handle nonroutine as well as unanticipated situations (Vicente, 2002). These situations can range from mundane disturbances requiring a work-around a few times an hour (Norros, 1996) to extremely sporadic catastrophic events. Few theoretical frameworks explicitly address worker adaptation in response to these situations (Vicente, 2002), which can have disastrous outcomes if not handled correctly.

Again, this is a clear advantage of EID over focusing on well-defined tasks; no matter how well-crafted the interface is, by virtue of concentrating on pre-defined tasks, problem solving is bounded, which could preclude successful adaptation to novel events (Naikar, Hopcroft and Moylan, 2005). This corresponds with a review of maritime autonomous surface ships by Dreyer and Oltedal (2019), where unanticipated undesired events were identified as a design challenge. They argued that is impossible to design for all potential scenarios, and therefore design focus should shift to providing appropriate design resilience, citing (Ahvenjärvi, 2016; Wróbel, Montewka and Kujala, 2017;2018).

EID is well poised to address this, as it has a core premise of promoting an understanding of the system to users to enable them to use, manage, and diagnose a system effectively to understand their work domain to plan and execute tasks in a goal-oriented manner. This contrasts with traditional user- and task- focused approaches, which place their focus on the operator and the completion of specified, well-defined, tasks (Vredenburg *et al.*, 2002; Burns, Kuo and Ng, 2003; Naikar, Hopcroft and Moylan, 2005) to achieve the same aims. This approach limits flexibility to model the work of operators where it is driven by behavioural variability and emergent behaviours arising from sociotechnical systems (Vicente, 1995; Chalmers, Easter and Potter, 2000). Consequently, while there would be merit in improving the individual tasks that operators may choose to complete to achieve their goals using UCD, this might not have sufficiently addressed the need to handle the multitude of situations that would need to be designed for, especially if they do not exist yet. This necessitates an ecological approach, as software should not be built on models incongruent with how the real-world functions based on programmed case-by-case event handling, but rather a holistic ecological domain analysis to ensure adaptation for new work constraints (Man *et al.*, 2018). The "*new*" aspect is particularly important when considering that this research is designed for an environment that is not fully defined yet and is likely to be a fast-moving objective. It was the change in work, moving operators to a centralised location for remote vessel control, that led Man *et al.* (2018) to argue for an ecological approach as conventional designs could not just be copied (Dreyer and Oltedal, 2019). While submarine interfaces are exceptionally well designed, this does not preclude changes to improve them (Stanton, 2014). Given that future requirements are likely to fundamentally change the work completed by submariners, it would be appropriate to avoid replicating current interfaces, albeit with generational improvements, and move towards an ecologically aware approach.

Furthermore, there is evidence across the literature that EID also yields several benefits during routine situations, which should account for the majority of a system's operating time. Burns, Kuo and Ng (2003) identified three main benefits of using EID over other interface methods such as UCD: a model of system functioning allows critical information to be identified, a display structure can be derived from the system's structure, and operators typically demonstrate faster, more accurate, problem diagnosis behaviour. These benefits are achieved by representing work domain information as physical and functional information, which allows the interface to take advantage of human perception and psychomotor abilities (Dinadis and Vicente, 1996). It is not possible to explicitly confirm the alternative approaches used for these 'traditional' interfaces outside of literature that explicitly identifies them, such as Harre (2019) and Harre and Lüdtke (2018) stating that User-Centred Design (UCD) was a conventional approach in their research. However, given the popularity of UCD (Vredenburg *et al.*, 2002), and its derivative methods (HCD and ACD), it would not be illogical to consider them as physical Information only interfaces, and thus conventional from the standpoint of EID literature. No matter how well a submarine interface has been designed, if functional information is not included, then something has been 'left on the table', as using both can lead to better performance than either in isolation (Torenvliet, Jamieson and Vicente, 2000; Vicente, 2002). EID includes this information, allowing the HMI to take advantage of human perception and psychomotor abilities (Dinadis and Vicente, 1996), which could address the ongoing challenge of ever-increasing data threatening to exceed an operators capability (Woods, Patterson and Roth, 2002; Dominguez *et al.*, 2006).

In the context of sociotechnical systems in complex work domains, the benefits of EID over traditional interface design methodologies are evident. Capitalisation on an operator's innate psychomotor capabilities to enhance performance and facilitate system understanding during routine operations is a clear benefit. However, where EID truly distinguishes itself is a core principle of utilising these benefits to serve a modern requirement for knowledge workers in complex sociotechnical systems, in addition to supporting work adaption in nonroutine, and potentially catastrophic, situations. Thus, the choice of EID over UCD was made to align with the principles of submarine control room operation, where there is a significant focus on understanding the submarine's environment through observation to understand the constraints of action in a multitude of complex and unforeseen situations.

2.5 Conclusion

This chapter introduced the underpinning theory for the thesis, covering the SRK Taxonomy, CWA (focusing on WDA and WCA as precursors to EID), and EID itself. EID is a design method for complex sociotechnical systems, aiming to make the constraints and affordances of a work domain apparent (Van Dam, Mulder and van Paassen, 2008). Taking an ecological approach (Brunswik, 1956) to representation, EID provides users with HMIs that allow for direct perception of their environment without intermediary cognitive tasks where possible (Gibson, 1979) to support goal-oriented behaviour (Rasmussen and Vicente, 1989).

It builds on the SRK Taxonomy (Rasmussen, 1983; Rasmussen and Vicente, 1989), which classifies perception and response to different categories of tasks (Drivalou and Marmaras, 2009). The aim is to support all levels of the SRK Taxonomy (McIlroy and Stanton, 2015b), primarily enabling operators to utilise SBB as much as possible, reducing their cognitive workload. This benefits them for routine operations, but also frees up cognitive capability to handle non-routine, unexpected, or novel events.

EID also builds upon WDA (Rasmussen, 1985; Vicente, 1999b) and WCA to understand the constraints which define and bound the work domain. The outputs of these analyses are used to inform the EID design. Physical and functional information from the AH (WDA output) is included in the design and how it is organised. While traditional interfaces typically only present physical information, EID interfaces include functional information, the combination of which can lead to better performance than either alone (Torenvliet, Jamieson and Vicente, 2000; Vicente, 2002). The WCA informs design decisions to ensure that all user skill levels are accommodated for, and that behaviour across the SRK Taxonomy is supported.

This chapter also presented the reasoning for choosing EID over UCD, a prevalent design approach (Vredenburg *et al.*, 2002). It was argued that while UCD could improve the Sonar and TMA HMI, choosing EID would be better suited to the challenges posed. Key to this was EID’s focus on supporting goal-oriented behaviour within bounds of possibility, as opposed to focusing on specific tasks, providing required flexibility to address the multitude of situations that would be encountered, and especially those that would not be conceived at design. That is not to discount UCD as a suitable approach to yield improvements however, rather that EID was chosen as it explicitly accounted for the challenges identified for the redesign of Sonar and TMA.

Contemporary sociotechnical systems continue to advance in capability and complexity, with ever-increasing benefits. These benefits are realised by enabling the social and technical components of these systems to leverage each other; technology provides vast processing capability and humans can provide intelligence, adaptability, and creativity. However, all usage eventualities cannot be accounted for, especially as these systems continue to expand in remit and complexity. EID helps to address this by concentrating on bounding the work domain using constraints that users can operate within to achieve their goals. Despite constraints being key to EID, there is not currently literature that presents what types of constraints have been found across the literature for others to use in their own systems analysis. Chapter 3 addresses this issue, detailing the creation of a taxonomy of constraints across the EID literature to support analysis activities.

Chapter 3 **A Taxonomy of Constraints**

3.1 Introduction

Chapter 2 introduced the theory of Ecological Interface Design (EID), which explicitly recognises that there are multiple solutions to control problems in complex sociotechnical systems and concentrates on facilitating operator autonomy within a constrained action state-space (Borst, Flach and Ellerbroek, 2015) to achieve required goals. This fundamental contribution of constraints, shaping an operator's action state-space, makes them an integral aspect of EID and its underlying theory; without proper representation and comprehension of constraints, operators may select improper courses of action, or be unaware of any violations.

Therefore, this chapter further explores constraints, undertaking a systematic literature review to build a taxonomy of constraints that are present in the literature itself that can be used for systems analysis purposes when completing Work Domain Analysis (WDA) and/or an EID design. The taxonomy represents an additional tool for analysts, either standalone or to be used with other concepts for practitioners to enhance their methodological process. To the author's knowledge, no current approaches to undertaking CWA or EID incorporate a taxonomy of constraints sourced from the literature in their approach, and no such taxonomy is forthcoming in the literature. There is also a dearth of literature showing how constraints are connected and the strength of these connections. Social Network Analysis (SNA) was therefore applied to reveal how different categories of constraints were connected, making it possible for practitioners to identify if they should explore other types of constraints if they are finding constraints of a linked category. Overall, the taxonomy is designed to be used by practitioners to maximise elicitation of constraints for a given work domain.

Maximising understanding the constraints of a given work domain is vital, as violations of constraints can create the potential for catastrophic outcomes (Rasmussen and Vicente, 1990), such as the Three Mile Island incident or the Ladbroke Grove rail crash, in which broken constraints were fundamental causal factors. For Three Mile Island, the constraint violation was the loss of sufficient coolant such that safe system operation could be maintained, exacerbated by a design flaw that did not alert operators that they were in a situation where they had violated this constraint (Hopkins, 2001). For Ladbroke Grove, the constraint violation was the passing of a signal at danger (Cullen, 2001). In both instances, the violation of a constraint in the work domain led to an accident occurring, demonstrating that constraints are an integral aspect of sociotechnical systems. However, the impact of constraints is not limited to the critical path for accidents;

constraints can influence all aspects of sociotechnical systems, including design, training, and operation.

This importance of constraints, both to EID and the safe operation of complex sociotechnical systems in general, is not lost on human factors practitioners. There is a plethora of literature mentioning constraints in the EID corpus, covering the underpinning theory (e.g., Rasmussen and Vicente, 1990; Vicente and Rasmussen, 1992; Burns and Hajdukiewicz, 2004), domain-specific applications (e.g., Hall and Miller, 2009; Ellerbroek *et al.*, 2013b), reviews of its application (e.g., Vicente, 2002; Borst, Flach and Ellerbroek, 2015; McIlroy and Stanton, 2015b; Bennett and Flach, 2019), or literature that seeks to assist practitioners with elicitation of constraints during systems analysis (e.g., Rasmussen, Pejtersen and Goodstein, 1994; Naikar, Hopcroft and Moylan, 2005). Furthermore, the key underpinning analysis method, Cognitive Work Analysis (CWA), is focused on formulating an understanding of work domain constraints, as was covered in Section 2.2.

This literature, combined with a track record of successful applications (Vicente, 2002; Bennett and Flach, 2019), establishes EID as a mature and credible design method. However, there are still challenges to be addressed to establish wider adoption. When outlining unaddressed issues for EID, Vicente (2002) identified the time and effort to conduct an analysis could be prohibitive and could constrain analysis scope, a sentiment shared across the literature (Salmon *et al.*, 2007; Stanton *et al.*, 2013; Stanton *et al.*, 2017a; Read *et al.*, 2018). Several potential solutions were identified, which have been addressed, either explicitly or implicitly, across the years, such as the creation of software tools (Human Factors Integration Defence Technology Centre, 2007; Hingu *et al.*, 2017). Vicente (2002) also recommended domain-specific templates and libraries for various objects in domains. This has been addressed by several seminal works (Rasmussen, Pejtersen and Goodstein, 1994; Burns and Hajdukiewicz, 2004; Naikar, Hopcroft and Moylan, 2005; Jenkins *et al.*, 2009; Stanton *et al.*, 2017a). These works provide practitioners with a comprehensive grounding in the theory, with extensive examples of prior WDAs conducted, and comprehensive methods for eliciting constraints during CWA. Combined with a substantive amount of published CWA and EID work, of which McIlroy and Stanton (2015b) provide an overview of applications, there is an ever-growing library that practitioners can consult to synthesise a starting point for their applications.

The taxonomy presented in this chapter adds value for practitioners by addressing these challenges in three ways. Firstly, there is a dearth of EID applications to intent-driven domains, when compared to mixed and law-driven domains (Bennett and Flach, 2019), potentially limiting the available library of constraints from previous applications. A detailed list of constraints found across the literature where there are limited domain-specific examples might serve as a useful starting point in this instance. Secondly, it is a core tenet of EID to account for unanticipated events by presenting the work domain to create an action state-space. The combined expertise of practitioners and appropriate subject matter experts will likely identify almost all pertinent constraints related to a work domain. However, there remains a possibility that constraints will not be identified by either, due to their unanticipated nature. This could affect the efficacy of the designed interface. Presenting a systematic summation of all constraints found across the EID literature in the form of a taxonomy, drawn from how work is performed, could serve as verification to ensure that all types of constraints have been considered. Finally, and closely related, the summation could be incorporated into various stages of a practitioner's existing CWA process. For example, subject matter experts could be prompted about each different type of constraint or guided towards constraints that link closely to constraints already identified. Furthermore, constraints of related types could be considered together when designing the EID to further enhance understanding of how the work domain is constructed.

3.2 Constraint Coding and Category Creation

As described in the introduction, the taxonomy was constructed using the results of a systematic literature review. This methods section details how the review was completed and how the results were used to compile the taxonomy.

3.2.1 Collection, Collation, and Exclusion of Literature

Five scholarly search engines were used to obtain the literature: Microsoft Academic, CrossRef, Google Scholar, Scopus, and Web of Science. For each engine, the phrase "Ecological Interface Design" was searched for using the default settings provided. The default results view was used for all search engines, except for Web of Science. This is because the engine displayed relevant papers in all views, and as such, the literature was indexed for each type of result view (relevance, by date, by citations).

For the first one-hundred items in each result view, citation data, and a copy of the full text, if available, was downloaded. A cut-off was required as modern search engines can potentially return thousands of results, which will decrease in relevance. One-hundred was chosen as an arbitrary cut-off point as most search engines defaulted to displaying ten (initially) pages of ten results. There were no exclusion criteria at this stage. Citation data was loaded into an existing Clarivate Analytics EndNote X9 library (a library is a collection of literature in EndNote) and grouped by source search engine. A further group was created to store relevant literature that existed in the EndNote library beforehand from prior EID research but had not been returned by any searches. Furthermore, seven relevant (in the field of EID) literature reviews were identified, and their citations were stored in EndNote as groups under the first authors last name (Vicente, 1996; Reising and Sanderson, 2000; Vicente, 2002; Momtahan and Burns, 2004; Giang *et al.*, 2010; Read, Salmon and Lenné, 2012; McIlroy and Stanton, 2015b). Sources were considered relevant if they presented a literature review considering EID, CWA, or the SRK Taxonomy.

Finally, a de-duplication process was then carried out in multiple stages. For each stage, every match was reviewed manually to prevent erroneous deletion. The first stage utilised the EndNote "Find Duplicates" functionality, using the author(s), year, and title as match criteria. Spacing and punctuation were ignored. The second stage removed the author(s) as a criterion, to match instances where the author(s) name was recorded differently, such as the use of their initials or full name. Finally, the year was removed as a criterion, leaving the title, to match instances where the year had erroneously been recorded.

Once all literature was in EndNote, a smart group (automatic grouping of sources) was created. The criterium for the group was that either the title or full text should contain one of the following phrases: Ecological Interface Design, Ecological Interface, and Ecological Design. From this dynamic group, all journal papers were manually selected and moved to a separate group. Journal papers were chosen for inclusion in the taxonomy as they were peer-reviewed and typically contained a self-contained description of the application. Information from other sources was also likely to appear in journal papers, such as conference papers or theses, and as such these were excluded so as not to duplicate content. Figure 4 shows the breakdown of source types at the second stage (i.e., before narrowing of focus to journal articles). As can be seen, journal papers were by far the largest group. Further exclusion was then performed for journal sources that:

- Had no full text available.
- Existed as another source with a different name.
- Did not use English.
- Had unresolvable gaps in required bibliographical information, making it impossible to identify a source, and subsequently access full text.
- Presented contributions that were not EID, such as ecology design papers.

This process left 202 journal papers deemed suitable for inclusion in the analysis. The data for all sources was exported to an eXtensible Mark-up Language (XML) file from EndNote, which was imported into QSR International's NVivo 12. The number of sources at each stage is shown in Figure 5, starting from all sources found, with the number of sources being systematically reduced to a manageable number for further analysis in NVivo. Additionally, Figure 6 shows the domains of the remaining files in the analysis, derived from how the authors had tagged their research.

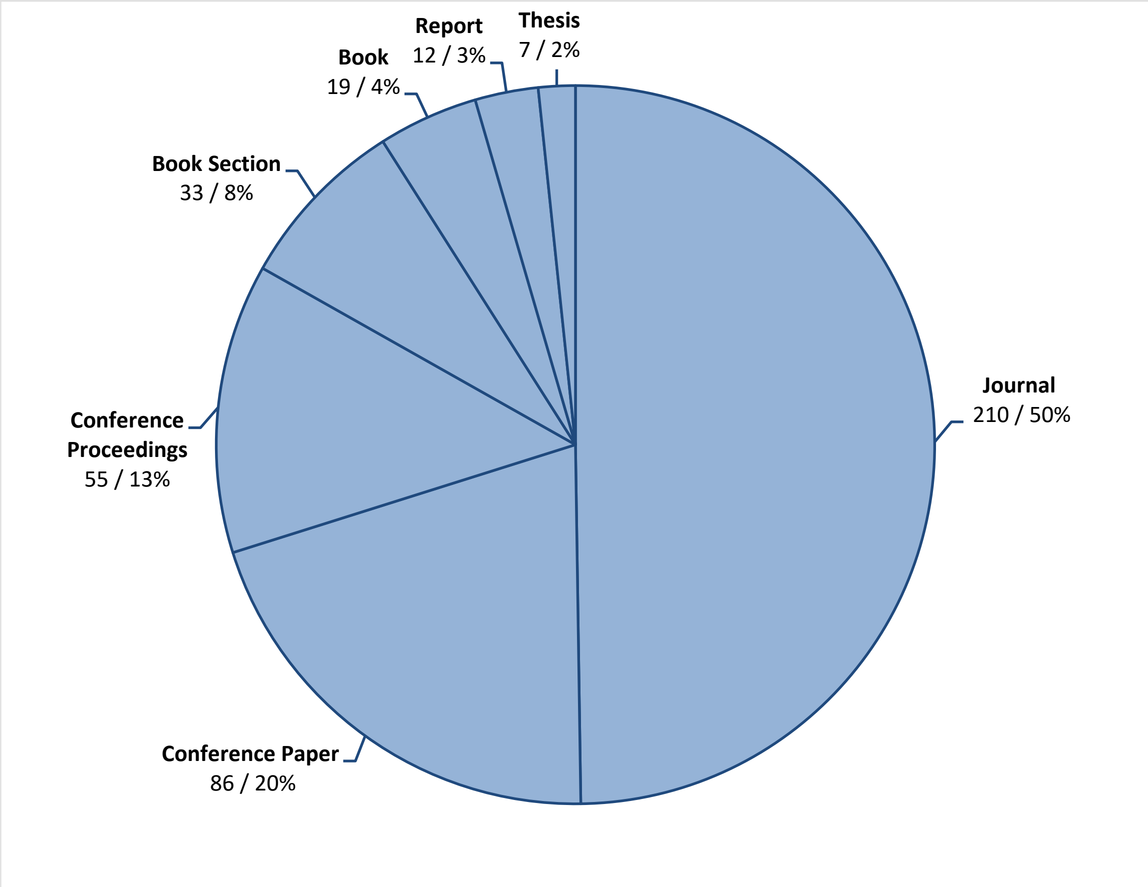


Figure 4 – The distribution of source types from the systematic literature search

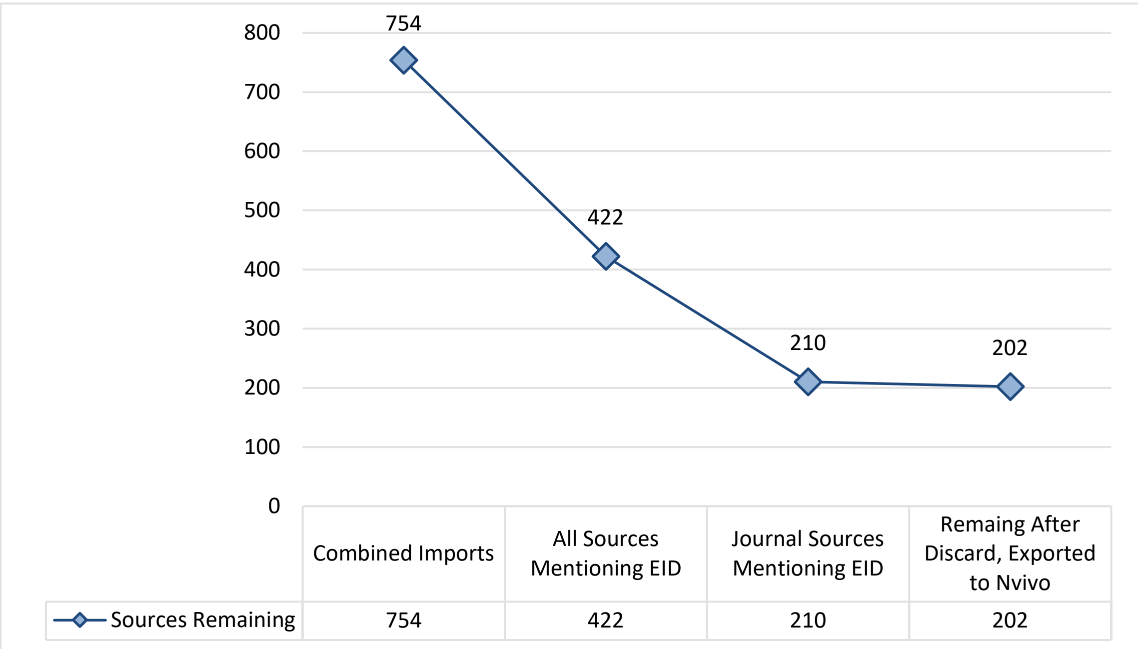


Figure 5 – Line chart showing the number of sources remaining at each stage

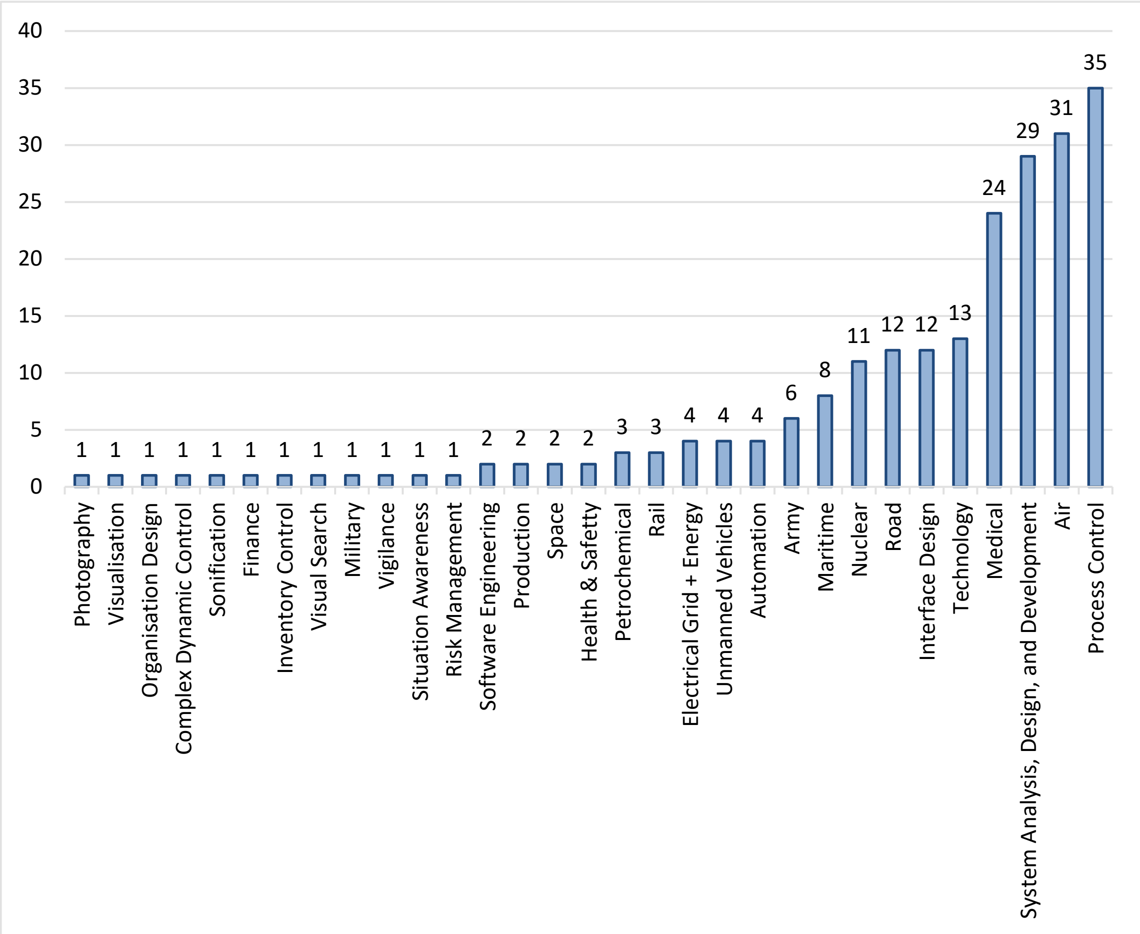


Figure 6 – Column chart showing domains in the literature, and the number of files in each

3.2.2 Coding of Context for Constraint Mentions

To locate the different types of constraint mentioned in the literature, two coding steps were taken: a wildcard textual search and contextual coding (coding prose around the actual match, known as its context). A wildcard search is one where matches are returned based on exactly matching any provided text and returning anything in lieu of a special character (or multiple characters, usually '*' or '?') in the search term. For example, "constrain<u>*</u>" could match "constrain<u>t</u>", "constrain<u>ed</u>", "constrain<u>s</u>", or even "constrain" itself. Contextual coding is where the prose from around the match is marked for further processing. To perform the wildcard textual search the text search query function was used to automatically code matches. The search term was "constrain*", searching in "Files & Externals" for "Exact Matches". The asterisk in the search term was interpreted as zero or more non-whitespace characters by the search engine. The search returned 179 journal articles, with 2770 matching references to "constrain*" within them.

Context coding was then performed to compile a list of prose that could be reviewed to identify constraint types within them. This was performed to ensure as many different constraints as possible were captured, as the word constraint might not have been near the type(s) of constraint it was describing (e.g., a table of constraints, a list of constraints, a paragraph that introduces constraints and then details specifics without repeating "constrain*"). This was completed manually as the automated feature in NVivo did not work as expected for the purposes required. This resulted in the prose associated with the matched reference to "constrain*" to be coded. As an example, consider the fictional sentence from a paragraph "… . The operators were constrained by time. …". The search would have returned the word "constrained", and the context would have been the surrounding paragraph. Relevant context for each match, typically the smallest selection of sentences possible, was manually coded to a "Constrain* Context" code. Exclusion criteria at this stage were mentions in the bibliography, which were not coded. Once complete, the "Constrain* Context" code contained 173 journal articles and 1,833 references to review for specific types of constraint. The reduction in files was due to the exclusion criteria and the reduction in nodes from "constrain*" codes that were co-located in contexts (i.e.: one context contained multiple matches from the search for "constrain*").

3.2.3 Definition of a Constraint

At this point, it is pertinent to detour slightly to understand how a constraint is defined across the literature, and whether this is aligned with the definition provided by Vicente (1999b), which was relationships between, or limits on, behaviour (Salmon *et al.*, 2010). This is pertinent as the taxonomy will be based on constraints identified in the literature, which should have congruent definitions to Vicente to best comply with the underlying theory.

There are two prevailing definitions of constraints across the literature: that constraints are the product of relationships between variables, or that they are the limits and boundaries placed on individual variables. However, the definitions are not dichotomous across the literature. Segall *et al.* (2013) specified that single variables in the work domain served as a basis for defining single-variable constraints, with upper and lower limits being used as bounds. In addition to single variable constraints, multivariate constraints were defined between two or more variables that could be expressed as equations. This is similar to other joint definitions (Jamieson and Vicente, 2001; Mazaeva and Bisantz, 2014; Flach, 2017), which identify limits as being simple constraints and relations forming complex constraints. Furthermore, Bennett and Flach (1992) also describe multivariate relations as being constraints (high-level constraints) and identify single variables (low-level data), but without reference to constraints. Instead, these are two opposite ends of a continuum of increasingly complex relations (properties or constraints) that categorise system states.

Table 4 provides a summary of these definitions, grouped by their type. As can be seen, there are many different definitions for constraints. However, they all have a relational aspect in common, where what is being constrained is related to the boundaries created by any number of variables, an equation, or requirement(s). This is congruent with the definition provided by Vicente (1999b), providing confidence that the definition of a constraint has been upheld from the originating literature.

Table 4 – Definitions of constraints across the literature, modified for British English spelling and brevity

	Author(s)	**Definition**
Relations	Flach, Stappers and Voorhorst (2017)	"The construct of affording is intended to draw attention to relations that constrain action possibilities."
	Vicente and Rasmussen (1992)	"First, when the system is functioning correctly, it can be described by a set of constraints…. These relationships can be described as constraints." "This relationship provides a very important source of constraint that can be exploited in problem solving."
	Bennett (2017)	"… the abstraction hierarchy, provides a complementary dimension for modelling domain constraints in terms of "means-ends" relations…"
	Bennett, Posey and Shattuck (2008)	"It provides a template, …, that can be used to categorize the critical characteristics (sometimes referred to as the relational invariants, constraints, …, or means-ends) of a domain."
	Flach (2017)	"representing the functional meaning or deep structure of a work domain in terms of a hierarchy of 5 layers of constraints that reflected both the top down relations between function and form … and the bottom up relations between form and function …"
	Vicente (1992)	"These constraints describe the redundant relationships that exist between process variables at a single point in time."
	Effken, Kim and Shaw (1997)	"The major constraints shown describe the structural relationships, or connectivity, of the system." "… designers must (a) identify relevant process variables in the task, (b) organize those elements so that their relationships correspond to higher-order system constraints, …"
	Effken (2001)	"If not, then the experts evaluate the constraints on SVO2, such as the relationship between the patient's oxygen delivery and oxygen consumption (a higher order variable)."
	Talcott, Martinez and Stansifer (2007)	"The abstraction (means-ends) and aggregation (part-whole) hierarchies are analytical tools that have been developed to discover the constraints (i.e., the relational invariants) of a work domain."
	Shier *et al.* (2018)	"The WDA previously developed [7] showed relationships between variables and the variable constraints…"
Limits	Burns, Bryant and Chalmers (2005)	"These constraints show the limits on the information that can be sensed and processed by the ship."

	Drivalou (2005)	"Through it we identified constraints related to the capabilities and limitations of the operators' cognition…"
	Ellerbroek *et al.* (2013b)	"Together with the velocity vectors, the horizontal and vertical SVEs relate to the safety goal, by showing how internal constraints (available power, structural limits, etc.) limit possible velocity vectors."
	Ellerbroek *et al.* (2011)	"For flight in general, comfort poses constraints such as upper limits on manoeuvre accelerations."
	Vernon, Reising and Sanderson (2002)	"Constraints are boundaries that distinguish what is physically or procedurally possible or desirable from what is impossible or undesirable. Constraints may be mechanical limits or physical limits, or they may be safety boundaries."
	Van Dam, Mulder and van Paassen (2008)	"Limitations to aircraft performance (constraint at physical level), such as maximum and minimum values for aircraft velocity, can be applied. Due to productivity (a more functional workspace constraint), the heading change is limited to 90° port and starboard…"
	Olsson and Lee (1994)	"Operation is constrained by: Maximum and minimum operating limits on the evaporator pressure…"
	Naikar and Sanderson (1999)	"The functional structure of a work domain may be illustrated in an abstraction hierarchy (AH)" … "Each layer identifies different kinds of functional constraints; the upper layers identify intentional or purposive constraints, and the lower layers identify physical constraints."
Both	Mazaeva and Bisantz (2014)	"Then, a set of constraints, or performance capabilities and limits, for each variable was identified, such as maximum and minimum limits for exposure or graininess… Some constraints involve relationships between several variables…"
	Jamieson and Vicente (2001)	"Also visible on both Paulsen displays are constraints imposed by physical or relational limitations…"
	Flach (2017)	"In addition, however, regulatory and other pragmatic constraints are considered at this level to the extent that they place limits or bounds on system performance. Physical constraints are typically represented at this level as algebraic or differential equations."
	Segall *et al.* (2013)	"Variables identified in the work domain model were used as a basis for defining constraints and to guide the display design process. For example, single variable constraints usually place upper and lower bounds on a variable. Information on such constraints can be used to display ranges of scales, determine alarm limits, and so on. Multivariate constraints are relationships between two or more variables that can be expressed as equations… Finally, means–end relationships describe the implication of one variable in the value of another."
	Bennett and Flach (1992)	"The term low-level data refers to the measured values of individual process variables. In contrast, the term high-level constraints refers to relations that exist between these process variables."

3.2.4 Constraint List Synthesis

After all contextual prose for each search match had been coded, the specific constraints were coded. All prose coded as "Constrain* Context" was read, and constraint codes were added where a specific type of constraint was mentioned, where factors constrained other factors, and where context otherwise dictated (constraints in tables, constraint types not necessarily near the "constrain*" highlight, etc). A small number of constraints were excluded from coding, due to being descriptors, or referring to experimental method constraints.

3.2.5 Category Generation

Coding all types of constraints in the literature yielded a list of 904 types of constraints that appeared in the 1,833 pieces of prose coded to review. To make this list more manageable for analysis, they were grouped into categories. Categories were flexible and not limited in number. The categories were not preconceived and were driven by the identified constraints. An iterative, bottom-up, approach was taken to group constraints together into groups. As many similar constraints existed, such as 'time' and 'temporal', these were first grouped together and given a generic name, such as 'temporal'. Once immediately similar constraints had been grouped, further iterations sought to add related constraints to groups, such as adding 'schedule' to the 'temporal' group. At each iteration, constraints were added to the existing categories if they were suitable. If a suitable category did not exist, a constraint was temporarily skipped, and a candidate new category was kept in mind until a few related constraints could be identified and given a group name. This process continued until all constraints were categorised.

The final list contained twenty categories of constraints, of which three contained subcategories. These were included in their parent category in the taxonomy, as they were a subset of their parent. Figure 7 presents the sum of coded files in each top-level category, with categories containing an average (M) of 54.10 files (SD = 32.04). The categories are detailed in Table 6, which lists the categories in descending order by the number of articles coded to them. The top three nodes for each category are listed, based on the number of files they appear in, with examples of the category from the literature. In cases where the third place is tied between several nodes, it was decided by the reference count. Should this not have separated them, third place was omitted.

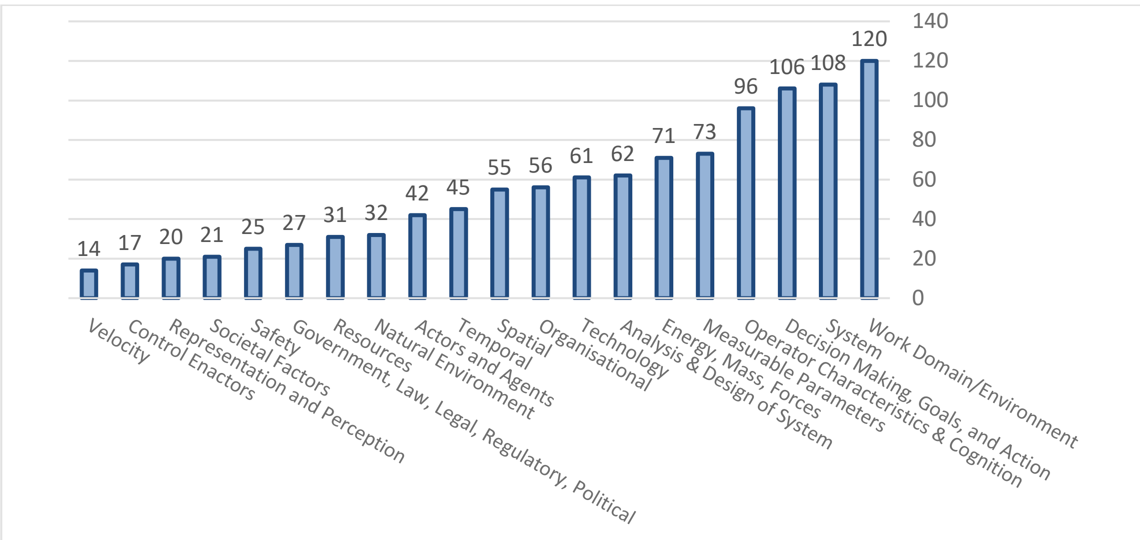


Figure 7 – Sum of Files Coded for Each Category

3.2.6 Inter-rater Reliability

Inter-rater reliability was conducted with another member of the research group, who had a background in Psychology, with one year of experience in Human Factors. They were provided a copy of the NVivo project with all constraints removed from categories, and a table containing a list of definitions. They were provided with a list that consisted of the Top 10% and bottom 5% of constraint types across all codes. This was to explore the effect, if any, on agreement with constraints that had more literature to define them. Additionally, they were asked to select a further 5% of nodes to code themselves, based on their education and experience. The additional rater was asked to categorise the nodes into categories using NVivo and make a note of where constraints could have fit more than one category, as NVivo does not permit a code to be in multiple categories. For these constraints, they were determined to be a match if at least one category chosen matched. Cohen's κ was calculated for each category and overall using SPSS. The results are shown in Table 5, which shows that for all categories but the bottom 5% a moderate level of agreement was observed using the categories of agreement defined by Landis and Koch (1977).

Table 5 – Results of inter-rater reliability for the constraints taxonomy using Cohen's κ

Group	**N**	**Cohen's κ**	**Agreement Category**	**p**
Bottom 5%	46	.307	Fair	< .001
Random 5%	45	.426	Moderate	< .001
Top 10%	92	.476	Moderate	< .001
Total	202	.496	Moderate	< .001
*Uncalculatable**	*19*	*N/A*	*N/A*	*N/A*

** SPSS reported that "no statistics were computed due to constants", a message shown when there is no variability between rater scores (IBM, 2020).*

Table 6 – Categories of constraint identified for the constraint taxonomy and associated definitions. '→' denotes a child of the last category without a prefix.

Category (Files / References) Notes/Examples	Top 3 Concepts (Files / References)	Description
Work Domain / Environment (120/591)	1. Work Domain (103/376) 2. Physical (37/93) 3. Functional (17/31)	Constraints regarding the work domain and its environment, excluding the system itself. This distinction is present in the theoretical foundations of EID, describing the system itself and the work domain as separate entities, such as "the human-machine system must take into account, or embody, the constraints inherent in the work domain" (Vicente and Rasmussen, 1992), citing from (Stassen, 1989).
Work Domain constraints appeared across a majority of domains and were general, encompassing all constraints that might be present, providing a direct link to the EID through understanding and displaying them (Vicente and Rasmussen, 1992; Talcott, Martinez and Stansifer, 2007; Miller and Feigh, 2019). Physical constraints, both in the sense of corporal objects and physical information, related to WDA and EID (Naikar and Sanderson, 1999; Rasmussen, 1999; Burns, Bryant and Chalmers, 2005; Borst, Flach and Ellerbroek, 2015). Functional constraints also related to WDA and EID, being revealed by them (Naikar and Sanderson, 1999; Effken, 2006; Upton and Doherty, 2008; Klomp *et al.*, 2016).		
System (108/531)	1. System (56/134) 2. Purpose (19/25) 3. Design (Of System) (15/29)	Constraints directly related to the system, including its affordances, function, structure, and purpose.
As system constraints referred to all constraints within a system, it was expected that they would be general. The constraints detailed covered basing a systems design on constraints (Niskanen, 2018), internal and external constraints (Drivalou and Marmaras, 2009), system constraints contextually limiting functionality (Jenkins *et al.*, 2010b), and differing levels of constraints (Effken, Kim and Shaw, 1997). Purpose constraints included a system's purpose and reason for design (Vicente and Rasmussen,		

Category (Files / References)	Top 3 Concepts (Files / References)	Description
Notes/Examples		
1992), restrictions on achieving its purpose and associated design constraints (Leveson, 2000), and the ecological constraints of a domain (Naikar *et al.*, 2003). Design (Of System) constraints appeared to be the result of prior analyses that shaped a systems design (Leveson, 2000), including the outputs from CWA (Niskanen, 2018), and how design choices might affect further choices (Read *et al.*, 2016).		
→ Limits, Capacity, Restrictions (24/48)	1. Locomotion Capabilities (7/11) 2. Performance Limitations (Aircraft) (6/11) 3. Operating Capacity (3/3)	Factors that indicate the limits, capacity, and restrictions of a system. This also includes alarms or other notifications.
Locomotion Capability constraints spanned domains such as interface design (actions being possible based on capabilities; Flach, Stappers and Voorhorst, 2017), eco-driving (road grip; McIlroy and Stanton, 2015a), and RLX design (wheel-rail interface; Read *et al.*, 2016). Performance Limitation (Aircraft) constraints were largely physical (Borst *et al.*, 2006; Borst *et al.*, 2008; Van Dam, Mulder and van Paassen, 2008), but also included productivity requirements and passenger comfort (Ellerbroek *et al.*, 2013b). Operating Capacity constraints were physical, being constrained by pump curves (ranges of pressures and flow rates for nuclear power plant pumps; Lau *et al.*, 2008), transformer feeding line capacity and compatibility (Drivalou and Marmaras, 2009), and power transmission flows (Tran, Hilliard and Jamieson, 2017).		
→ Entities (15/29)	1. Traffic (10/33) 2. Aircraft (6/19) 3. Weapon (3/6)	Entities that form a system, are part of a system, or are the system. Different to actors and agents as could be entities they are controlling, or not have sentience themselves (i.e.: artefacts).
As domains associated with these entities are quite prevalent throughout the literature, this is expected. Traffic constraints were present for the road (McIlroy and Stanton, 2015a), rail level-crossing (RLX) design (Read *et al.*, 2016), and air (Ellerbroek *et al.*, 2013b) domains. Aircraft constraints related to the aircraft and its sub-systems (van Marwijk *et al.*, 2011), with extensive detail being provided in some cases (Borst *et al.*, 2006; Klomp *et al.*, 2016; van Paassen *et al.*, 2018). Weapon constraints		

Category (Files / References)	Top 3 Concepts (Files / References)	Description
Notes/Examples		
existed in the maritime (Burns, Bryant and Chalmers, 2005), unmanned vehicles (Calhoun *et al.*, 2018), and military domains (Bennett, Posey and Shattuck, 2008), mainly concentrating on law-driven constraints regarding a weapon's capabilities.		
Decision Making, Goals, and Action (106/493)	1. Goal-Relevant (18/37) 2. Strategies (For Activity Execution) (16/20) 3. Process (15/30)	Constraints affecting or describing the decision-making process, the goals to be met, and actions executed because of this. It aimed to capture constraints related to the decision-making process that were not solely tied to the operator and their cognition. These constraints were placed under the operator category.
Goal-Relevant constraints were largely mentioned as a generalised collective (Vicente, 1992; Vicente and Rasmussen, 1992; Xu, Dainoff and Mark, 1999), and some made links between display aspects and operator goals (Dinadis and Vicente, 1999). Similarly, constraints on strategy consisted of generalised collectives (Jamieson, 2007), and mentions of work domain strategies, from Strategies Analysis in CWA (Watson and Sanderson, 2007; Jenkins *et al.*, 2010b). Process constraints were typically referred to as a generalised collective, with references to process variables (Bennett and Flach, 1992) and higher-order constraints that constricted processes (Bennett and Flach, 1992; Pawlak and Vicente, 1996; Effken, Kim and Shaw, 1997).		
Operator Characteristics & Cognition (96/355)	1. Cognitive (30/47) 2. Behaviour (20/26) 3. Decision Making (20/35)	Constraints that are part of, or describe, the characteristics and/or cognitive processes and capabilities of the operator.
As cognition is vastly encompassing, including a variety of factors, it is expected that cognitive constraints was a high-level catch-all to refer to them in most cases (Naikar *et al.*, 2003; Watson and Sanderson, 2007; McIlroy and Stanton, 2015b; Man *et al.*, 2018). Behavioural constraints are similarly generic and encompassing (Vicente and Rasmussen, 1990; Rasmussen, 1999; Sanderson *et al.*, 2004; Naikar, 2017). Finally, decision-making constraints follow the same generic pattern (Trentesaux, Moray and Tahon, 1998; Effken, 2001; Araujo, Davids and Serpa, 2005; Flach, 2017).		

Category (Files / References) Notes/Examples	Top 3 Concepts (Files / References)	Description
Measurable Parameters (73/251)	1. Information (36/51) 2. Relationships (Between Variables) (20/48) 3. Variables (9/12)	Constraints associated with data, information, variables, and metrics. It does not include the processes associated with transforming data into information, and information into knowledge, rather what can be measured within the system, like the concept of a variable in computer programming. These processes were included in other nodes, such as Operator Characteristics & Cognition or Representation and Perception; this was to keep this category as aligned to the questions of "what can be measured about the system?" as possible.

Information constraints were either references to information that might be present (Effken, 2006; Jenkins *et al.*, 2008b) or constraints on information itself (Flach and Dominguez, 1995; Burns, Bryant and Chalmers, 2005). Constraints concerning the relationship between variables were largely derived from descriptions of the work domain, likely involving WDA, such as: aircraft and surrounding traffic (Ellerbroek *et al.*, 2013b); camera exposure and focal length (Mazaeva and Bisantz, 2014); and, nuclear condenser pressure and condensate temperature (Lau *et al.*, 2008). Relationships were also drawn between variables and their perceptual forms on an interface (Bennett, 2017; Bennett, Bryant and Sushereba, 2018). Also, as identified above, these relationships were in some cases termed work domain constraints themselves (Segall *et al.*, 2013; Mazaeva and Bisantz, 2014). Discrete variables were mentioned as computing variables (Wright, Mathers and Walton, 2013), the result of an equation on data (Vicente, 1992), or, as identified above, work domain constraints (Effken, Kim and Shaw, 1997; Seppelt and Lee, 2007).

Energy, Mass, Forces (71/279)	1. Energy (15/46) 2. Temperature (10/18) 3. Mass (8/8)	Constraints related to energy, mass, and forces. This included the laws of nature, energy flows, and energy in action.

Category (Files / References)	Top 3 Concepts (Files / References)	Description
Notes/Examples		

References to energy constraints were largely made in a practical sense whilst discussing the work domain, such as during flight (Amelink *et al.*, 2005; Borst *et al.*, 2008), power grid management (Tran, Hilliard and Jamieson, 2017), and process control (Vicente, 1992). Similarly, temperature constraints were also largely mentioned in a practical fashion, pertaining to the domains of petrochemical (Jamieson and Vicente, 2001), process control (Ham, Yoon and Han, 2008), and nuclear (Lau *et al.*, 2008). Finally, mass followed suit, with domains including process control (Vicente, Christoffersen and Pereklita, 1995) and petrochemical (Vicente, 1999a).

→ Laws of Nature (41/80)	1. Causal (Laws of Nature) (13/24) 2. Inherent (11/14) 3. Invariant Constraints (8/10)	The laws of nature.

Causal constraints were mentioned as the laws of nature (Rasmussen, 1999), physical laws (Borst, Flach and Ellerbroek, 2015), and the structure of a process (mass, energy, information, value flows; Vicente, 1996). Inherent constraints were less encompassing and referred to those naturally present in the work domain itself, providing examples in for air traffic management (van Paassen *et al.*, 2013), spacewalking (Feigh *et al.*, 2018), and military command and control (Talcott, Martinez and Stansifer, 2007). Invariant constraints were similar, being used to describe relationships in the environment, as well as action and effect (Waterson, Le Coze and Andersen, 2017), global goals of a nuclear power plant (Vicente and Rasmussen, 1990), and using the AH, as a key component of EID, to identify the invariant work domain constraints for process control (Ham, Yoon and Han, 2008).

Analysis & Design of System (62/175)	1. Work Domain Analysis & Abstraction Hierarchy (32/62) 2. High Level (12/24) 3. Cognitive Work Analysis (7/16)	Constraints arising from the system analysis, which can feed into the design process. As CWA is used to analyse a system for EID, it, and its constituent stages, were the primary nodes included. These nodes do not include the constraints arising from performing a CWA analysis, but constraints on the process itself and constraints related to it.

<table>
<tr><th>Category (Files / References)</th><th>Top 3 Concepts (Files / References)</th><th>Description</th></tr>
<tr><th colspan="3">Notes/Examples</th></tr>
<tr><td colspan="3">Work Domain Analysis & Abstraction Hierarchy constraints largely described the hierarchy itself (Flach, 2017), including what purposes WDA was deemed to have across the literature (Durugbo, 2012), and/or presented a WDA that described the domain (Salmon et al., 2007). This category would be useful for exploring whether any prior analysis has been done, or for gathering feedback on the analysis under construction. High-Level constraints were discussed across papers similar to “intentional” constraints, using it to refer to a group of constraints (Bennett and Flach, 1992). However, high-level constraints were typically explained in extensive detail to inform the reader of the work domain (Leveson, 2000; Ellerbroek et al., 2013b). Similar to WDA and AH constraint category, the Cognitive Work Analysis category presented CWA (Jiancaro, Jamieson and Mihailidis, 2014) and discussed results from individual stages (Jenkins et al., 2008a; Jenkins et al., 2010b).</td></tr>
<tr><td>→ Analysis Considerations (13/13)</td><td>1. Access to End-Users and subject matter experts (3/3)
2. Trial order (2/2)</td><td>Factors that could constrain the analysis and design of the system. These factors do not belong to the system itself, but to the analysis and experimental process associated with evaluating or implementing a new or improved system.</td></tr>
<tr><td colspan="3">Access to the subject matter experts (Hettinger, Roth and Bisantz, 2017) and their availability (Naikar et al., 2003) were issues that could hamper or limit analysis. Trial order constraints referred to experimental design restrictions during EID trials (Vicente, 1992; Vicente, Christoffersen and Pereklita, 1995).</td></tr>
<tr><td>Technology (61/175)</td><td>1. Technology/ Technical (20/35)
2. Design (Of Interface)/ Interface (9/14)
3. Sensor (8/9)</td><td>Constraints relating to technology, including hardware and software. Whilst certain nodes, such as software, could be considered agents, this would be contingent on the scope of CWA analysis, whereas it would always be technology.</td></tr>
<tr><td colspan="3">Technology constraints existed across varied across domains including technology (Jenkins et al., 2010b), military (Jenkins et al., 2008a), and electrical grid + energy (Tran, Hilliard and Jamieson, 2017). Specific constraints included software that would work with music devices (Jenkins et al., 2010b), networking advances enabling airborne collaborative working (Jenkins et al., 2008a), as well as screen size and resolution (van Paassen et al., 2013). Design (Of Interface) constraints were present across a</td></tr>
</table>

Category (Files / References)	Top 3 Concepts (Files / References)	Description
Notes/Examples		
smaller number of domains, and included constraints introduced by design requirements (Vernon, Reising and Sanderson, 2002) and display screen size (Miller, Scheinkestel and Steele, 2009). Sensor constraints were largely military based, such as law-based sensor constraints (Bennett, Posey and Shattuck, 2008), type of sensor (Calhoun *et al.*, 2018), as well as a sensor's range and coverage (Fay, Roberts and Stanton, 2019).		
Organisational (56/157)	1. Organisational (including "Social Organisational") (16/31) 2. "Operations" (including "Operating" and "Operational") (8/29) 3. Work (7/9)	Constraints that arise from the working environment, such as organisational, work-based, or managerial factors.
Organisational constraints included the allocation of work tasks (Naikar and Elix, 2016), the Values and Priorities (from an AH) of a work domain (Burns, Bryant and Chalmers, 2005), organisational structures (Naikar, 2017), and organisational influences recorded in an AH (Hettinger, Roth and Bisantz, 2017). Operations constraints tended to refer to specific work within systems, such as space walking (Feigh *et al.*, 2018), nuclear power plant operation (Naito *et al.*, 1995), and naval ship operation (Burns, Bisantz and Roth, 2004). In comparison, work constraints were not generally specified, rather encompassing all work that might be done in a system (Niskanen, 2018; van Paassen *et al.*, 2018), and factors that could affect this (Flach, 2017; Niskanen, 2018).		
Spatial (55/456)	1. Spatial (27/115) 2. Manoeuvring (11/81) 3. Trajectory Planning (6/12)	Constraints that are spatial in nature, including dimensions.
Spatial constraints were largely mentioned in relation to the air domain (Borst *et al.*, 2006; Ellerbroek *et al.*, 2011; Ellerbroek *et al.*, 2013b; van Paassen *et al.*, 2013). Other domains included display design (Mazaeva and Bisantz, 2014), military command and control (Bennett, Posey and Shattuck, 2008), and maritime (Fay, Roberts and		

Category (Files / References)	Top 3 Concepts (Files / References)	Description
Notes/Examples		
Stanton, 2019). Time was linked with spatial constraints, such as aircraft separation being always maintained (Ellerbroek *et al.*, 2011) or calculating a permitted travel space for aircraft (van Paassen *et al.*, 2013). Manoeuvring constraints were largely related to the air domain, detailing calculation of an aircraft's manoeuvre space and how various factors affected capabilities to effect manoeuvres (Borst *et al.*, 2006; Ellerbroek *et al.*, 2011; Ellerbroek *et al.*, 2013a; Ellerbroek *et al.*, 2013b). Trajectory planning was largely in the air domain, relating to how planning how the aircraft could move given their constraints, including those of a spatial and manoeuvring nature (Leveson, 2000; van Marwijk *et al.*, 2011; van Paassen *et al.*, 2013).		
Temporal (45/90)	1. Temporal (41/80) 2. Speed (Temporal) (2/2) 3. Time-Independent (1/1)	Constraints of a temporal nature.
Temporal constraints included communication delays (Miller and Feigh, 2019), constrained time for decision making (Trentesaux, Moray and Tahon, 1998), and time constraints for achieving goals (van Marwijk *et al.*, 2011). Speed (Temporal) constraints included the fastest time in which aircraft actions could be carried out (Ellerbroek *et al.*, 2011) and the time taken to complete tasks (Jenkins *et al.*, 2008b). Time-Independent constraints were described by Vicente (1992) as relationships between variables at a single point in time, such as "Total Energy Stored" or "Mass Input Flow Rate".		
Actors and Agents (42/147)	1. Actors (11/14) 2. Intentional (11/46) 3. Communication (9/12)	Actors and agents within the system, both abstract and concrete, including their intentions. The definitions are that of sociotechnical systems theory, actors are the social components, and agents refer to an entity that may be either social or technical (Stanton *et al.*, 2017a).
The constraints on actors were split between describing assessing function allocation during CWA (Salmon *et al.*, 2007; Jenkins *et al.*, 2008a) and their roles in systems, such as how they would affect airborne separation displays (Ellerbroek *et al.*, 2011; Ellerbroek *et al.*, 2013b). Intentional constraints were used to generically describe		

Category (Files / References)	Top 3 Concepts (Files / References)	Description
Notes/Examples		
the constraints that actor intentions might impose on a given system and was often used with examples specialised examples to illustrate specific contextual constraints, such as organisation (Borst, Flach and Ellerbroek, 2015), military laws (Burns, Bisantz and Roth, 2004), and company objectives (Rasmussen, 1999). Closely related was “Intent” constraints, which in addition to mentioning the intent of actors (Ellerbroek *et al.*, 2011; Ellerbroek *et al.*, 2013b), also mentioned intent-driven domains (Bennett, Posey and Shattuck, 2008). Communication based constraints mentioned communication structure (Houghton *et al.*, 2015), and communication (social and technical) itself (van Paassen *et al.*, 2013).		
Natural Environment (32/88)	1. Environmental (12/36) 2. Terrain (11/14) 3. Weather (8/11)	Constraints associated with the natural environment, such as terrain or weather.
Environmental constraints were mentioned generally in most cases, such as in reference to in-depth models (Burns, Bisantz and Roth, 2004), and in some cases, specific examples were provided (Leveson, 2000; Calhoun *et al.*, 2018). Terrain constraints were largely in the domain of aviation, specifically collision avoidance (Borst *et al.*, 2006; Borst *et al.*, 2008; Ellerbroek *et al.*, 2013b; van Paassen *et al.*, 2013). However, other domains were also included, such as military command and control (Bennett, Posey and Shattuck, 2008; Bennett, 2017) and solar racing (Hilliard and Jamieson, 2008). Similar to environmental constraints, weather constraints were largely mentioned in general (Naikar and Sanderson, 1999; Ellerbroek *et al.*, 2013b; van Paassen *et al.*, 2013), and had specific examples, such as ice or rain affecting car tyre grip (McIlroy and Stanton, 2015a) and the priority of wind shear alerts for aircraft (Leveson, 2000).		
Resources (31/68)	1. Resources (11/18) 2. Financial (5/7) 3. Physical Resources (5/18)	Constraints relating to any type of resource available within the system, or that can be consumed by it.
Resource constraints refer to all resources available to a system in general (Bennett, Posey and Shattuck, 2008; Strauch, 2017) and constraints that these resources face, such as the life duration of tools (Trentesaux, Moray and Tahon, 1998). Financial constraints are identified explicitly (Jiancaro, Jamieson and Mihailidis, 2014), and are		

Category (Files / References)	Top 3 Concepts (Files / References)	Description
Notes/Examples		
mentioned with time constraints (Jamieson *et al.*, 2007; Salmon *et al.*, 2007; Read *et al.*, 2015a). Financial constraints during CWA analysis were also mentioned (Read *et al.*, 2015a; Naikar, 2017). Physical resources are also referred to as a collective, however, they refer to tangible resources and are typically tied to WDA (Naikar *et al.*, 2003; Lintern, 2006; Naikar, 2006; Read *et al.*, 2015a; Naikar, 2017).		
Government, Law, Legal, Regulatory, Political (27/64)	1. Regulations (11/15) 2. Government (7/9) 3. Rules of Engagement (4/8)	Constraints associated with the government, laws, and regulatory frameworks that govern the system.
Regulations governing specific work domains were mentioned across domains such as aviation (Borst, Flach and Ellerbroek, 2015), road transport (Salmon *et al.*, 2007), solar racing (Hilliard and Jamieson, 2007), and power grid management (Tran, Hilliard and Jamieson, 2017). Government constraints, mainly as laws passed by them, were also mentioned across domains, such as aviation (Borst, Flach and Ellerbroek, 2015), eco-driving (McIlroy and Stanton, 2015a), and military command and control (Bennett, Posey and Shattuck, 2008). Whilst not explicit in the literature, regulations appeared to be more specific to the work domain itself, whereas government/laws were general and imposed on the domain. Rules of Engagement constraints were military specific, and included specific uses such as training system procurement (Naikar and Sanderson, 1999), missions and culture (Burns, Bisantz and Roth, 2004), and constraints on action (Burns, Bryant and Chalmers, 2005).		
Safety (25/53)	1. Safety (20/44) *There were no clear second or third place constraints (same file and reference count)*	Constraints relating to the system safety and the conditions of this safety, including conditions that should not be broken, or should be maintained.
Safety constraints were typically related to a high-level goal of maintaining safety within the system across domains such as aviation (Leveson, 2000), eco-driving (McIlroy and Stanton, 2015a), and solar racing (Hilliard and Jamieson, 2007). Other constraints topping this category included “Danger” (Mantel, Hoppenot and Colle, 2012), “Failure Modes (Introduced in Experiment)” (Christoffersen, Hunter and Vicente, 1996), and “Fault Diagnosis (Described as Algorithm)” (Jamieson and Vicente, 2001).		

<table>
<tr><th>Category (Files / References)</th><th>Top 3 Concepts (Files / References)</th><th>Description</th></tr>
<tr><th colspan="3">Notes/Examples</th></tr>
<tr><td>Societal Factors (21/33)</td><td>1. Social (11/14)
2. Social Values (2/2)
3. Language (2/3)</td><td>Constraints created by explicit and implicit societal factors, such as culture or customs.</td></tr>
<tr><td colspan="3">Social constraints are typically presented generally (Burns, Bisantz and Roth, 2004; Jenkins et al., 2008a; Read et al., 2015a) in the text, and specified in WDAs presented. Burns, Bisantz and Roth (2004) mention that social constraints affect the use of force and authority when describing the WDA for a Canadian ship. Other CWA stages were also mentioned as having social constraints, such as SOCA (McIlroy and Stanton, 2011) or ConTA (Jenkins et al., 2008a). Social values constraints are general, and it was specified that they could affect military operations (Lintern, 2006) or command and control Burns, Bryant and Chalmers (2005). Again, the specific constraints are presented as WDAs. Finally, language, among other constraints, could limit humans through constraining and dominating them (Read et al., 2016; describing the “radical humanist paradigm”). Language constraints are also mentioned by Borst, Flach and Ellerbroek (2015) to argue that sensor failure is not necessarily a show-stopper as suggested in the literature (Vicente and Rasmussen, 1992; Vicente et al., 1996; Vicente, 2002); they argue that the failure of a key when typing written text, thus being subject to language constraints, would be easier to detect than with random typing. Applying this assertion to EID, they put forth that by making the constraints explicit, sensor failures, which violate them, will be more detectable.</td></tr>
<tr><td>Representation and Perception (20/39)</td><td>1. Display (4/5)
There were no clear second or third place constraints (same file and reference count)</td><td>Constraints that pertain to the aesthetics and representation of entities or data, and the perception of these factors. Nodes pertaining to the operator’s perceptual process in their mind were categorised into Operator Characteristics & Perceptions, and nodes that contributed to, or triggered, these processes were categorised here.</td></tr>
<tr><td colspan="3">Display constraints were those on the display overall (Talcott, Martinez and Stansifer, 2007; Bennett, Posey and Shattuck, 2008) and information that should/can be displayed (Effken, Kim and Shaw, 1997; Vicente and Ethier, 2000). Other constraints topping this category included the “Visualisation (Of Constraints)” (Bass, 2014), and</td></tr>
</table>

Category (Files / References)	**Top 3 Concepts (Files / References)**	**Description**	**Notes/Examples**
			"Representation" of signals, signs, and symbols (Flach, 2017). Signals are also referred to as "Time-Space Signals", signs are also referred to as "Perceptual Forms", and "Symbols" are defined as meaningful relational structures (Vicente and Rasmussen, 1992).
Control Enactors (17/35)	1. Control (7/9) 2. Control Strategies (2/3) 3. Manipulation (2/2)	Entities and processes associated with enacting command and control, both civilian and military. The category was not named Command and Control (C2), as it was decided during inter-rater reliability discussions that the term could suggest a military context.	Control constraints were referenced in general for EID principles and design (van Paassen *et al.*, 2018), tasks as part of Control Task Analysis (Naikar, Moylan and Pearce, 2006), and work domain constraints that affect control (Klomp *et al.*, 2016). Control strategy constraints talked about how EID visualisations would show work domain constraints to allow operators to implement effective control strategies (van Paassen *et al.*, 2013) or how a control enactor (an Air Traffic Controller in this instance) should respond to complex situations (Klomp *et al.*, 2016).
Velocity (14/64)	1. Speed (13/32) 2. Velocity (6/29)	Constraints involving velocity.	Speed constraints included aircraft (Ellerbroek *et al.*, 2013b), cars (McIlroy and Stanton, 2015a; Beanland *et al.*, 2018), and boats (Araujo, Davids and Serpa, 2005). Operating bounds for speeds were specified, as well as the relationship between speed and altitude for aircraft (Borst *et al.*, 2008; Ellerbroek *et al.*, 2011). As these constraints applied to vehicles in transit, they could be considered velocity constraints. Explicit velocity constraints were based in the air domain, namely how internal and external constraints were affect by, and affected, aircraft velocity (Ellerbroek *et al.*, 2011; Ellerbroek *et al.*, 2013b; Bass, 2014; van Paassen *et al.*, 2018).

3.3 Category Connection and Social Network Analysis

At this stage in the process, there was a list of constraint categories, but this did not provide any information on how they linked together. It was decided to conduct a Social Network Analysis (SNA) on outputs from NVivo to reveal this information. A non-directed adjacency matrix was generated using a matrix query in NVivo. A matrix query tabulates a specific value for each pair of categories, which in this case was the number of articles that were coded to both. As the matrix describes a graph, it is known as an adjacency matrix, which is non-directed as the order of connections does not matter. An example of this is the 'city to city' distance tables found in atlases. All taxonomy categories were added as headers for both rows and columns. Sub-categories were included with their parents, instead of being discreet, due to NVivo limitations. They were not removed as categories as they still provide prompts as to the types of constraints expected. Categories were deemed to be connected if they appeared in an article together. Whilst this might have slightly inflated relationship numbers in some cases, reducing the unit of analysis could have wrongly excluded valid relations. For some of the smaller categories, this could have completely removed their relations to another category. It was decided that links would be non-directed, as each manuscript was structured differently, meaning that a single measure of link direction could not be used. A representation of the output generated by NVivo is shown in Table 10, with the second, identical, half of the table (e.g., Safety → Spatial = Spatial → Safety) omitted for clarity.

SNA was performed on the matrix shown in Table 7 using the Applied Graph & Network Analysis (AGNA; Benta, 2003) tool version 2.1.1. Global network metrics, defined in Table 8, and nodal (category) metrics, defined in Table 10, were calculated. The definitions are adapted from Stanton and Roberts (2018) and Benta (2003), who use the commonly accepted definitions from AGNA. This was to facilitate an understanding of the entire network, in addition to understanding the specific statistics of each category. This revealed information about the network to understand its composition, in addition to how it would affect the taxonomy usage. The global network metrics, calculated from Table 7 are presented in Table 9, and the node-specific metrics are presented in Table 11.

Table 7 – One-half of the non-directed adjacency matrix generated by NVivo. The other half is the same and has been omitted for clarity.

	Analysis & Design of System	Control Enactors	Decision Making, Goals, and Action	Energy, Mass, Forces	Government, Law, Legal, Regulatory, Political	Measurable Parameters	Natural Environment	Operator Characteristics & Cognition	Organisational	Representation and Perception	Resources	Safety	Societal Factors	Spatial	System	Technology	Temporal	Velocity	Work Domain/Environment
Actors and Agents	6	0	12	2	2	3	1	7	3	0	0	1	2	7	5	4	3	4	6
Analysis & Design of System		2	18	6	1	7	1	9	5	1	2	0	1	4	18	2	2	1	17
Control Enactors			2	2	0	2	0	1	0	0	0	1	0	2	4	2	0	0	2
Decision Making, Goals, and Action				7	6	14	3	22	10	2	3	4	1	17	22	8	6	1	29
Energy, Mass, Forces					1	16	2	6	2	0	1	2	0	11	16	3	0	2	24
Government, Law, Legal, Regulatory, Political						1	1	1	2	0	0	0	2	0	2	2	0	0	1
Measurable Parameters							3	14	1	5	5	6	1	13	11	10	3	4	17
Natural Environment								1	0	1	1	1	0	7	8	1	1	2	8
Operator Characteristics & Cognition									7	4	3	5	1	7	17	12	3	1	25
Organisational										0	2	1	2	1	9	4	1	0	10
Representation and Perception											1	0	0	3	3	1	0	1	1
Resources												0	0	1	0	2	1	0	2
Safety													0	6	4	1	0	1	4
Societal Factors														0	1	1	0	0	1
Spatial															16	6	10	10	12
System																12	2	2	26
Technology																	1	2	9
Temporal																		1	3
Velocity																			0

Table 8 – Definition of global SNA metrics

Metric	Definition
Nodes	The number of categories of constraint present in the taxonomy.
Edges	The total number of connections between the categories.
Density	Number of connections between categories observed, as a fraction of the total possible connections. Used to determine how connected the categories are compared to the maximum possible number of connections (every category is connected to all others). A higher density indicates that more connections link the categories.
Diameter	Maximum number of connections followed to traverse the network. Used to determine how many connections might be required to move from one category to another, if following a linear path through the connections.

Table 9 – Global network metrics

Metric	Value
Nodes	20
Edges – Non-Directed	152
Density	0.80
Cohesion	0.80
Diameter	2.00

Table 10 – Definition of nodal SNA metrics

Metric	Definition
Emission	Number of outgoing connections from a category to other categories. Used to determine how well a category is connected.
Sociometric Status (Sociometric)	Number of emissions from a category, relative to the amount of network nodes. Used to identify the most important categories in the network.
Centrality (Bavelas-Leavitt)	The sum of all distances in the network, divided by the sum of all distances to and from each node. Used to identify the central categories that would serve as possible starting points for analysis.
Farness	The sum of the shortest geodesic paths to each category in the network. Used to determine how close the category is to other categories.
Betweenness	The number of times a category is in the shortest geodesic paths between other categories, divided by the total number of paths. Used to determine how often this category would appear in the connections between other categories.

Table 11 – Nodal metrics for each category within the taxonomy

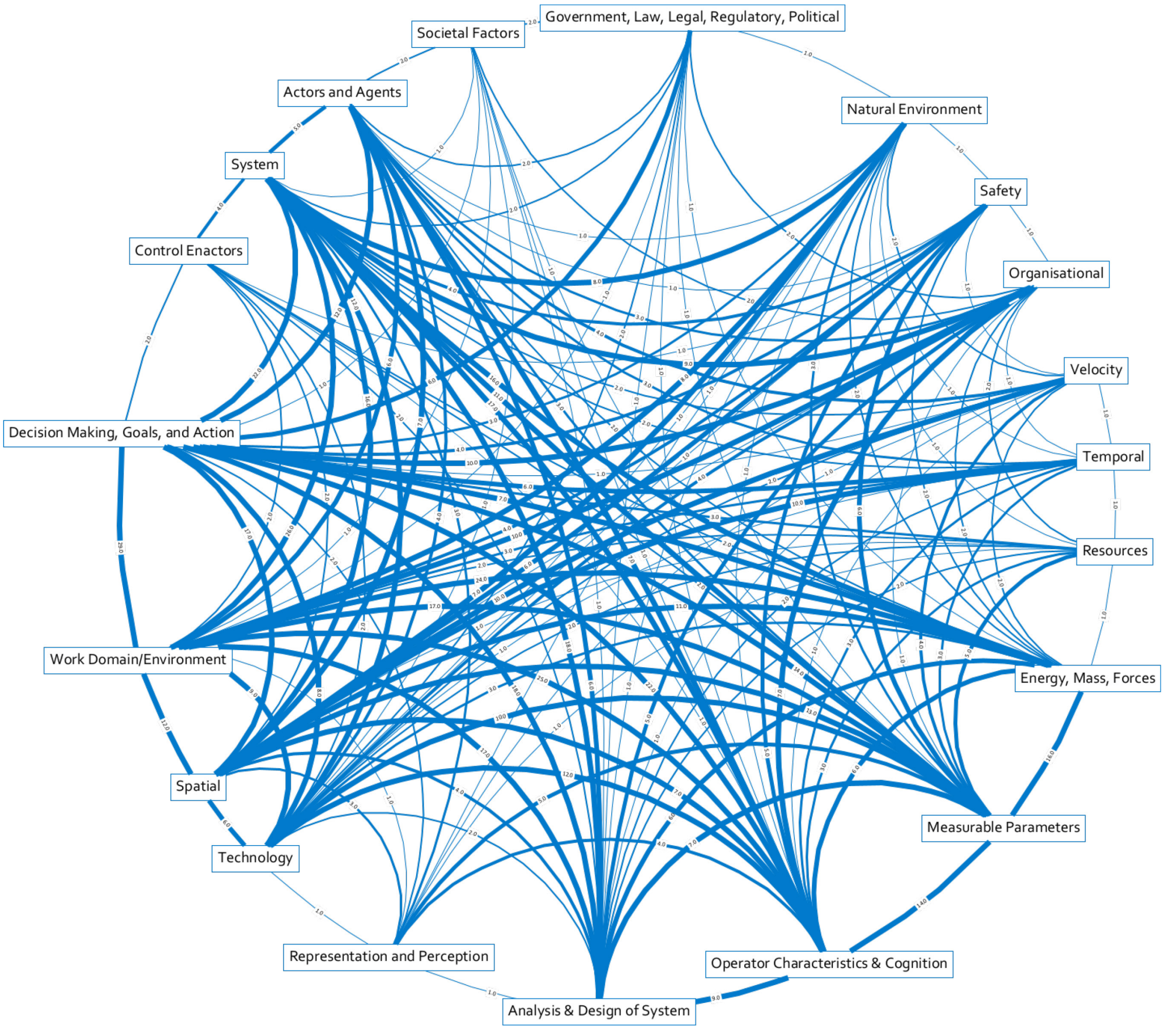
	Emissions	Status	Centrality	Farness	Betweenness
Societal Factors	13	1.37	8.14	28	0.00
Control Enactors	20	2.11	8.14	28	0.15
Government, Law, Legal, Regulatory, Political	22	2.32	8.77	26	0.57
Representation and Perception	23	2.42	8.44	27	0.52
Resources	24	2.53	8.77	26	0.88
Velocity	32	3.37	9.12	25	1.28
Safety	37	3.89	9.12	25	1.30
Temporal	37	3.89	9.12	25	1.03
Natural Environment	42	4.42	10.36	22	3.48
Organisational	60	6.32	9.91	23	3.04
Actors and Agents	68	7.16	10.36	22	3.17
Technology	83	8.74	12.00	19	8.10
Analysis & Design of System	103	10.84	11.40	20	6.87
Energy, Mass, Forces	103	10.84	10.36	22	3.55
Spatial	133	14.00	10.86	21	4.32
Measurable Parameters	136	14.32	12.00	19	8.10
Operator Characteristics & Cognition	146	15.37	12.00	19	8.10
System	178	18.74	11.40	20	6.58
Decision Making, Goals, and Action	187	19.68	12.00	19	8.10
Work Domain/Environment	197	20.74	11.40	20	6.84

3.4 Visualising the Constraint Taxonomy

Calculating a matrix of links between categories provided information on how each category was connected, and the strength of these connections. By conducting SNA on these connections and interpreting the results, it was possible to build and understanding of how the constraint taxonomy should be compiled and used. This section will detail this process and how to use the finished taxonomy.

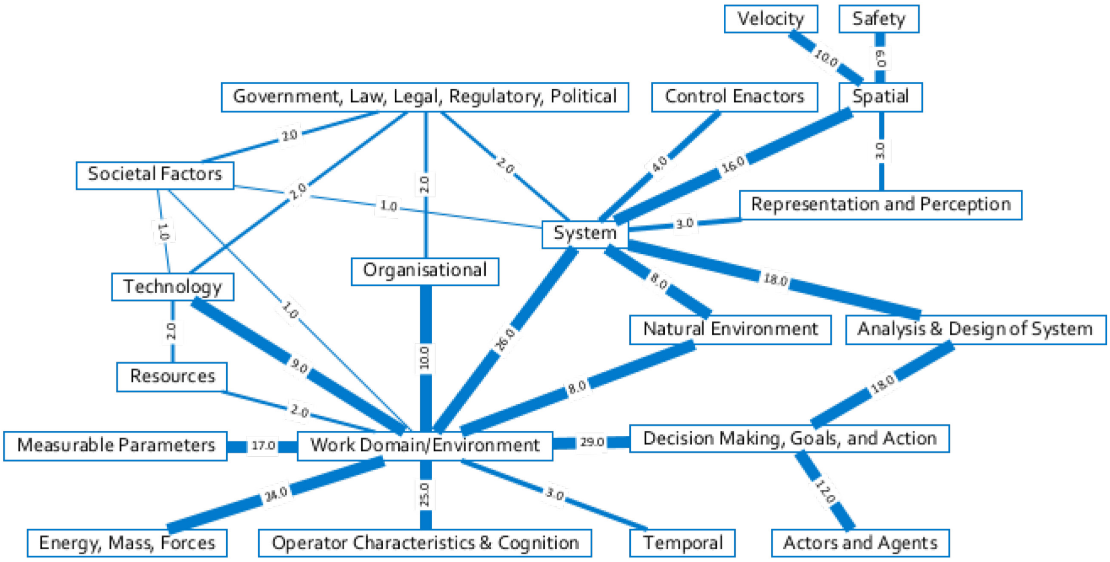
A visual representation of the taxonomy was generated by importing the matrix into yWorks yEd version 3.22. The SNA adjacency matrix was modified by removing symmetrical edges, and those with a zero cardinality, as these were not automatically excluded by yEd and would generate incorrect edges if not omitted. Built-in functionality to read SNA adjacency matrices was used to open the modified adjacency matrices. The ‘Adopt’ functionality was used to automatically enter all data, which was rendered using default aesthetics (shapes, connectors, and routing options). Custom properties were then used to change the styling of the nodes (colour, font) and edges (end styles, thickness, labels) depending on the data. Automatic layouts were then applied, with subsequent manual changes were made for aesthetics. The generated visualisation is shown in Figure 8. The connectivity identified in the SNA is immediately apparent and makes the diagram hard to interpret; most categories connect to most other categories, with varying emission counts.

Figure 8 – Visual representation of connectivity between constraint categories with all edges in the taxonomy shown



While an accurate representation of links present within the literature, Figure 8 was not a human-readable representation of the taxonomy, limiting its usability outside of demonstrating the complexity of the 'big picture'. To make the visual representation easier to interpret, another adjacency matrix was created, taking only the top-rated edge (or edges if multiple edges had the highest cardinality) and rendered using the above process. Figure 9 shows this final taxonomy visualisation. As can be seen, it is vastly easier to understand, and central nodes can be identified; "Work Domain/Environment" and "System" connect to most nodes, including each other, as expected from their sociometric status, highlighting their importance within the network.

Figure 9 – Visual representation of connectivity between constraint categories with only the top edge shown



3.5 Using the Constraint Taxonomy

With the constraint taxonomy visualised, it becomes possible to propose how to utilise it during the analysis process. The first possibility is to systematically work through the list of ordered categories provided in Table 12 as a checklist, eliciting information from the source material(s) and/or subject matter experts for each one using the detail in Table 6. This process should involve using the provided descriptions and top individual constraints as prompts where possible, see Figure 4 as an example for the first three categories. This approach would likely be suitable for performing analyses on new domains, or one where the analysis will cover a large sociotechnical system, both requiring in-depth analysis. It would also be useful for providing sample answers during interviews to the prompts provided across the CWA literature (e.g., Naikar, Hopcroft and Moylan, 2005; Jenkins *et al.*, 2009; Read *et al.*, 2016), to illustrate to the interviewee what kind of constraints and categories would be suitable to include.

Another possibility is to use the visualisation to guide the analysis using a guided checklist approach, see Figure 11. A practitioner would start at the Work Domain/Environment category, having the highest sociometric status (see section 3.6.2.2). They would still use Table 6 to elicit information. While doing so, they should be cognisant of discovering constraints that belong to other categories and seek to traverse the connections to those categories to discover related constraints, jumping to the relevant item in the taxonomy, instead of working through. Continuing to traverse the taxonomy and exploring the constraints, should provide a structured process to acquire constraints relevant to the system being analysed.

Table 12 – A checklist for the taxonomy of constraints

Category	**Top Constraints to Query**
☐ Work Domain / Environment	☐ Work Domain ☐ Physical ☐ Functional
☐ System	☐ System ☐ Purpose ☐ Design (Of System)
☐ Limits, Capacity, Restrictions	☐ Locomotion Capabilities ☐ Performance Limitations (Aircraft) ☐ Operating Capacity
☐ Entities	☐ Traffic ☐ Aircraft ☐ Weapon
☐ Decision Making, Goals, and Action	☐ Goal-Relevant ☐ Strategies (For Activity Execution) ☐ Process
☐ Operator Characteristics & Cognition	☐ Cognitive ☐ Behaviour ☐ Decision Making
☐ Measurable Parameters	☐ Information ☐ Relationships (Between Variables) ☐ Variables
☐ Energy, Mass, Forces	☐ Energy ☐ Temperature ☐ Mass
☐ Laws of Nature	☐ Causal (Laws of Nature) ☐ Inherent ☐ Invariant Constraints
☐ Analysis & Design of System	☐ Work Domain Analysis & Abstraction Hierarchy ☐ High Level ☐ Cognitive Work Analysis
☐ Analysis Considerations	☐ Access to End-Users and subject matter experts ☐ Trial order
☐ Technology	☐ Technology/ Technical ☐ Design (Of Interface)/ Interface ☐ Sensor
☐ Organisational	☐ Organisational (including “Social Organisational”)

Category	Top Constraints to Query
	☐ "Operations" (including "Operating" and "Operational") ☐ Work
☐ Spatial	☐ Spatial ☐ Manoeuvring ☐ Trajectory Planning
☐ Temporal	☐ Temporal ☐ Speed (Temporal) ☐ Time-Independent
☐ Actors and Agents	☐ Actors ☐ Intentional ☐ Communication
☐ Natural Environment	☐ Environmental ☐ Terrain ☐ Weather
☐ Resources	☐ Resources ☐ Financial ☐ Physical Resources
☐ Government, Law, Legal, Regulatory, Political	☐ Regulations ☐ Government ☐ Rules of Engagement
☐ Safety	☐ Safety
☐ Societal Factors	☐ Social ☐ Social Values ☐ Language
☐ Representation and Perception	☐ Display
☐ Control Enactors	☐ Control ☐ Control Strategies ☐ Manipulation
☐ Velocity	☐ Speed ☐ Velocity

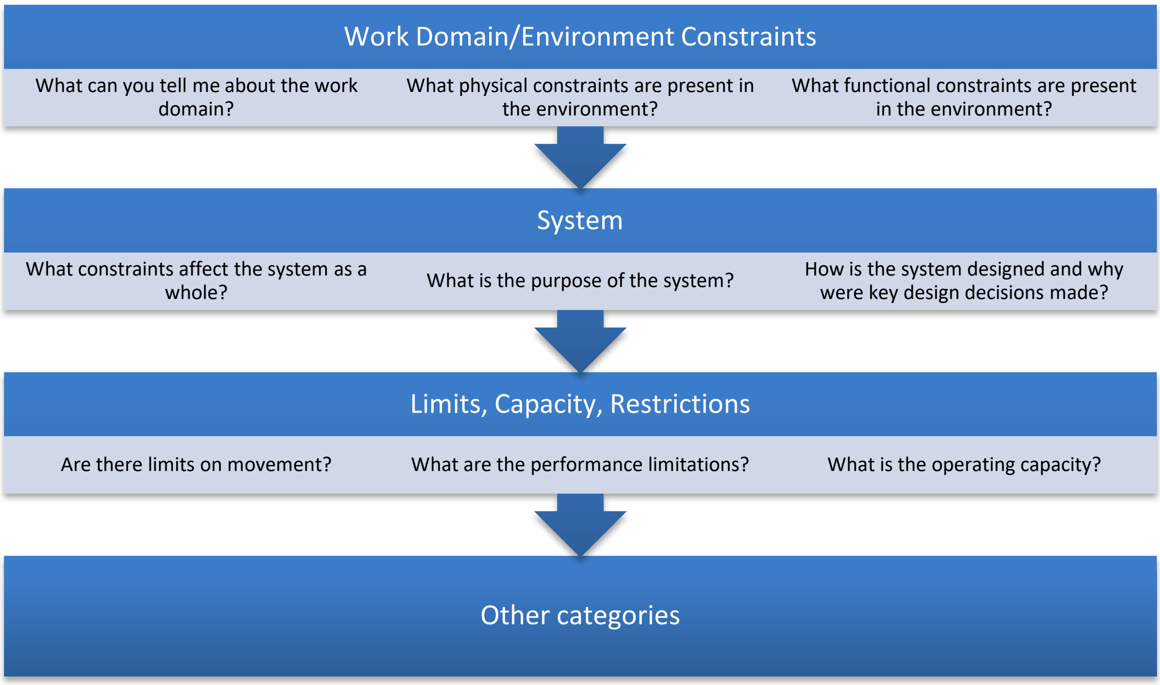


Figure 10 – A checklist approach to using the constraint taxonomy

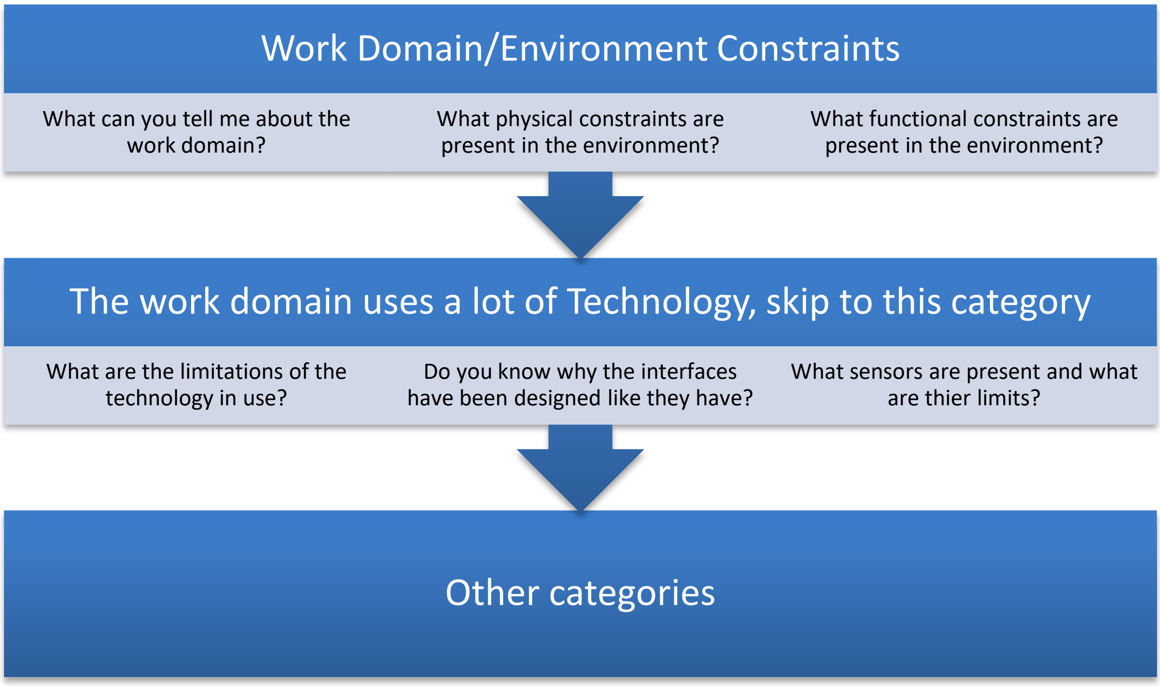


Figure 11 – A guided checklist approach to using the constraint taxonomy

A traversal approach could also be used to imply where constraints might be missing from an analysis, see Figure 12. Again, a practitioner would start at the Work Domain/Environment category, or possibly the category that matches the most constraints identified in previous iterations for ongoing analyses. They would then assess whether they have identified constraints that belong to connected categories. If constraints have not been identified, or only a small number are present, this could suggest that more are present based on the links exhibited in the literature.

The emissions of each category provide a mechanism to order the exploration of neighbours. If enough constraints have been identified in a neighbouring category, then analysis could iteratively move across the taxonomy to address all neighbouring categories.

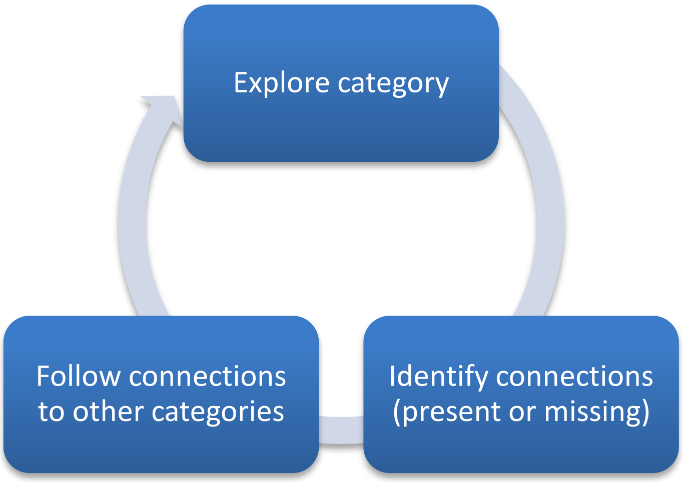


Figure 12 – A traversal approach to using the constraint taxonomy

Finally, the above use cases have been focused on exploring categories where the most constraints would likely be found. An alternative approach could be to start at categories with low sociometric status or ones on the edge of the taxonomy. This could be useful to practitioners with an established analysis who are seeking to explore less-obvious constraints, which could be important for ensuring that as many novel and/or unanticipated situations are accounted for in the resulting design(s).

In summary, there are two ways to use the taxonomy, as a systematic checklist, or as a traversal map to identify connections between constraints. The proposed use cases above are suggestive, and practitioners are encouraged to use the taxonomy as appropriate within the context of their specific analysis.

3.6 Discussion

This chapter aimed to provide a constraint taxonomy that could be used for analysing systems when performing CWA and/or EID. The taxonomy has been constructed throughout this chapter, synthesised from a systematic literature review conducted on the topic of EID. Journal articles from a search of multiple search engines were loaded into NVivo and all recorded constraints were coded. These constraints were subsequently grouped using an iterative bottom-up approach to create the categories of constraints that form the taxonomy. SNA was conducted to analyse how the categories connected, and to inform the use cases, such as starting at the Work Domain/Environment category. The completed taxonomy takes the form of a tabular checklist and network visualization, both of which can be used as required based on the proposed use cases in the previous section. This section presents a discussion of the taxonomy and the analysis used to construct it.

3.6.1 Constraint Categories

The generated categories are aligned with the topics a practitioner would expect to encounter in EID, indicating the suitability of the results to support systems analysis and design processes. All categories were expected from the author's previous reading and experience with EID. For example, the Technology and Actors and Agents categories were expected, given that complex sociotechnical systems were a focus. This suggests that while this review concentrated on mentions of "constrain*" words and their context, the theoretical underpinnings of EID were successfully captured within the taxonomy. This is readily apparent in Figure 9, in which Work Domain/Environment and System are central highly connected constraint categories. For Work Domain/Environment, the categories' importance is thought to be linked to CWA, which has a pivotal role in the systems analysis stages of EID, appearing across the literature (McIlroy and Stanton, 2015b). As the system is the focal point of CWA and EID, it follows that the System category should also be a major category, and also as it exhibits such strong connectivity to Work Domain/Environment. Other prevalent aspects of EID also manifested themselves as categories, such as Operator Characteristics & Cognition (Worker Competencies Analysis and SRK Taxonomy) and Representation and Perception (e.g., direct perception of environmental variables as a goal of EID). Aligning closely with constraints identified by expert practitioners, and with the community consensus of EID, is a strong indicator of the robustness of the taxonomy's contribution.

The categories also help affirm previous literature that seeks to provide a list of constraints, or questions to understand what constraints might exist within a system, such as the work of Rasmussen, Pejtersen and Goodstein (1994), Naikar, Hopcroft and Moylan (2005), Jenkins *et al.* (2009), Naikar (2013), and Stanton *et al.* (2017a). This aligns with the aim of this chapter to establish a list of constraint categories found across the literature as part of the taxonomy, serving as a reference for practitioners to ensure that EID applications cover all types of constraint appropriately. This is achieved by clearly demonstrating the presence of constraints mentioned in the theory across the literature. For example, Naikar, Hopcroft and Moylan (2005), citing Rasmussen, Pejtersen and Goodstein (1994), present a table (table 12, p. 69) of constraint types across the domain continuum, split into five categories. This table contains sources of regularity for each domain, driven by types of constraints associated with them. These sources of regularity appear in the conducted analysis, such as the causal constrains for automated systems governed by the laws of nature, or intentional constraints based on a Decision Making, Goals, and Action, and preferences for systems used by autonomous casual users. It is important to note that the contributions of domain-leading authors were not under scrutiny; rather this review sought to fill a gap in the literature by systematically analysing constraints across the EID literature to explicitly demonstrate their existence, rather than utilising knowledge gained from extensive experience with system design. It is also important to emphasise that correlation does not equal causation, the categories were not derived from these works, rather informed by the body of literature. In doing so, this chapter's contribution to knowledge is bolstered by incorporating the practical findings of various practitioners into the taxonomy, which might not have been elicited otherwise.

The taxonomy exhibits complexity in the form of the number of categories and the connectivity between these categories, a reflection of the EID literature. While it is accepted that twenty top-level categories could be unwieldy, it must be considered that these constraints pertain to complex sociotechnical systems, and that this will be commensurate with complexity in a taxonomy based on literature based on them. This aligns with the proposition of Borst, Flach and Ellerbroek (2015), who argue that Ashby's Law of Requisite Variety (Ashby, 1956) is fundamental to EID's consideration of ecology (Vicente, 1991a), necessitating that effective interfaces be as complex as the domain under consideration. Table 6 demonstrated that each category can be clearly defined and is present within the literature with examples, providing reasoning for their inclusion. It could be possible to reduce the number of categories to simplify the taxonomy by merging similar categories. However, given the encompassing definitions of some categories, this may make the resultant categories too broad for meaningful analysis. Conversely, certain categories are smaller and can also exhibit part-whole relationships with others. Future work could simplify the taxonomy by utilising techniques from Work Domain Analysis, such as the Abstraction Decomposition Space, which decomposes the Abstraction Hierarchy into different component levels (e.g., system, sub-system, component). This would add extra dimensions to taxonomy with fewer categories at each level. Despite these changes, it would be expected that it would remain heavily connected. Again, the nature of complex sociotechnical systems is present, with constraints affecting each other throughout the literature; there may be discussion about which category constraints are placed in but is inescapable that categories will be highly connected. To alleviate this for future iterations, different methods of representation should be concentrated on, to ensure that the links are understandable (the complexity of Figure 8 compared to the understandability of Figure 9 for example).

3.6.2 Social Network Analysis

3.6.2.1 Global Network Metrics

The global network metrics are indicative of how numerous and complex constraints are in modern sociotechnical systems. 904 constraint types were categorised into twenty top-level nodes for the analysis. The IRR results indicate that whilst there was a level of agreement between the original coder and the secondary rater, there was room for improvement to achieve even better agreement. However, the categories themselves were deemed to be encompassing, as no constraints were categorised outside of the provided categories. This suggests that the categories themselves are suitable for the taxonomy. During coding discussions and discussions with the secondary rater, it became apparent that whilst some nodes could be more appropriately categorised, a portion of codes could fit into multiple categories. For example, 'operating niche' *("… take into account constraints … from the particular ecological niche in which the robot is to operate (home, underwater, forest, rubbles, etc.)."* (Mantel, Hoppenot and Colle, 2012)) could be in the Natural Environment or Work Domain/Environment categories, depending on the perception of the coder. Both could be considered correct, an operating niche is both the Natural Environment a system could be deployed to, as well as the Work domain/Environment. Due to the limitations of NVivo, it was not possible to categorise a node into more than one category. Future work could seek to use different software to categorise nodes into more than one category to account for their complex nature, potentially improving inter-rater reliability scores. Another method to improve the inter-rater reliability scores would be to make the categories more specific and defined. However, it is believed that the complexity of some constraints might make this idea difficult; there are already twenty categories and there is no guarantee that creating more constraints, adding complexity to the taxonomy, might result in mutually exclusive coding.

The taxonomy exhibits high connectivity, as demonstrated by the number of edges and density values. The extent of the connections can be seen from Table 7. Only 39 non-directed edges (76 directed edges) are absent (equal to 0) from a fully connected network. This high connectivity demonstrates how constraints are often inter-related and cannot be considered in isolation. This affirms the final proposed benefit of this chapter, which was to understand how constraint categories were connected, enabling practitioners to use the connected constraint taxonomy to further enhance their analysis using the taxonomy over utilising the checklist alone. It would be pertinent to consider these identified relationships during the construction of an Abstraction Hierarchy (and other stages of Cognitive Work Analysis). For example, the Decision Making, Goals, and Action and Work Domain/Environment categories are highly connected, meaning that constraints in these categories are very likely to be connected. If interviews with subject matter experts describe several constraints relating to how they make decisions and set goals, it would be sensible to elicit what components of the work domain affect these constraints and proceed using connections identified in the taxonomy.

3.6.2.2 **Nodal Metrics**

The Work Domain/Environment, Decision Making, Goals, and Action, and System categories had the highest emissions and sociometric status. This is to be expected as these constraints define and bound work systems. Thus, the presence of these types of constraints would be expected throughout the literature, connected to application-specific constraints. The categories exhibiting the next highest emissions were Operator Characteristics and Cognition, Measurable Parameters, and Spatial. While not as all-encompassing as the top categories, they are pervasive throughout the application of EID and the literature. It is a core tenet of EID that an operator's capability should be considered, representing spatial layout is key to representing the work domain, and understanding a system's measurable parameters allows for them to be represented in the EID to show the status of the system.

In contrast, the Societal Factors, Control Enactors, and Government, Law, Legal, Regulatory, Political categories had the lowest emissions. For Societal Factors and Government, Law, Legal, Regulatory, Political, these categories are widespread factors that undoubtably will influence sociotechnical systems. However, their effect might not be direct, nor relevant at the scope of the interface. For example, consider a legal requirement for a minimum screen size. A constraint on screen size might be mentioned, but would be categorised as a technological constraint, instead of the legal constraint it is derived from. Similarly, Control Enactors might not be explicitly mentioned as such, instead being referred to via Actors and Agents. Applied to the usage of the taxonomy, this suggests that explorations in these categories do not need to be as exhaustive as other categories, unless dictated by a specific domain where these categories are known to be prominent.

The categories Decision Making, Goals, and Action, Operator Characteristic and Cognition, Measurable Parameters, and Technology had the highest centrality/farness/betweenness. This suggests that they were important concepts throughout the literature, that were central to defining the constraints for a work system. Supporting Decision Making, Goals, and Action, Operator Characteristics and Cognition is a key aspect of EID, and thus it makes sense that other types of constraints would be connected. As the work systems in question are sociotechnical systems, it is also expected that constraints featuring the social (Operator Characteristic and Cognition) and technical (Technology) aspects would appear as central categories. Finally, it is believed that the Measurable Parameters of a system possessed high centrality due to its key role of providing information about the status and functioning of a system, which is in keeping with EID and its goal of showing both physical and functional information. Applied to the taxonomy, this would suggest that practitioners should seek to work towards these categories if they have not already been explored.

By contrast, the categories Societal Factors and Control Enactors were the least central nodes, and it is believed that this is for the same reason as their low connectivity (emissions); whilst pervasive in a greater context, their effect on constraining the work system may be achieved through other constraint categories that have more effect in a system context. Despite the network's high density and small diameter, there was a high range of farness (9), with the most central categories being connected to other categories by only one edge, and the least connected categories being connected by an average of 1.47 edges. This strongly affects the betweenness of each category. For the categories with the least farness, possessing connections to all other nodes of 1 edge, there is a high probability that they will be in various shortest paths and will be considered when using the taxonomy. Less central categories will likely be at the end of a route, connected to only specific categories, increasing their farness, as they can only be accessed through more edges. Consequently, practitioners may have to take explicit steps to consider these categories if they are not organically arrived at during the analysis.

3.7 Conclusion

A key component of EID is the constraints it represents, framing the action state-space in which operators are free to utilise to achieve their planned goals. Constraints are generally well understood across the literature, but literature did not exist that provided a systematic overview of those identified from actual applications of CWA. This chapter addressed this by performing a systematic literature review using the search term "Ecological Interface Design" to create a taxonomy of constraints. Constraints were identified, classifying them into twenty categories, and Social Network Analysis (SNA) was used to identify relations between the categories to build a constraint taxonomy. The taxonomy was designed to support practitioners in discovering as many constraints as possible during their analyses, as they are key to designing an EID that is fit for purpose.

The main contribution to knowledge of this chapter was the creation of the twenty categories of the taxonomy, which can be a foundation for new analyses and a checklist for ongoing/completed analyses. A further contribution was using SNA to demonstrate how these constraint categories were connected to enable practitioners to follow these connections, especially if they are eliciting several constraints in one category, but not from linked categories, which suggests that these should be explored. It was demonstrated that many categories are highly inter-connected. Whilst the number of constraints in each category varied, they were each derived from the literature, and clearly show the types of constraints that are found in contemporary sociotechnical systems. The taxonomy's connectivity reflects this, showing that each category of constraint could not be considered in isolation and that practitioners should consider how categories are connected.

The benefits make the taxonomy a strong contribution to the literature, providing practitioners with a literature-underpinned understanding of constraints that can be applied at different stages of their EID applications. At the systems analysis stage, the understanding of constraints will facilitate a more in-depth assessment of the system and its capabilities. For complex systems such as Sonar and TMA, this is vital, as constraints might not immediately be apparent, being abstracted away at the HMI level. At the design stage, it will enable consideration of how constraints should be represented, based on their prevalence and connections. Should constraints be heavily connected, it would be appropriate to display them together, further capitalising on the operator's innate psychomotor capabilities. Finally, at the experimentation stage, it could permit an insight into whether a constraint's prevalence or connectivity has any bearing on operator performance.

As the underlying theory and approach to analysis have been covered at this point, the rest of the thesis will detail the specific application of EID to Sonar and TMA. Chapter 4 presents a review of how submarine control rooms should operate as a pre-cursor to the CWA completed in subsequent chapters. This was needed as the operation of submarine control rooms is not generally well known, meaning that they had to be understood before working with subject matter experts on a redesign of Sonar and TMA.

Chapter 4 How submarine control rooms operate, and the Case for Ecological Interface Design

Previous chapters have concentrated on introducing the theoretical knowledge required for this thesis (Chapter 2) and presenting a taxonomy of constraints (Chapter 3). This chapter marks the start of the analysis process conducted to derive an Ecological Interface Design (EID) for Sonar and Target Motion Analysis (TMA). These stations were chosen due to their prevalence within the control room, both in terms of their role in ensuring ownship safety, and their communications forming a significant proportion of command team activities (Stanton and Roberts, 2018). The chapter presents a consolidated overview of how submarine control room sociotechnical systems operate, providing context to the work of Sonar and TMA. Their work is briefly explored to set the foundation for CWA chapters on each role (Chapter 5 for Sonar, and Chapter 6 for TMA). Challenges facing the control room will also be detailed, highlighting why change is needed to ensure that adequate capability is maintained. Finally, the case for applying EID to the redesign of Sonar and TMA is presented, showing that there is synergy between the challenges faced by the control room, and those that EID aims to address.

4.1 The Submarine Control Room Sociotechnical System

To understand how the Sonar and TMA interfaces could be improved, it is also important to understand the operational context in which they are utilised. As Sonar and TMA are part of the wider control room sociotechnical system their operation is affected by, and affects, the control room. The following section will detail broader control room operation and the command team's work. The subsequent section will discuss potential shortcomings with current HMI designs and offer Ecological Interface Design (EID) as a paradigm that could optimise the next generation of HMIs.

4.1.1 Social: Command Team

The structure of a command team varies, although typically it is formed of operators and officers, and is led by a senior officer (a person with managerial responsibilities). When creating a tactical picture this will be the Officer of the Watch (OOW), although other officers may lead the control room depending on seniority (e.g., if the captain is present) and context (e.g.: a specialised activity). A tactical picture is an overview of the submarines' environment, including the perceived positions of contacts. A contact is a vessel or object that has been detected by sensors such as Sonar, radar or periscope, and is being analysed in the control room (Maranda, 2008; Stanton, 2014; Wang, 2016). An example subsection of a Royal Navy command team structure that focuses on tactical picture generation, adapted from Roberts, Stanton and Fay (2017b), is described in Table 13 and Figure 13 respectively. Other navies might have different specifics but are likely to have commonalities with the structure described. Operators will collect and process information under the guidance of senior operators/officers, who will use the information for strategic as well as tactical decision making (Roberts & Stanton, 2018).

Table 13 – Roles present in a typical Royal Navy picture compilation team described by Roberts, Stanton and Fay (2017b)

Role	Description
Officer of the Watch (OOW)	Responsible for guiding tactical picture generation and guiding submarine activity to best complete mission objectives whilst maintaining the three tenets of submarine operation (see section 4.1.1.1).
Operations Officer (OPSO)	Generation of the tactical picture for OOW, incorporating information from all available sensors. The OPSO supervises TMA1/2 (below) and is responsible for quality checking their work.
Sonar Controller (SoC)	Assimilation and interpretation of sonar data being received in the sound room. This data is then used to generate the tactical picture by OPSO and TMA1/2. The SoC supervises SoP1/2 and is responsible for quality checking their work.
Sonar Operator 1/2 (SoP 1/2)	Detection, Classification, Localisation and Tracking (DCLT) of surrounding vessels, using sonar to do so.
Target Motion Analysis Operator 1/2 (TMA 1/2)	Compilation of environmental information into a tactical picture.
Periscope Operator (Peri)	Operation of the periscope mast to provide visuals of the surrounding environment.
Ship Control (ShC)	Enacts changes to submarine parameters, ensuring safety whilst doing so.

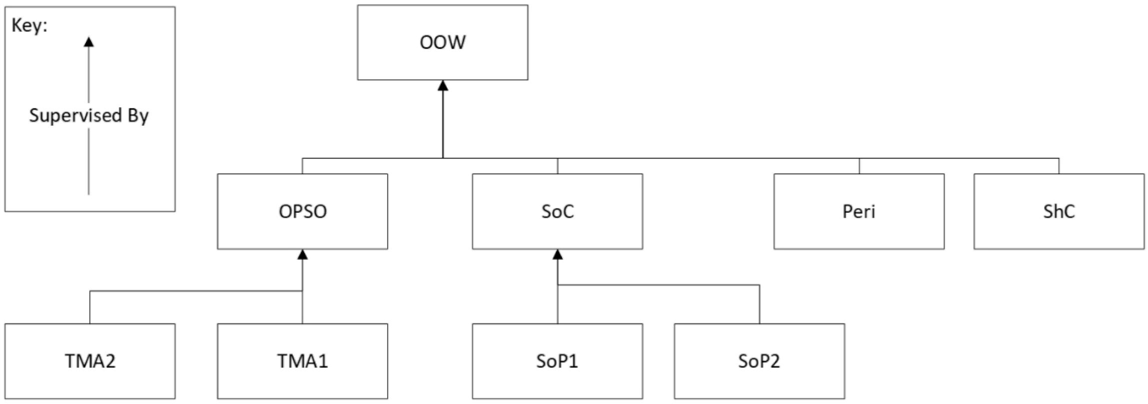


Figure 13 – Structure of the picture compilation team described by Roberts, Stanton and Fay (2017b)

4.1.1.1 **Objectives**

Command teams have three main objectives: remain safe, remain undetected, and complete mission objectives (Mack, 2003; Mewett, 2014; Fay, Stanton and Roberts, 2017). Remaining safe is considered to be the most important objective (Mack, 2003), as the consequences of a poorly maintained tactical picture can have profound consequences, such as the USS Greeneville crash or the Royal Navy submarine incident previously discussed (National Transportation Safety Board, 2001; Drumheller and Benoit, 2004; Marine Accident Investigation Branch, 2016). With submarines frequently operating in more complex environments, such as littoral waters (Duryea, Lindstrom and Sayegh, 2008; Bateman, 2011), and the global marine environment becoming busier (Davies, 2013), increasing the number of potential vessels a platform could encounter, safety must be continuously considered. The team must constantly assess the submarine's position in the water, how this relates to its surroundings, and current threat vulnerability.

Command teams must gain information about their environment to inform possible courses of action, and it is vital that the combat system provides the capability as well as flexibility to support this. This is especially important given that submarines operate with significant amounts of uncertainty (Brolese, 2005; Hunter, Hazen and Randall, 2014), meaning that data must be continuously evaluated to ensure that it is valid and reliable. Mission objectives are fluid and can vary, but routine operations typically include (Stanton and Roberts, 2018) dived tracking of entities (following something while underwater to remain undetected), returning to periscope depth (safely moving to near the surface from underwater), or inshore operations (conducting operations in littoral waters); a command team must be proficient in all mission types to ensure safe operations in a contemporary political climate (Bateman, 2011).

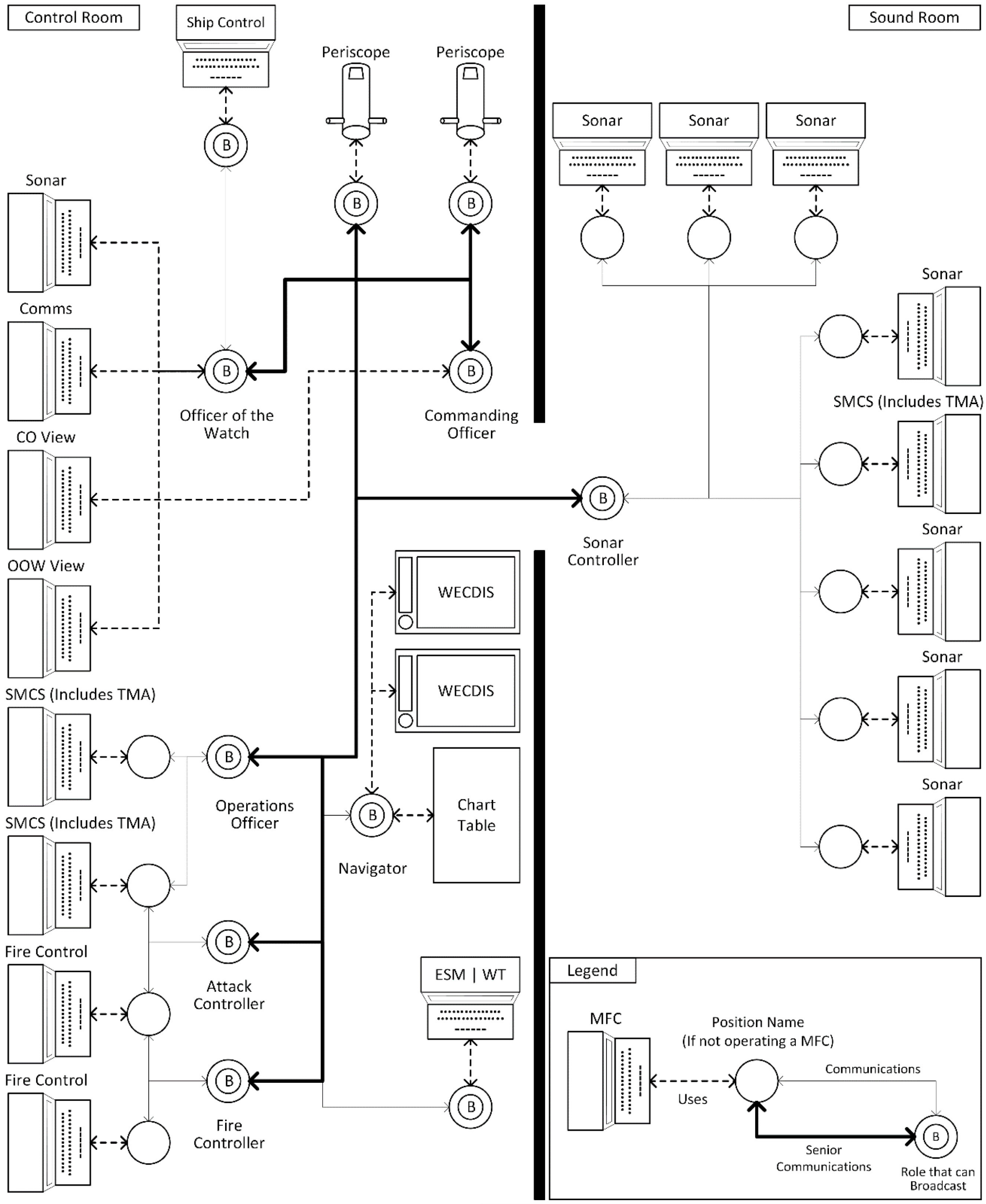
4.1.1.2 **Control Room Location and Layout**

Control rooms are typically located amidships (centre-ship), directly underneath the conning tower. This placement is a requirement of hull penetrating periscopes, which impose physical restrictions on a control room's location (Duryea, Lindstrom and Sayegh, 2008). As the periscope is a solid tube that breaches the hull, a control room has to be directly underneath, as with the USS Nautilus (Ven, 1956). Newer submarines use optronics masts, which transfer image data via fibre optic links, and as such do not require hull penetration (Hamburger, Miskimens and Truver, 2011). Consequently, a control room could potentially be located anywhere on the submarine. This occurred on the USS Virginia (SSN 774), whereby the control room was moved down a deck, which created more space, allowing sonar operators to occupy the same space as other watch-standers (Hamburger, Miskimens and Truver, 2011).

Hamburger, Miskimens and Truver (2011) noted that a control room relocation appeared to improve Situation Awareness (SA) and communications in Virginia class submarines by allowing all watch-standers to occupy the same space. This could potentially be a result of communication bottlenecks being removed, identified by Roberts and Stanton (2018), resulting in operators being able to share information quicker and more frequently. At the level of the individual, SA is defined as 'the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future' (Endsley, 1995). However, it is argued that such a definition does not appreciate the complexity of control room environments. In such environments SA is not held solely in one operator's mind, rather it is distributed across many system agents, both social and technical. It is therefore proposed that the control room has Distributed Situation Awareness (DSA), as each actor contributes to overall awareness (Stanton, 2015; Stanton *et al.*, 2017b). Stanton *et al.* (2006) define DSA as 'activated knowledge for a specific task, at a specific time within a system'. This definition is congruent with control room operations, whereby the knowledge of various agents within the sociotechnical system is continuously shared and represented to achieve mission objectives.

An adapted representation of HMS Drakes' Talisman Trainer is shown in Figure 14. The trainer is split into two rooms: control and sound. All personnel except those related to Sonar are situated in the control room. Sonar personnel are in the sound room to reduce ambient noise as far as possible to aid aural detection. Some modern platforms have combined these rooms, such as the Canadian Victoria Class (Hunter, Hazen and Randall, 2014) or the USS Greenville (National Transportation Safety Board, 2001), capitalising on sound cancellation technology in headsets to reduce interference from ambient control room noise (Arrabito, Cooke and McFadden, 2005). The co-location of all operators provides an opportunity for shared information screens that could display relevant information, potentially improving DSA. However, these displays will need to be designed appropriately to ensure each operator can assimilate the information they require, for the completion of individual sub-tasks, which may come from several technologies and operators, for the achievement of higher-order team objectives.

The Officer of the Watch (OOW) leads the command team during tactical picture creation and is responsible for making decisions regarding submarine safety based on the teams' collective SA (DSA). A contemporary submarine control room has dozens of screens available to the OOW to make these decisions, see Figure 14. The figure shows the layout of a control room, including the picture compilation operators from Table 13/Figure 13 as a subset of overall capability. The addition of more screens may negatively impact their SA, as their cognitive capacity may be overwhelmed whilst using these screens (Hamburger, Miskimens and Truver, 2011). This is not limited to the OOW, as operators are expected to utilise multiple screens displaying complex information, with SA being highly distributed between humans and technology, both at the level of the individual and the team. Dominguez *et al.* (2006) found that data integration from multiple sources was a challenge for commanding officers, adding cognitive workload, and requiring near-constant communication with operators. Dominguez *et al.* (2006) also found that vital information was often not displayed appropriately for processing from a commanding officer's vantage point (standing at a distance). Operators could provide this information, however, they may be impeded by the command teams' complex communication structure, or erroneously disregard information that should be shared (Carrigan, 2009). Thus, interface designers must consider their location within the control room, their intended purpose, and current mission type, to facilitate maintenance of DSA and effective decision-making.



CO	Commanding Officer
ESM	Electronic Support Measures
MFC	Multi-Function Console
SMCS	Submarine Command System
WECDIS	Warship Electronic Chart Display and Information System
WT	Warner Transmissions

Figure 14 – A typical layout, adapted from Stanton and Bessell (2014), showing both the control and sound rooms

4.1.1.3 **Work and Communication**

During picture compilation, the command team works together to continuously update a tactical picture under direction from the OOW (National Transportation Safety Board, 2001; Dominguez *et al.*, 2006). The tactical picture is used to directly inform operational as well as strategic decisions; therefore, it is imperative to ensure its accuracy (how the environment is perceived vs its actual state). However, the complexity of modern control rooms poses challenges that may negatively affect tactical picture accuracy, namely the volume of sensor data and the communication of this data across the command team via which shared SA is attained (Roberts and Stanton, 2018). Another key challenge is managing uncertain data to compile a tactical picture; while data can be made more accurate, these uncertainties are a core aspect of submarine control room operation, meaning that information must be continuously evaluated for its validity and reliability.

As control rooms become more advanced, the amount of data being presented is increasing (Chalmers, Easter and Potter, 2000), this has the potential to exceed an operator's capacity to interpret such information effectively (Woods, Patterson and Roth, 2002). This has the potential to negatively impact console operation or verbal communication, which may cause operators to interpret or process data incorrectly, degrading the accuracy of the tactical picture. In highly demanding situations the cognitive capacities of operators have been demonstrated to reduce in a modality-specific fashion, this has the potential to increase attentional focus but also to reduce the volume of information operators can handle (Roberts and Cole, 2018). Data is presented to operators on Multi-Function Consoles (MFC), which are computer workspaces (Rhie *et al.*, 2017). Physically, they represent large cabinets, with space cut out to mount multiple monitors and a shelf for input devices (Bowden and Grosse, 2011). Most operators will use a single MFC however, some personnel, such as the OOW, will have access to many different MFCs (National Transportation Safety Board, 2001; Stanton and Bessell, 2014). Utilising their expertise and MFC, personnel will create and contribute information to the command team's DSA, which manifests in the overall tactical picture (Stanton, 2014). As detailed by Roberts and Stanton (2018), this information is the result of cognitive processing being applied to data perceived on an operator's MFC. Operators will perceive data, such as a sensor readout, and then gain an understanding through processing the data to create information for the tactical picture. Table 14 provides a list of typically available MFCs within a control room, with sample screens. The OOW views are composites of other MFCs and as such shall not be described.

Table 14 – Description of control room MFC roles, the screens that comprise their functionality, and the overall purpose of each MFC.

MFC / Role	Possible Configuration(s)	Purpose
Sonar	Medium Frequency Oscillations (MFO) Towed Array Narrowband (TANB) Flank Broadband (FKBB) Bow Broadband (BBB) Intercept	Detection, Classification, Localisation, and Tracking (DCLT) of contacts
Command System	Command System	Generating solutions (bearing, course, range, speed) for contacts to generate the tactical picture
Warship Electronic Chart Display and Information System (WECDIS)	WECDIS	Plotting, mapping, navigation
Electronic Support Measures (ESM) Warner Transmissions (WT)	ESM	DCLT and analysis of electromagnetic energy (emitted signals from contacts, communications towers, etc.)
Fire Control	Fire Control	Priming, targeting, controlling, and verifying weapon strikes
Periscope / Optronics	Search Attack	Providing visual DCLT of surface contacts
Ship Control	Helm Controls Plane Controls	Operating ship control surfaces to steer the submarine
All MFCs are paired with a communications screen that allows operators to select channels of communication from their headset.		

The complexity of MFCs and the quantity of screens present a vast amount of data for operators to process. Some control rooms currently provide HMIs designed to assist operators in maintaining SA/DSA (National Transportation Safety Board, 2001; Stanton and Bessell, 2014), such as the repeater screen onboard the USS Greeneville. However, there may be room for improvement to maximise the utility of these interfaces (Dominguez *et al.*, 2006). For example, Ly, Huf and Henley (2007) propose that effective watch leader decision making could be supported by interfaces that facilitate information accessibility. Such interfaces are being made possible with modern combat systems, which could represent all information required by operators in one interface (National

Research Council, 1997), instead of being dispersed across the control room. For example, information from all sensors could be displayed on one screen, removing the need for operators to manually corroborate data about contacts from each sensor. By implementing changes to simplify the HMIs, and the information they display, the potential for overwhelming operators could be avoided.

An increased amount of data also poses communication challenges, such as missed, incorrect, or non-timely communications. Such failures have been recognised as a risk factor in control rooms across many domains (Gibson *et al.*, 2005), including aviation (Cushing, 1994) and nuclear power (Lee, Ha and Seong, 2011). This challenge of effective communication is further exacerbated by the command rooms' complexity, and the variety of communication modalities, see Figure 14. Information is communicated via a combination of headsets, inter-MFC sharing, loudspeakers, and face-to-face conversations (Stanton and Bessell, 2014; Nakashima, Chow and Wang, 2015). Verbal messages received are acknowledged and repeated back, ensuring that the correct information has been received (Murphy, 2000b). All communication follows the rank hierarchy, with subsequent levels of leadership aggregating and filtering information until it reaches the OOW (Carrigan, 2009). The information is emergent and generated from interactions between social as well as technical agents, a key trait of DSA (Stanton *et al.*, 2017b). Changes to HMIs could support more effective control room communication, such as allowing operators to send information digitally to improve speed and reduce the potential for mistakes. In turn, this could mitigate communication-related challenges affecting control room operation, potentially reducing the risk presented by communication issues.

4.1.1.4 **Work of Sonar**

Sonar is a system for the location and ranging of objects using sound propagation and listening. Its four main functions are Detection, Classification, Localisation and Tracking (DCLT; Hughes *et al.*, 2010). To detect a vessel an operator will either hear discreet noise against the environment's background noise and/or see a concentration of sound on their waterfall forming a line, at a specific Direction of Arrival (DOA), using broadband Sonar, see Figure 15a. A waterfall refers to the main representation of sonar data in HMIs. Each time that a Sonar array returns data, it is plotted as a line on the display. When a new line is added, it moves all others down, creating a 'waterfall' effect over time (Asplin and Christensson, 1988; Matthews *et al.*, 2006), see Figure 15d and Figure 15e. Due to advances in modern boats (such as quieter engines, more efficient anechoic tiles, and advanced hull designs), it is possible that some signals from them will be quiet and intermittent, which could result in them not being detected. This is because they may not be readily discernible as clear traces (Matthews *et al.*, 2006). The system does not currently highlight such traces to an operator, which may affect submarine safety if they fail to detect them.

When a vessel is detected, operators assign it an identifier, allowing the system to automatically track and update its location, communicating details to other MFCs such as TMA. This identifier is a tracker (Fillinger *et al.*, 2010). Detection, and therefore localisation, is mostly performed using broadband Sonar (Zarnich, 1999), which detects sounds within a large frequency range.

Classification of a vessel is typically performed using narrowband (Plumb and Kendrick, 1981), which gives individual frequencies of a signal (Zhiyin and Lin, 2009), see Figure 15b. These individual noises are unique and can be compared against databases to identify a vessel. The classification is not validated by the system, meaning that operators are not advised of potentially incorrect classifications, despite the system potentially having the capability to do so. This can affect ownship safety and that of surrounding boats, as demonstrated by the Karen accident, in which an incorrect classification led the command team to believe that they could safely approach at a much closer distance than was safe.

Tracking of a vessel is performed by analysing broadband trends over time to determine actions a vessel is taking. Speed is also used to track a vessel, which is calculated using a DEMON (Detection Envelope Modulation On Noise (de Moura, de Seixas and Ramos, 2011)/Demodulation of Noise (Mill and Brown, 2005)). The calculated speed is passed verbally to be used by TMA. A DEMON waterfall shows a broken down (demodulated) broadband signal, allowing representations of individual shafts and propellers to be viewed (de Moura, de Seixas and Ramos, 2011), see Figure 15c. The frequency of the shaft combined with a Turns per Knot (TPK) value (how many times a propeller turns for one knot of speed) allows speed to be calculated. The TPK is obtained from a classification database or can be estimated based on the type of vessel. As the system does not validate classifications, TPK values can also be wrong, which will invalidate the contact's known position owing to its predicted movement(s) being inaccurate; it may be closer than thought, potentially on a collision course with ownship. The speed, together with the contact cuts, are passed to TMA with the tracker designation.

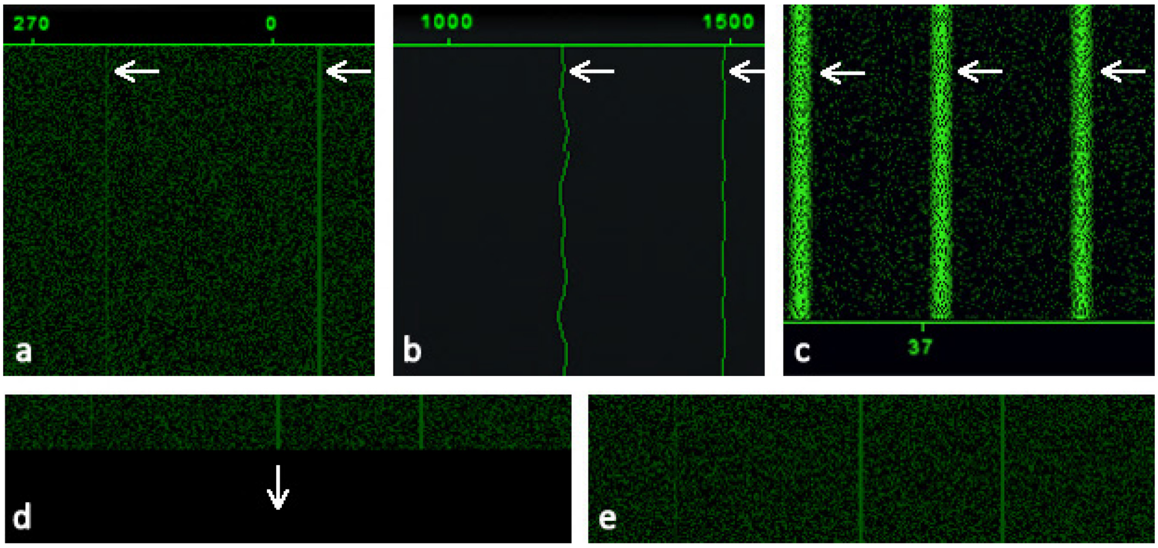


Figure 15 – a) A broadband waterfall showing ambient sound, with two vessels being detected,

b) A narrowband waterfall showing the frequencies making up a broadband signal,

c) A DEMON waterfall showing the number of propeller shafts and blades detected for a contact. Each group of lines represents one shaft, with the number of lines representing blades,

d) A broadband waterfall display that has just started to show data, notice the bottom is empty

e) Data at the top of ‘d’ has been pushed to the bottom by newer data, giving a waterfall effect

4.1.1.5 **Work of Target Motion Analysis**

TMA is the process of analysing positional data from contacts derived from passive sensors to produce a location and predicted movements (Murphy, 2000c). This is called a ‘solution’, comprising the speed, course, range, and bearing of a vessel (Genç, 2010). Bearing is the direction to a contact from ownship. To generate solutions, a Local Operations Plot (LOP) can be used (see Figure 16a). A LOP is a chart with a contact’s previous detections plotted, allowing solutions to be calculated (Clarke, 1999).

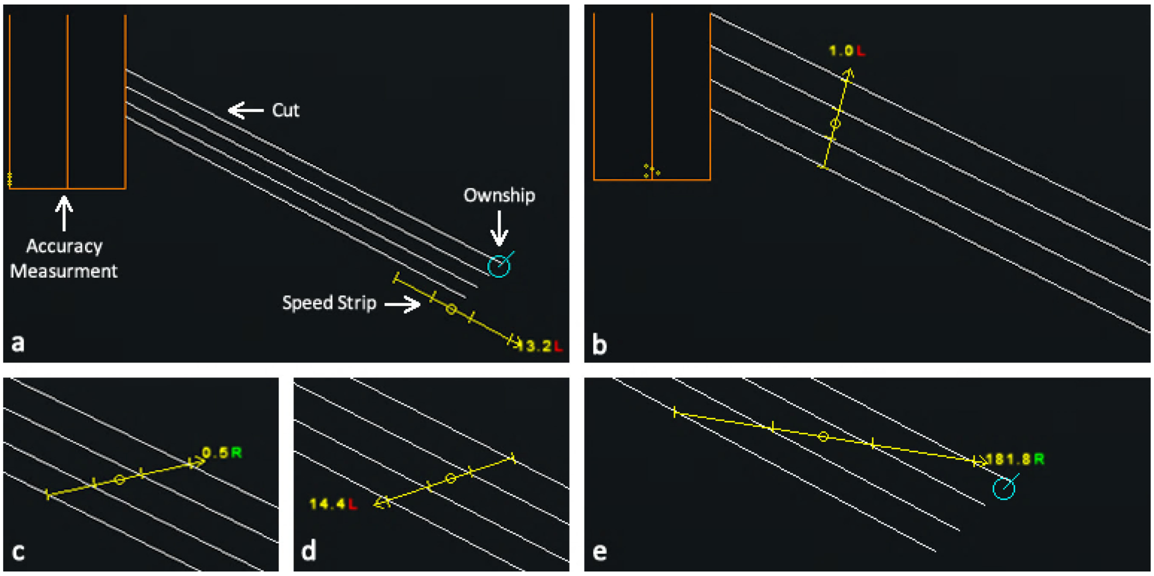


Figure 16 – a) A labelled Local Operations Plot interface screenshot b) A finished solution, ready for sharing c) d) e) Different matching potential solutions

When a sensor such as Sonar makes a detection, a ‘cut’ will be sent through from the detecting MFC to TMA (Stanton and Roberts, 2018). A cut is a straight line that represents the Line of Bearing (LOB) for a signal; it is plotted on the LOP between the submarine’s position and the maximum detection range of a sensor, with an angle equivalent to the detection bearing. Cuts from each sensor are grouped by detected vessel, with one shown at a time. Operators can merge cuts from two sensors and treat them as one contact (Huf, Arulampalam and Manning, 2006), which provides more information on its behaviour. However, if the system does not verify merges, it is possible for operators to perform a merge using two (or more) discrete contacts, which will result in incongruent information from multiple sources being displayed to an operator. This will degrade their ability to effectively process that information unless the mistake is identified and rectified.

Once enough cuts have been accrued for a vessel, an operator can start analysing the vessel’s path using a speed strip. A speed strip is a manoeuvrable visualisation of a vessel’s historic path in the water (DeAngelis and Green, 1992), with optional marks to represent where a vessel would be if the path was correct. Whilst the strip is manoeuvrable, it can be cumbersome for operators to move to the desired solution position. A mark is added to the strip for each cut. The strip is aligned over the cuts by an operator, and if the marks intersect the cuts, the vessel *could* have travelled in the manner indicated. An accuracy measurement view represents the spatial difference between cuts and a speed strip, with dots representing the strips’ intervals, see Figure 16b. As a solution becomes more accurate, the dots will form a vertical ‘stack’, showing they are close to the cuts they represent and align with each other (Huf, Arulampalam and Manning, 2006). The term ‘*could*’ is used because there are multiple speed strip configurations that may align, however, they would not all be correct (DeAngelis and Green, 1992).

Whilst the system provides a method to validate the accuracy of a solution, it does not assist the operator in creating an accurate solution. For example, there is no option to fix cut marks to their matching cuts, significantly reducing the number of available solutions. Figure 16c-e shows sample LOPs, demonstrating how the speed strip matches cuts in a variety of configurations. Once a solution that matches all cuts is derived, see Figure 16b, it is shared (Huf, Arulampalam and Manning, 2006). Sharing is where contact information is made available to other nodes within the system. When this occurs, it is plotted as a contact marker onto the geographical view, see Figure 17b, allowing relevant personnel to view the vessel's location. On some systems, a contact marker already exists, but it is marked with a cut to indicate that only the bearing is known, see Figure 17a. Unless a vessel changes its parameters (how it is moving) or the accuracy degrades, the solution does not need to be manually updated, dead-reckoning will be used to plot further movements. Dead-reckoning is plotting positional data by extrapolating previous trends (Murphy, 2000a). For example, if a solution had a vessel travelling at ten knots, the contact marker would move as if travelling at that speed without needing operator intervention, illustrated by contact markers moving between Figure 17b and Figure 17c.

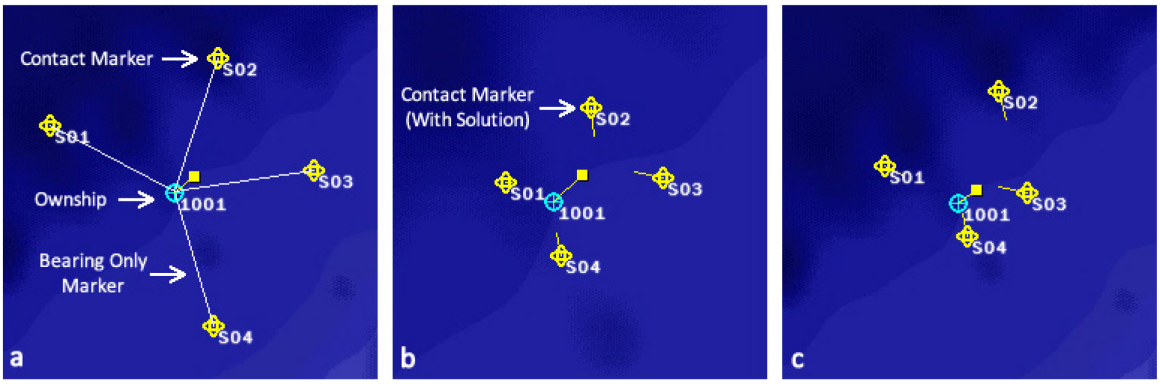


Figure 17 – a) A geographical view with four contacts marked when only their bearing is known, b) A geographical view with four contacts marked with full solution data, c) Illustration of automatic movement of contacts from 'b' over time

4.1.2 Technical: Combat System

The activities of the command team are supported by the combat system, which encompasses a wide variety of technology available within the control room. A combat system comprises multiple software subsystems including Sonar, periscope, radar, and command systems such as the Submarine Command System (SMCS; Dominguez *et al.*, 2006). Combat systems facilitate the management of information about contacts and make such information available to the command team. Systems are partially networked, facilitating data sharing between the Sonar and control room (Owen *et al.*, 2006; Emery, 2010). Each software subsystem could also be considered a technical agent in the control room sociotechnical system, as they engage in goal-directed behaviour (providing sonar data, tracking targets), and interact with other agents, both socio (displaying and receiving information) and technical (networked communication). This provides further evidence of control room DSA (Stanton, 2015), with technological agents also sharing information to facilitate the maintenance of SA.

4.1.2.1 **Backend: Open and COTS Systems**

Typically, combat systems have been created to bespoke military specifications, with subsystems supplied by various vendors. However, defence agencies are now adopting Consumer off the Shelf (COTS) systems to create open architectures, such as the Royal Navy (Owen *et al.*, 2006) or the United States Navy (Womble *et al.*, 2011). Common Core Combat System (CCCS) is the Royal Navy's' COTS-based combat system and was created to lower lifetime costs, support reusability, modularity, and capability progression (Owen *et al.*, 2006). Information sharing is no longer restricted to pre-programmed communications between systems, defined in a military specification, rather all information is available and systems can subscribe to receive what they require (Owen *et al.*, 2006). For example, as soon as a sonar speed is calculated, it could be shown on all TMA screens. This could address the communication bottlenecks identified by Roberts and Stanton (2018), removing the need for multiple sociotechnical agents to be involved with the transfer of data via verbal communication. Furthermore, it may also make the attainment of DSA more efficient by allowing MFCs, which can communicate almost instantly, to share information at a faster rate than what is encountered when information is verbally passed between multiple command team operators.

A cluster of COTS servers is used to create a Shared Computing Environment (SCE), which powers all control room technology, including legacy systems. Using a COTS cluster has resulted in increased processing power and a reduced footprint in comparison to previous systems. For example, onboard HMS Astute, Sonar 2076, command, and navigation functionality are now powered using sixteen processor parts, instead of over 200, housed in only a few cabinets (Defence Equipment and Support, 2010). This has created the physical space and computing capability for the addition of more systems that could assist operators, such as artificial intelligence agents to solve complex TMA datasets. Additionally, newer capabilities allow deployment of updated HMIs, such as the new Sonar 2076 update designed to capitalise on cutting-edge sonar technology (Royal Navy, 2015).

The benefits of adopting a new generation of combat systems are clear, driving futuristic capability with consideration for organisational factors, such as cost and risk. It is imperative, however, that HMIs are developed at the same pace to make full use of the systems.

4.1.2.2 **Frontend: User Interfaces**

Adoption of open combat systems has included the upgrade of MFCs, and changing HMIs to improve functionality and usability. In doing so, two key problems are addressed: the stasis of core HMI concepts and the addition of new features.

Despite considerable capability advances across the history of submarines, certain HMI aspects have remained largely unchanged or seldom updated for several decades. This is illustrated in Figure 18, a collection of Sonar HMI images ranging from 1989 to 2015. While spanning three decades, the waterfall from 1989 is visible on each. Leftmost is from the USS Hyman G. Rickover (SSN-709) and rightmost is Sonar 2076, used across the Royal Navy's' contemporary submarine flotilla. Exact dates and implementations are classified; however, little difference is apparent. Whilst specifics such as screen layouts or different waterfall aesthetics have changed, large commonalities remain (such as the use of green and black, or the 'waterfall' display: new data moves old data down the screen, causing a waterfall-like aesthetic).

Each interface, a product of their time, met requirements within the available budget, computing capabilities, and core considerations for submarine design (Burcher and Rydill, 1995). As time progressed, budget sizes (Fallon, 2015) and computing processing power increased, allowing more advanced systems to be procured and more processing power to be achieved, but the previous interface style tends to have been retained. This has resulted in highly advanced systems, but ones that employ legacy design paradigms. Reasons for this could be the retention of existing training programmes, subject matter expert familiarity with existing systems, or the potential risk associated with 'buy-in' to next-generation HMI (Gosling, 2008; Hall, 2012).

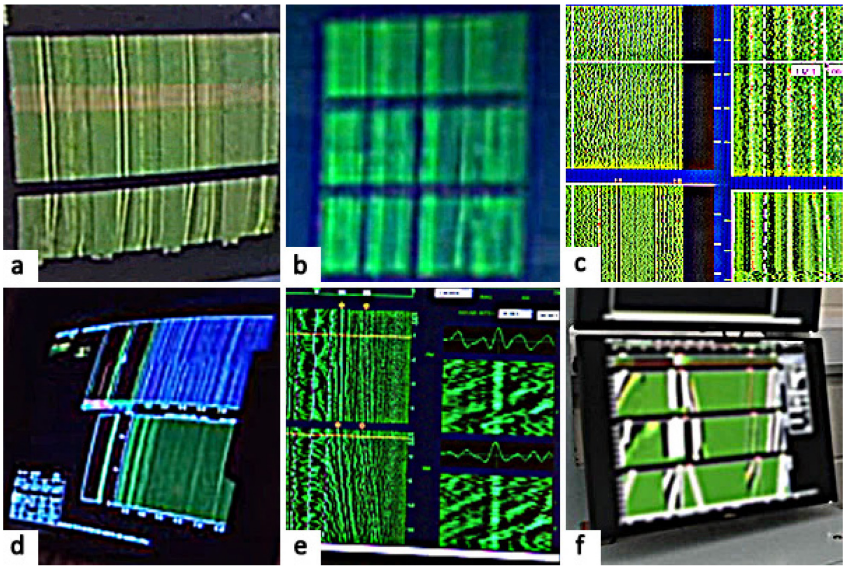


Figure 18 – Chronological screenshots of Sonar user interfaces across time, showing their similarities, from: a. Hoffman (1989), b. Bosner and Oxley (2009), c. Jones (2009), d. Canadian Armed Forces (2013), e. Webber (2015), f. Royal Navy (2015)

This resistance to change may be due to maintaining training readiness (Hall, 2012), reducing costs, and reducing risk (Gosling, 2008). Additionally, onboard factors include space requirements for supporting equipment, processing power, system compatibility, and implementation time (Defence Equipment and Support, 2010). This has led to evolutionary improvements (Roberts, Stanton and Fay, 2015), but this does not preclude further improvements; comparison with different ways of working would determine if the current paradigm is suitable for future requirements, or whether a step change is required. Nevertheless, submarines are at the forefront of technological innovation. For example, Sonar 2076, while having a similar HMI to predecessors, is vastly more capable, with increased detection capabilities (Royal Navy, 2008). While individual changes seem small, their combined effect can be observed by comparing the Turtle (Clautice, 1978) to HMS Artful (BAE Systems and Launderdale, 2016). HMS Artful, like all submarines in the Royal Navy, is highly advanced, allowing command team personnel to carry out a multitude of deployment objectives. This success has been engineered from continuous improvement of the entire submarine, including user interfaces, such as Sonar 2076, to maximise the exploitation of these modern capabilities in an evolutionary fashion (Roberts, Stanton and Fay, 2015).

4.2 The Case for Ecological Interface Design

While submarine control rooms are currently an advanced product of evolution, afforded by constantly trained operators and ever-changing combat systems, it does not mean they cannot be improved upon (Stanton, 2014). In particular, the lack of substantial HMI development discussed above is an area that has the potential to maximise current and future capabilities. The challenges of a contemporary and future global maritime environment may necessitate changes to remain at the vanguard of capability as well as safety. As a complex sociotechnical system, there are a variety of aspects that could be assessed to see where changes could be implemented, such as crewing requirements or sensor capability. However, the USS Greeneville and Royal Navy submarine incidents have highlighted that HMIs can play a critical role in control room operation, performance, and safety, creating an impetus to assess how they may be improved to avoid future incidents and to capitalise on new advances and capabilities provided by modern combat systems.

Maintaining a tactical picture relies on operators understanding and creating information about their environment. Despite this, information from HMIs often takes on a different form, unrepresentative of its physical manifestation, requiring it to be perceived and processed. For example, despite representing 360° of aural signal, the Sonar waterfall is not circular, requiring operators to translate the plot of their surroundings. For TMA, details about a cut strength are not represented, and the last cut is not marked. Furthermore, design issues such as transient signals in the sonar waterfall not being highlighted, or TMA solutions not being constrained may add further complexity and cognitive workload. Interactions between perceptual and cognitive processes can affect operator performance (Hanisch, Kramer and Hulin, 1991; Masakowski and Hardinge, 2000), as can overly complex screens (Coll and Wingertsman, 1990). As most control room tasks focused on a submarine’s immediate environment, especially with Sonar and TMA, the EID paradigm could offer a novel approach for future development. This section will make the case for using EID as a design paradigm for the development of future Sonar and TMA interfaces.

4.2.1 Existing Ecological Interface Designs

The potential of EID as a design paradigm has already been demonstrated across a variety of domains, including power generation and petrochemical production (McIlroy and Stanton, 2015b). The parallels that can be drawn between the issues faced by complex control rooms in other domains and the maritime domain strengthen the case for EID to be applied; it would be an egregious oversight to ignore these parallels and the substantial body of literature (McIlroy and Stanton, 2015b) highlighting the application of EID.

Lau *et al.* (2008) identified that while conventional nuclear power plant interface design had led to interfaces with acceptable performance and safety records, unanticipated events had been a causal factor identified in accident investigations. They noted that unanticipated non-routine events were precursory to these accidents, arising because of the complexity of modern nuclear power plant control rooms. To mitigate this, EID interfaces were designed that would aid operators in handling the complexities of modern control rooms. These EID interfaces were designed from completed WDAs, with the environment and its constraints being represented in the resultant designs (Lau *et al.*, 2008). An assessment of these designs by Burns *et al.* (2008), showed increased SA in certain circumstances, demonstrating the potential for EID to be successfully applied as a design paradigm in nuclear power plant control rooms.

In the petrochemical industry, Jamieson (2007) identified a failure of interface technologies to match technological advances as a contributory factor towards an estimated billions of dollars (Bullemer and Nimmo, 1994) in annual avoidable losses. They designed and tested EID interfaces for production control and found evidence of potential benefits for application in an industrial setting. It was also noted that these benefits demonstrated that EID should contribute to improving safety and productivity.

In the military domain, Hall and Miller (2009) tested their Representation Aiding Portrayal of Tactical Operations Resources (RAPTOR) tool, designed to support military decision-making, and noted that it was more effective than a baseline interface in all areas.

In the maritime domain, organisations are making design changes, and consequently, systems that appear to encompass some of the principles of EID to varying degrees already exist for Sonar (Atlas Elektronik, 2016a) and Combat Management (General Dynamics, 2014; Havelsan, 2015; Atlas Elektronik, 2016b). However, it is unclear as to whether EID was explicitly followed as a design paradigm, or whether the companies' designs reflect a desire to reflect information in a highly usable manner. In the case of the former, the principles could be applied to further extend the usability of the interfaces and offer a theoretically underpinned approach to their design.

Research is also being conducted on different design paradigms for future command interfaces that assist the OOW with managing the amount of information contained within the tactical picture (Ly, Huf and Henley, 2007; Hunter, Hazen and Randall, 2014; Danczyk *et al.*, 2015). Comparing the proposed navigation-aiding interfaces by Danczyk *et al.* (2015), see Figure 19, and Ly, Huf and Henley (2007), see Figure 20, it is evident that both interfaces achieve their functionality using substantially different HMIs. As interface design can be subjective, the ecological representation of objects may differ, producing differing levels of usability in certain aspects; whilst not performing the same functionality, they have a core set of features they both complete, as would be expected from a navigation system. Thus, while it is important to build upon already successful designs, it is also useful to explore alternative designs within the EID research space. Furthermore, navigation is a different work domain to Sonar and TMA, although does take information to map the environment; what has been demonstrated to work well for one work domain might not be applicable to another.

In doing so, the issue of design concepts being evolved, constraining future designs, can be avoided. Furthermore, these designs concentrate on using EID to enhance the OOW's SA. This is a worthwhile endeavour, as it is the OOW who is responsible for maintaining the tactical picture, and ultimately maintaining submarine safety. However, they do not fully address the needs of operators compiling the tactical picture; command team DSA/SA may be compromised due to these operators expending more cognitive workload to determine information being made readily available to the OOW. Thus, it is pertinent to research how the pace of change can be kept across the control room, ensuring maximum utility of benefits.

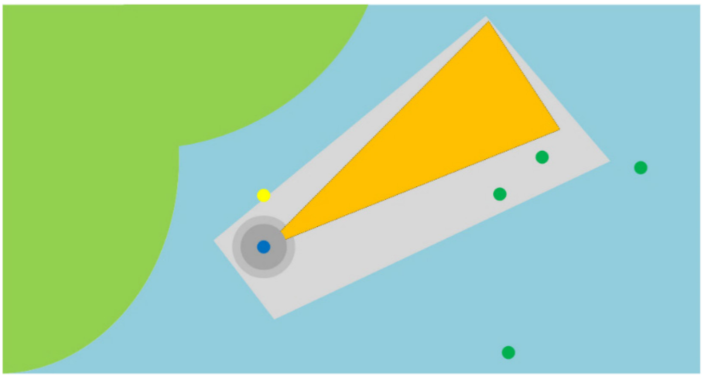


Figure 19 – A simple representation of a navigation assistance interface proposed by Danczyk *et al.* (2015).

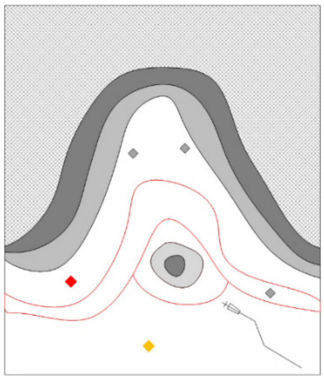


Figure 20 – A simple representation of a navigation assistance interface proposed by Ly, Huf and Henley (2007).

Previous work (Bisantz *et al.*, 2001; Bisantz *et al.*, 2003; Burns, Bisantz and Roth, 2004; Burns, Bryant and Chalmers, 2005) in the maritime domain has demonstrated that CWA can be used to understand naval command and control functionality. In their work, they used WDA to derive system design requirements for a Canadian frigate (Burns, Bryant and Chalmers, 2005) and a USN surface combatant (Bisantz *et al.*, 2001; Bisantz *et al.*, 2003). Due to the scale of the analyses, recommendations could be made for various aspects of operation, ranging from control room to interface design. Bisantz *et al.* (2003) developed prototype interfaces that yielded acceptable performance, indicating the suitability of applying WDA to derive design recommendations. Whilst the design was not explicitly identified as an EID the utilisation of CWA and the referenced work implicitly implies the display is of an ecological nature. This work comprehensively analyses surface vessels, but it may not be suitable for deriving design requirements for submarines due to the different work domains.

Interfaces utilising EID principles are not just limited to research. Warship Electronic Chart Display and Information Systems (WECDIS) are already present in modern submarine control rooms, presenting chart data in an ecological fashion to aid navigation. An example WECDIS interface, adapted from Offshore Systems Ltd (2007), is shown in Figure 21. The ownship icon and land representations are physical information, showing the current environment. Functional information is overlaid to show areas of danger, calculated from factors such as sea depth, current and geographical features. Whilst some interpretation must still be performed, the environment and its constraints are immediately apparent. This is a core aspect of EID and demonstrates the suitability of the method for use in submarine control rooms. Furthermore, it establishes a precedence of EID, or similar/adjacent capability, being used by the Royal Navy, showing that it can successfully integrate and provide training on a new generation of interfaces.

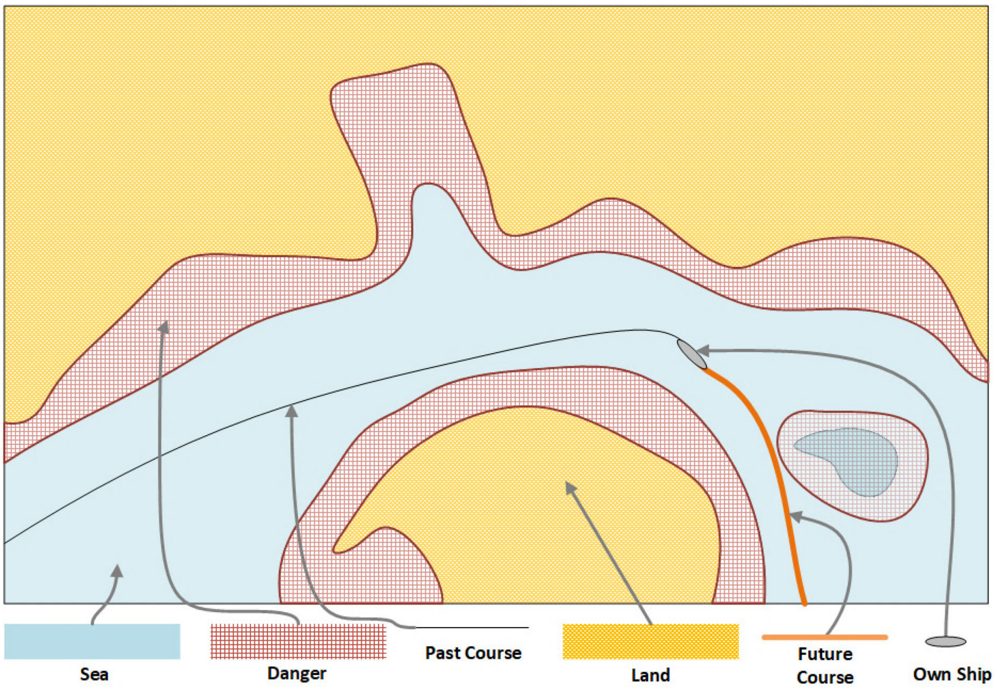


Figure 21 – A simple representation of a WECDIS product, made by Offshore Systems Ltd (2007).

4.2.2 EID Application to Sonar and TMA

Precedence for implementation in an operational submarine control room further strengthens the case for EID. Moreover, it addresses organisational as well as technical factors that may have affected redesign efforts previously, such as a desire to maintain training readiness (Hall, 2012), reduce cost as well as risk (Gosling, 2008), and technical capability; if these can be overcome for one system there may opportunity for other systems, such as sonar and TMA. The remainder of this section outlines initial ideas of how EID principles could be realised within these systems.

With modern combat systems such as CCCS affording MFCs greater access to data, displaying both physical and functional information could be achievable. For example, consider a TMA interface that is showing Physical cut information from sonar. A TMA operator may not be aware of the maximum sonar detection range at a given time and may attempt to place a solution outside of this range. Alternatively, they may attempt to place a contact in the sonar baffles (blind spots), which is improbable, as no cuts would be received.

In both instances, functional information regarding the sonar sensors' range and coverage could be overlaid onto the LOP to provide the operator with constraints to place the solution inside. Currently, the TMA system does not have this functionality, placing the onus on operators to perform these checks mentally. Whilst operators are well trained, a high workload situation may lead to mistakes or omission of sense checks. Furthermore, whilst sonar information could be retrieved from a sonar operator, bottlenecks in the system may impede the effective and timely communication of this data (Roberts and Stanton, 2018). Providing support directly in an interface developed using EID may relieve these issues, by reducing cognitive workload (e.g., mental manipulation of complex material) and additional communications across a complex network. In turn, this would benefit the entire control room sociotechnical system by ensuring tactical picture accuracy to ensure ownship safety, the top priority of any submarine (Mack, 2003).

For operators working directly with MFCs, such as Sonar and TMA, improvements could be made to assist operators in overcoming challenges with their usage. For example, transient sonar signals are not currently highlighted. Therefore, operators must continuously scan the waterfall history for these signals, or a pattern of signals (Matthews *et al.*, 2006). This increases workload and may distract the operator from other tasks. Following an EID approach to sonar design, support for RBB could be added, and transient signals could be highlighted to draw an operator's attention for investigation, removing the need for them to manually find these signals.

As sensors become more advanced, they will detect more contacts. Maintaining suitable SA about each of these contacts may exceed an operator's capability unless the HMI allows intuitive storage and interpretation of the data. For sonar, this could be achieved by adding support for SBB in the interface, by utilising a skeuomorphic representation of available arrays. Current representations require mental translation to a 360° representation, creating additional operator workload. Representing the waterfall data using a sensor's geometry could remove this workload, allowing operators to intuitively understand the location of traces in relation to ownship.

KBB behaviour could be supported by ensuring that operators have all the required information available to support their tasks within their HMI. This may be of substantial benefit to TMA operators, who incorporate information from a variety of social and technical agents within the control room to generate contact solutions. Whilst modern combat systems are advanced, certain information is still transferred manually, such as the speed calculated by sonar. The speed is generated within a sonar MFC, yet is not passed digitally, nor displayed to a TMA operator. This adds unnecessary communication and relies on the TMA operator to manually incorporate the speed into their solution. Given the amount of data being processed by TMA, this could add a substantial cognitive workload, which may be further exacerbated in the future by more advanced sensor capabilities. An EID design incorporating all required work-domain information could reduce the cognitive workload and difficulties (communication bottlenecks, many screens) associated with collating it manually; when information pertinent to TMA is generated by other MFCs, it could automatically be displayed to TMA operators, and their solutions validated against the information.

In addition to taking advantage of innate abilities, supporting appropriate multi-level control allows operators to behave effectively when faced with new, unfamiliar situations (Drivalou and Marmaras, 2009), assisting in ensuring a correct and safe outcome. With submarines undertaking more missions and mission types than previously, such situations may be encountered with a greater frequency, potentially leading to incidents such as the USS Greeneville (National Transportation Safety Board, 2001) or Karen (Marine Accident Investigation Branch, 2016) accidents. EID interfaces for operators that rely on summary screens, such as the OOW, could enhance their SA and decision-making process during these situations.

For example, if a proximity contact on a collision course with ownship were detected suddenly, then the OOW would be required to manoeuvre quickly and correctly to avoid the danger. If diving were not possible due to operating in busy littoral waters, the OOW would have to steer the submarine to a safe area of water. The OOW may be storing the tactical picture in their mind, without any effective assistance from command room screens (Ly, Huf and Henley, 2007); this would require the OOW to calculate a course of action mentally, which may incur a temporal penalty and not be the optimal course of action. Decision support screens, such as those proposed by Ly, Huf and Henley (2007) and Dominguez *et al.* (2006), designed using the EID paradigm could address these issues, allowing the OOW to plot a safe course of action with a comparatively low cognitive workload. This functionality could even be extended further to plot optimal courses to safety automatically. As a maritime environment can be challenging and unpredictable, it is advantageous to provide this support, reducing the risk posed by unfamiliar situations.

In summary, whilst evolutionary design has yielded capable control rooms, there are still difficulties associated with their operation. This issue is present across multiple domains, not just maritime. However, EID could provide a means to mitigate these issues. With the objective of EID being synergistic with submarine control room operation (understanding the environment), and its core design goals supporting key issues identified across the control room, there is a strong case for it to be utilised. Furthermore, the implementation of EID interfaces may not have been previously possible due to organisational (risk aversion, training requirements) or technical factors (computing capability). However, with the advent of modern combat systems addressing these factors, combined with precedence of adoption, an opportunity is afforded to capitalise on a design paradigm that could contribute to ensuring the suitability of future control rooms for challenges that lie ahead.

4.3 Conclusion

This chapter detailed how submarine control rooms operate as a complex sociotechnical system. The social aspects were described first, starting with the entire command team's objectives, before narrowing down the focus to the work of Sonar and TMA as a focus of this thesis. Focus then moved to the technology that supports the command team in their work, and the challenges associated with its utilization. Finally, the case for EID being applied to Sonar and TMA, accounting for the practical issues discovered throughout the chapter, was put forth.

The crux of this case was that while an evolutionary approach has kept submarines at the vanguard of capability, using contemporary means to achieve their mandate, emergent issues require addressing if they are to be capable of meeting future demands. Additional sensors, more data, more displays, and reduced crew numbers will create an onus for improvement that must be met to operate effectively. A need to address emergent issues stemming from increasing requirements is not necessarily limited to submarine control rooms. While control rooms across many domains continue to evolve to meet current requirements, future requirements may challenge this model of adaptation's success.

The interactions of the command team and the combat system to achieve a variety of missions define the control room as a sociotechnical system. One area of improvement may be the HMIs, which can be critical factors to success; current designs can accommodate contemporary submarine capabilities, however as the envelope of innovation is pushed, their efficacy is diminishing. This issue is not limited to the maritime domain. Avoidable losses for petrochemical control rooms or accident causation in nuclear power plant control rooms, serve as motivation to explore how best to meet future challenges. Decreased effectiveness has led to accidents where HMIs were identified as a contributing causal factor. Thus, it is pertinent to assess how they may be improved, concentrating on Sonar and TMA in the maritime control room domain due to their prevalence.

The command team works to generate a tactical picture, whilst maintaining the three main objectives of submarine operation: remain safe, remain undetected, and complete the mission. As missions can be varied and complex, it is important that the Sonar and TMA HMIs support operators in maintaining effective performance in all situations. This performance may be negatively affected by the complexity of control room operations and HMIs. In the past, organisational, and technical factors have impeded implementing a solution. However, modern combat systems could provide an opportunity to make HMI design changes. While this section has concentrated on the maritime domain, it is possible that other domains are making advances that could facilitate similar design changes. For all domains, a step change in HMI design could ensure optimal operator performance, and that they are fully supported by the interface.

EID is proposed to mitigate issues with current Sonar and TMA HMI design, as the nature of work within the control room is synergistic with what it provides as a design paradigm (McIlroy and Stanton, 2015b): a focus on the environment, its properties, behaviours, affordances, and constraints. While designs employing EID principles currently exist for both Sonar and Combat Management (which includes TMA), with some in use within fleets, it does not mean that they cannot be improved upon; the benefits, or otherwise, derived from redesigning using EID should be explored to ensure their continued future effectiveness. As with their predecessor interfaces, these EID systems should continue to evolve to stay at the vanguard of function and innovation, facilitated by domain-specific advances (such as the flexibility and power of open command systems for submarine control rooms). The potential benefits of EID, coupled with the capability to realise them provides an interesting research space, which this thesis will explore. This process is started in the next chapter, which details the application of Work Domain Analysis and Worker Competencies Analysis to the Sonar waterfall to identify where improvements could be made.

Chapter 5 **Cognitive Work Analysis of Sonar: What could be improved about the waterfall?**

5.1 Introduction

Having made the case for Ecological Interface Design (EID) as a design methodology to explore in Chapter 4, a completed systems analysis was required to proceed with the redesigns. While there was an understanding of how submarine control rooms operated from the review in the first half of the chapter, this was prescriptive (describing what the system should do), as opposed to the formative (describing what it could do) analysis required for EID. This is an important distinction as work completed is not necessarily reflective of the work domain, and tasks completed might have become routine and dissociated from their originating constraints (Rasmussen, Pejtersen and Goodstein, 1994; Naikar, Hopcroft and Moylan, 2005). Therefore, Cognitive Work Analysis (CWA) was applied to the Sonar role in ComTET (the simulator facility where experiments in this thesis were conducted) to determine where improvements could be made, and the outputs are presented in this chapter, along with design recommendations. There is a focus on the waterfall as a leverage point, described by Read *et al.* (2018) in the CWA Design Toolkit as a system aspect that could yield large changes across the system if changed in a small way. The importance of waterfalls is corroborated by their prevalence in the Sonar literature (Chen and Burns, 2007; Ericsson, 2009; de Moura, de Seixas and Ramos, 2011). This importance makes it vital to understand how they work to achieve their goals, and where changes could be made to better support them. CWA was selected to model the system because of its close association with EID, and the prevalence of literature demonstrating this (Burns and Hajdukiewicz, 2004; McIlroy and Stanton, 2015b). While all stages of CWA can be applied to EID, the originally required, Work Domain Analysis (WDA) and Worker Competencies Analysis (WCA; McIlroy and Stanton, 2015b), were selected.

The focus on the waterfall is because it is a core component of most Sonar implementations in some form, and changes to the way that it is implemented could have far-reaching benefits, especially given its key role in maintaining ownship safety. As Sonar is the origin of most information about a submarine's operating environment, it is key to maintaining distributed situational awareness, defined in Section 4.1.1.2, throughout the entire control room. Accidents involving sonar (National Transportation Safety Board, 2001; Marine Accident Investigation Branch, 2016) have shown that the consequences of this awareness being degraded can be catastrophic. The key factors identified in these accidents were operator workload and the availability of sonar information. With the constant evolution of sonar systems and the amount of data they provide (Dominguez *et al.*, 2006;

Gosling, 2008; Defence Equipment and Support, 2010; Jacobus, Yan and Barrett, 2012; Smith *et al.*, 2013), it is pertinent to focus design directions on reducing operator workload and making information more readily available, especially if there is a risk of the amount of data exceeding operators' ability to process it (Woods, Patterson and Roth, 2002) or their working memory capacity (Mason *et al.*, 1989). Working memory is defined as the necessary systems to keep items in memory while performing complex tasks such as reasoning, comprehension, and learning (Baddeley, 2010). This issue is potentially exacerbated by the disparity between the external environment and how aural data is represented by a sonar interface, with the interaction between perceptual and cognitive processes negatively affecting operator performance (Hanisch, Kramer and Hulin, 1991; Masakowski and Hardinge, 2000).

Consequently, this chapter focuses on how the waterfall could be redesigned with increased ecological validity, shifting cognitive workload to perceptual workload, in line with the principles of EID. It details the completed abstraction hierarchy and worker competencies analysis, proposing design ideas for a Sonar redesign. For the former, the analysis will break down the waterfall into its component Physical Objects to examine what work could be completed using them and propose where improvements could be made to better support operators in achieving nodes at higher levels of abstraction hierarchy. It is recognised that the WDA is a formative method, focussing on what could be achieved, although this is made more accessible by understanding how functionality is currently offered to contextualise improvement recommendations.

5.2 Method

The method for the creation of WDA and WCAs for both Sonar (this chapter) and TMA (next chapter) are presented below.

5.2.1 A Different Approach to WDA

The work by Rasmussen, Pejtersen and Goodstein (1994) and Vicente (1999b) is the basis for CWA, see Section 2.3.1. However, there have previously been differing theoretical approaches and methodologies for WDA, the first phase, which can limit its accessibility and applicability (Naikar, Hopcroft and Moylan, 2005). The key reasons Naikar, Hopcroft and Moylan (2005) identified for this were contrasting presentations of the theoretical underpinnings, work that is difficult to interpret or generalise, and limited methodological discussion in seminal work. Their work aimed to address the lack of coherency by addressing conceptual issues and proposing a methodology for performing WDA, although they acknowledged that a methodology could never be completely specified but must align with the underlying theory. Lintern (2013) shares this sentiment, explicitly stating that their method could be adapted as required so long as it fits their treatment of the underlying theory. Additionally, Kant and Sudakaran (2022) highlight how CWA theory has evolved with EID as the methodological application has changed over time.

The adaptability of the CWA method is reflected in seminal work that combines applications (Burns and Hajdukiewicz, 2004; Jenkins *et al.*, 2009; Stanton *et al.*, 2017a), showing variability is acceptable, as well as work that proposes modifying CWA to account for contemporary system characteristics (Burns, 2012). Thus, CWAs in this chapter were compiled using the approach outlined by Stanton *et al.* (2017a), based on the method described by Naikar, Hopcroft and Moylan (2005), supplemented by the direction Burns and Hajdukiewicz (2004) provided for creating a CWA. Briefly, this approach comprises:

1. **Establish the analysis purpose:** Define what will be considered in the analysis, and what the desired outcome(s) is/are.
2. **Identify project constraints:** Determine what factors will constrain the analysis, such as time or money.
3. **Identify analysis boundaries:** Determine an appropriate analysis size, that should capture what is required to achieve the desired outcome(s).
4. **Identify the types of constraints in the domain:** Identify what categories and types of constraints are present in the domain.
5. **Identify sources of information for the analysis:** Identify where information can be sourced from, aiming to utilise a variety of sources.
6. **Construct the abstraction hierarchy using readily available information:** The abstraction hierarchy should be created using information that can be readily sourced and interpreted.
7. **Conduct special data collection exercises:** Seek further information on the domain or from subject matter experts.
8. **Review with domain subject matter experts:** The output should be reviewed by domain subject matter experts.
9. **Validate the output:** Determine if the output is suitable through a validation exercise.

It was decided to modify this method for the purposes of the WDAs conducted in this thesis, with the aim of exploring if methodological improvements could be made. The method was changed to include the Sonar and TMA HMIs in the analysis and to bound the analysis to them, in contrast with the general expectation to not include the item being designed (Burns and Hajdukiewicz, 2004). Sonar and TMA are a mix of law- and intent- driven domains (Bennett, 2014). Additionally, they implicitly employ metaphors for the work domain as it can seldom be directly observed, so representations of the environment are used, such as the tactical picture or Local Operations Plot.

It became apparent while performing the analysis that this had caused representations of Physical Objects from the work domain to implicitly become present in the interfaces, along with metaphors of the work completed. Consequently, Physical Objects in the display could be used as analogues to represent the work domain at the Physical Objects level of the abstraction hierarchy. An example of this would be the use of metaphors representing real-world equivalents in computing, such as the Desktop (an actual desk with files on) or the Recycle Bin (a space to store unneeded files before their complete disposal). A new HMI being built from the abstraction hierarchy could then build on the upper levels using these implicit metaphorical representations, yielding information requirements on new Physical Objects that could better support nodes above them that offer representations with a higher ecological validity than those currently present. This could also be used to build on "Physical Object Cards" in the CWA Design Toolkit (Read *et al.*, 2015b), which was intended to promote the exploration of new ways to use existing Physical Objects within the system.

Scoping the WDA to the existing HMI provides a level of abstraction that is appropriate to end-goal of a prototype interface, similar to the approach advocated for by Kortschot *et al.* (2017). The work domains of Sonar and TMA are very complex, which is reflected in systems designed for these domains, such as Sonar 2076 (a submarine sonar suite used by the Royal Navy (Gosling, 2008)) or the Common Core Combat System (a modern combat system for Royal Navy submarines (Owen *et al.*, 2006; Scott, 2006; Defence Equipment and Support, 2010)). If one was designing for these systems, a WDA of the entire work domain (e.g., the physics of sonar and data processing pipelines) would be appropriate, capturing as much information as possible for consideration of novel designs. However, for exploratory work to establish the effectiveness of a design direction, as presented in this thesis, this would be an inappropriate level of detail. Abstracting away from specifics using the representations in the HMI allows the WDA to be more appropriately scoped to the level of detail and congruent with the exploratory nature of the work. Furthermore, this helped to address the issue of subject matter expert availability, cost, and time, addressing the constraining factor of these on the analysis, in line with step 2 above and potential downsides to the application of CWA (Vicente, 2002; Stanton *et al.*, 2013; Hou *et al.*, 2015). The entire work domains of Sonar and TMA are very complex and having access to enough suitable subject matter experts to address them in suitable detail would not have been feasible. By capitalising on the representative formats that had implicitly been included in the interfaces, it becomes possible to set a more appropriate boundary for the work domain, utilising a degree of abstraction to reduce the number of subject matter experts required. Outside of the context of this thesis, a reduction in subject matter experts required to populate an Abstraction Hierarchy is extremely beneficial, as companies might not have the expertise readily available, enough budget to hire lots of different subject matter experts, or means to contact the required subject matter experts (e.g., no direct contact with operators). Care was taken to thoroughly prompt for 'how' and 'why' factors in the work domain, to address the concerns that workers had become dissociated from their domain (Rasmussen, Pejtersen and Goodstein, 1994; Naikar, Hopcroft and Moylan, 2005). This was helped by the subject matter experts having training in the work domain as well (e.g., courses to understand sonar principles, and how TMA plotting could be completed in different ways).

The principle of describing categories instead of specific instances (Naikar, Hopcroft and Moylan, 2005) was still adhered to, with entities being referred to in generic terms. For example, there are multiple bearing tape instances in sonar, but the WDA only includes a generic representation for all of them. Including the HMI mainly affected the bottom two levels (Physical Objects and Object-Related Processes), when compared to the traditional approach, which would represent the submarine's operating environment (e.g., ocean, boats, sensors, etc) as Physical Objects and their affordances (e.g., carry sound, emit sound, and detect sound) as Object-Related Processes. The practical reasoning for this was to increase the value proposition of the Work Domain Analysis by maximising the opportunity possible usage. By using the existing HMIs as the analysis target, it was posited that the abstraction hierarchies could be used for the main purpose of designing new HMIs, but also be used as training and formative reference materials for the current HMIs. For example, operators could use the abstraction hierarchies to identify what a component could be used for, or their options for achieving a specific goal using the system. This would go beyond a training manual which details the system and how it should operate, instead allowing operators to understand how they could use the system to achieve their goals.

Alternatively, stakeholders developing the systems could be provided with a copy of the outputs to further their understanding of how the system operates. This could be especially useful when providing software engineers with a design to implement. As will be explored in more detail in Chapter 8, the provision of the original outputs to software engineers could be beneficial. Concerns regarding communication between human factors practitioners and software engineers are well recognised (Viller and Sommerville, 2000; Bruseberg, 2008; Baxter and Sommerville, 2010; Wells *et al.*, 2011; Dhukaram and Baber, 2016), especially with regard to maintaining the richness of conducted analyses. Furnishing software engineers with formative information about the system in the form of a CWA could allow them to understand requirements better (McIlroy and Stanton, 2012), especially if the requirements are typically expressed in descriptive terms and do not provide underlying information to understand why there are expressed as such.

The outputs could also be used by the practitioners themselves, to record how the current HMI works, in conjunction with other methods such as Hierarchical Task Analysis (HTA). This could be useful for Human in the Loop (HITL) experiments that compare performance, especially where training materials would have to be developed to allow for the inclusion of novice participants, as opposed to already trained subject matter experts. The abstraction hierarchies could be used to plan how the system could operate for novice participants, considering the amount of training that they would have. Finally, showing benefits using the first iteration may serve as leverage for gaining additional resources or access to expand and iterate the WDA.

5.2.2 Cognitive Work Analysis Creation

The Human Factors Integration Defence Technology Centre (HFI-DTC CWA) Tool versions 1.0 and 1.1 (Human Factors Integration Defence Technology Centre, 2007; Jenkins *et al.*, 2009) were used to produce digital abstraction hierarchies. Further software was developed by the author to validate the abstraction hierarchies and to automatically render them using Microsoft Visio Professional 2016 (and later 2019). The individual process for WDA and WCA is detailed below. The completed analyses were validated through submission of project reports, which were reviewed by domain experts, and as part of the validation process in Chapter 7.

5.2.2.1 **Work Domain Analysis**

The purpose of the WDA was to understand where improvements could be made to existing implementations of Sonar, specifically concentrating on the waterfall as a key leverage point. The analysis was bounded at the Sonar HMI for completeness, instead of focusing just on the sonar waterfall in isolation, which might cause factors pertinent to a redesign to be overlooked.

A key constraining factor was access to subject matter experts. While they were generous with their time when available, this was limited and had to be used optimally. Consequently, learning was conducted independently as much as possible, combined with research from the concurrently running ComTET project. This meant that information was not sourced from specific teaching events, but it was gained from research activities and information osmosis (i.e., on the job training) as the project continued. This included, but was not limited to: conducting experiments and research; reading reports, journals, and books; formal and informal conversations with experts; experimenting with software; observing operators where possible; and watching documentaries. Subject matter expert input was sought throughout the ComTET simulator’s creation to understand control room operation and layout (Roberts, Stanton and Fay, 2015). An additional source of information was project management meetings, which were held quarterly, with feedback being provided on work as it was completed, including the analysis efforts. The expertise provided was varied, ranging from command team leaders to submarine designers. Additionally, the subject matter experts were from a range of defence organisations, including the Royal Navy, Defence

Science Technology Laboratory, and BAE Systems. The continued variety of expertise and organisations facilitated an in-depth understanding of submarine control room operation, which contributed to the process, in addition to assisting with simulator creation itself. This was further supplemented by the research conducted for Chapter 3, and tutorial material (user manual, video guides, written guides) for Dangerous Waters.

At the point of creation for the Work Domain Analysis, the taxonomy of constraints from Chapter 3 was still a work in progress and the existing prompts to discover constraints present within the domain were used in conjunction with a list of expected constraints that one would expect to find from a Cognitive Work Analysis expert (Stanton). This was a starting impetus for the generation of the constraint taxonomy, as a gap was identified to update future applications of the method to use the constraint taxonomy for step 4. The analysis outcomes were not updated once the taxonomy was completed as they had been built using general categories of constraints provided by a leading CWA expert, who was familiar with the constraints that would be found. Naikar, Hopcroft and Moylan (2005) addressed this in their treatment of the method, noting that being apprenticed to an expert makes aspects of the process more accessible, which in this case was understanding categories of constraints that would be present before this was confirmed by the taxonomy.

The abstraction hierarchy was completed in the order recommended by Stanton *et al.* (2017a), starting from the top (Functional Purposes, Value & Priority Measures), moving to the bottom (Physical Objects, Object-related Processes), and then linking both halves using the Purpose-Related Functions. There were two different versions, an initial version using a traditional approach, and an updated version employing the principles from Section 5.2.1. The main difference between the two was the bottom two levels (Physical Objects, Object-related Processes).

The Functional Purposes in both versions were populated with station use-cases. These were derived from the understanding gained from the review of submarine control room operation in Chapter 4, along with discussions with submariner subject matter experts about what the core purpose for each role existing was. The result was a technical purpose for the role to exist (what it was needed for), along with the goal of supporting higher command activities, supporting the social concept that information is aggregated and abstracted through the command team in anticipation of its use by the Officer of the Watch (OOW).

Global submarine goals (remain safe, remain undetected, and complete the mission (Mack, 2003; Mewett, 2014; Fay, Stanton and Roberts, 2017)) were not included as Functional Purposes for both versions. Instead, they were added as Value & Priority measures, as station success can be evaluated by measuring fulfilment of the goals. Other measures were subsequently added, utilising input from submariner subject matter experts on how their role's success could be measured. Measures were prefixed with '*minimise*' or '*maximise*' as they were not easily quantifiable (Naikar, Hopcroft and Moylan, 2005), especially when presented outside of a task-based context as in WDA. If a task was provided, then specific parameters could have been furnished, but this would have been at odds with the principle of WDA being task-independent.

In the initial versions of the abstraction hierarchy, the Physical Objects were populated by items that had been identified from the pilot testing and knowledge familiarisation stage of ComTET. They were a combination of actual physical objects present in the system, such as sensors, and concepts, such as acoustic sonar information. These were subsequently connected to Object-Related Processes that detailed their task-independent affordances, and what they could be used for across all situations.

After the initial analysis was completed, it was realised that the alternative approach in Section 5.2.1 may yield a more complete analysis for the reasons outlined. Therefore, the bottom two levels were adjusted before continuing. For the Physical Objects level, each item in the interface was added in Western reading order (left to right, top-down). These items were generated from what could be given a specific name within the interface, such as specific buttons, or aspects of the waterfall. This ensured that the analysis was appropriately scoped and did not attempt to break the interface down into an irrelevant level of detail. This was also to ensure that items representing the work domain, a methodological choice for this analysis, were represented appropriately as a whole object. For example, a waterfall trace representing where an entity is in the water could be a representation of the entity itself, but the component pixels making up the line would not, and therefore not be useful in this context.

Items on the Object-Related Processes level were then changed to align with the new Physical Objects, consolidating affordances exhibiting high similarity. They described what would be possible with the Physical Objects present from a technical and social viewpoint, considering the actions that they could participate in, even if they were not currently utilised in that way by current doctrine. This was to bring out the formative aspect of the analysis, showing what the interface could be used for, without limiting the affordances to how it should or is used.

Finally, the Purpose-Related Functions level was constructed by listing functionality that linked purpose-independent processes (Object-Related Processes) to object-independent functions (Purpose-Related Functions). For both stations, this closely aligned with processes operators followed to achieve their aims, such as the Detect, Classify, Localise, Track (DCLT) initialism for Sonar (see Section 4.1.1.4). This functionality was expected from a Sonar station, but it was not tied to the Physical Objects, rather how the functionality they afforded was used.

The updated abstraction hierarchy was validated by submarine and human factors subject matter experts involved in the ComTET project. The experience of these experts ranged from novices to experts. Specific demographics were not collected as ethical approval was only gained to gather this information from people who volunteered to be part of studies, not people who were providing their expertise as part of the ComTET project. Subject matter expert experience was used to ensure the output was accurate and complete; human factors subject matter experts validated the output's construction, and domain-specific subject matter experts (technical partners and submariners) validated the content. subject matter experts were asked to clarify or review certain aspects, including during the validation visit to the Talisman training facility at HMS Drake, detailed in Chapter 7, with the final product being reviewed by the research sponsors. This review was conducted as part of a delivered report for the ComTET project, which detailed the CWA process for Sonar and TMA. Feedback was provided on the report by a technical partner and a military advisor, and the completed analysis outputs were accepted. Minimal changes were required throughout, such as small wording changes, and clarification of concepts that were represented. The contents of this report were adapted into the contents of this chapter and Chapter 6, which were again reviewed by a technical partner before publication.

5.2.2.2 **Worker Competencies Analysis**

WCAs were also constructed to understand what skills operators would require to effectively operate Sonar and TMA. A similar approach to Stanton and Bessell (2014) was used, but using a different level of the abstraction hierarchy; while Stanton and Bessell (2014) utilised Object-Related Processes (affordances of Physical objects) as situations, Purpose-Related Functions (Functions that can fulfil system goals) were used in this instance. This is because the specificity of Object-Related Processes in the Work Domain Analyses would have resulted in excessively specific detail, confounding the analysis.

A table was constructed with the taxa of the Skills, Rules, Knowledge (SRK Taxonomy) as column headers, and the Object-Related Processes nodes as row headers. For each row, the appropriate cell was populated as follows:

- Skill-based behaviour cells were populated using responses that reflected the behaviour of experts or people appropriately trained with the system.
- Rule-based cells were populated using actions that would be taken if the operator was reacting to stimuli using rules they had been provided.
- Knowledge-based cells were populated using prerequisite taught or experience-based theory that was required to carry out the task or could be utilised for handling unfamiliar situations.

Again, validation of the WCA was performed by subject matter experts associated with the ComTET project using the same process as the WDA outputs.

5.3 Work Domain Analysis

The initial abstraction hierarchy for Sonar is presented in Figure 22, and the subsequent completed version using the method from Section 5.2.1 that the analysis is based on is presented in Figure 23. They are both presented vertically so they can fit onto a singular page. This section will work through the different levels of the finished abstraction hierarchy, starting from the Functional Purposes.

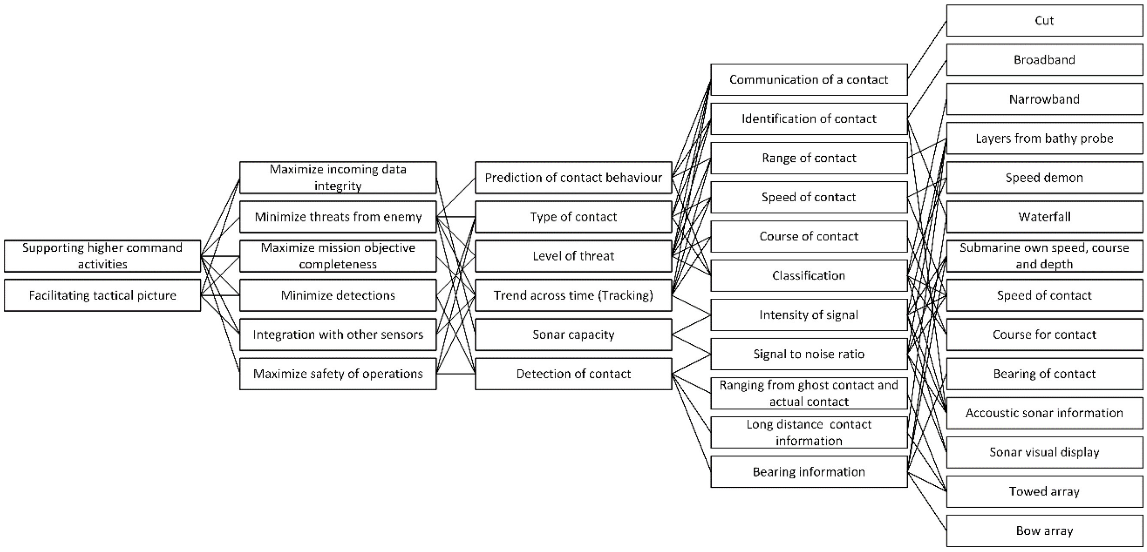


Figure 22 – Initial abstraction hierarchy of Sonar

5.3.1 Functional Purposes

Two Functional Purposes of Sonar were identified. They were to *'facilitate tactical picture generation'* and *'support higher command activities'*. Other stations contributed towards generating this picture, however, aural data from Sonar was particularly vital as there are limited detection methods that can overcome limited underwater visibility and maintain covertness. Hence, it was included in the submarine control room and had the purpose of supporting command activities, and providing information for the Officer of the Watch (OOW). These reasons for existing would not be changed by the introduction of a new interface, as this would require consideration of the entire control room sociotechnical system.

5.3.2 Value & Priority Measures

Two types of Values and Priority measures of success were present, global submarine goals, and local station-specific measures. Global measures of success are the submarine tenets (Mack, 2003), '*Maximise Safety*', '*Minimise Counter-Detection*', and '*Maximise Mission Objective Completeness*'. Of interest is that these goals are ordered by importance, with ownship safety being an absolute priority; therefore, any new design should support this goal as a fundamental consideration above all else. To '*Maximise Safety*', operators are required to detect other vessels, so that collisions can be avoided. This is closely related to '*Minimising Counter-Detection*', where operators avoided vessels that were equipped to detect the submarine and ensure the submarine was not emitting excessive noise. Completion of both goals, in addition to providing required information about other vessels facilitated '*Maximising Mission Objective Completeness*'.

There were local goals specific to Sonar, which were '*Maximise Contact Detection*', '*Maximise Signal Clarity*', and '*Maximise Known Contact Information*'. To '*Maximise Signal Clarity*', operators can change sensors and settings to provide the clearest picture of their surroundings. In doing so, they could '*Maximise Contact Detection*' by making the noises from surrounding vessels easier to detect. With the signals easier to detect, and therefore analyse, they could '*Maximise Known Contact Analysis*'. It is clear from these goals that the success of Sonar can be measured be how well they collect data and process that data into information. This suggests that any redesign of Sonar would need to ensure that operators can manage data effectively, and that they are provided the tools to process data into information.

Most goals do not contradict any others, although there is a potential disagreement between *'minimise counter-detection'* and the local goals of sonar aimed at increasing the information being detected about contacts. They can all be achieved at the same time, although getting clearer information on a contact might require manoeuvring, deploying the towed array, or another action that increases that chance of counter-detection. This is mostly addressed by having more capable sensors that do not require such actions, although this will place a requirement on operators to manage the increased information. This does not change lower levels of the abstraction hierarchy, but does provide an impetus to move towards information representations that reduce the cognitive workload required to ensure that ownship safety is maintained.

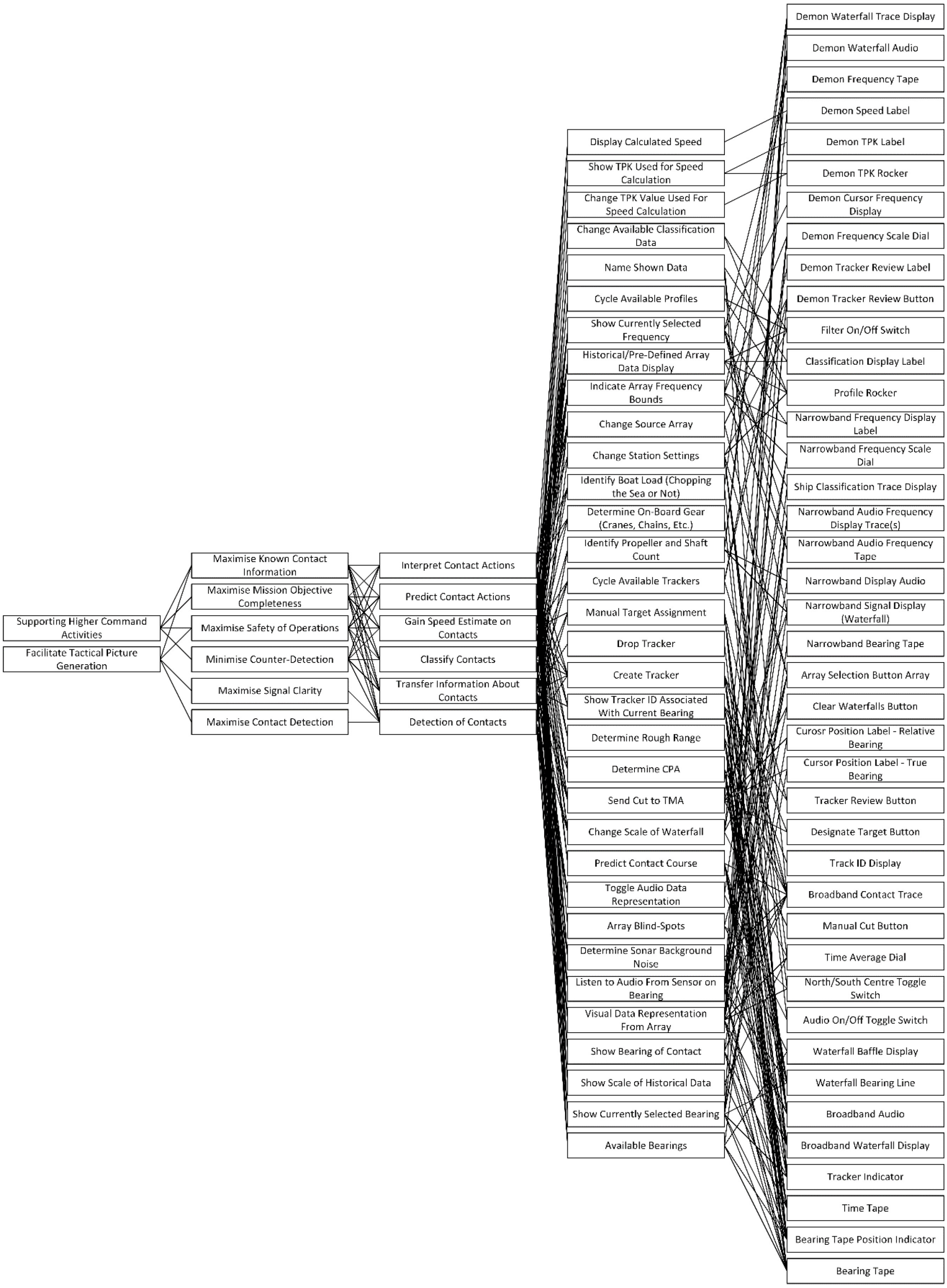


Figure 23 – Full Abstraction Hierarchy of Sonar

5.3.3 Purpose-Related Functions

Each global goal was connected to all Purpose-Related Functions. Conversely, local goals each connected to a subset of Purpose-Related Functions. This indicated that functionality was not included if it did not facilitate meeting submarine tenets, and local goals did not dictate function groupings. Instead, functions were correlated with components of the DCLT (Detection, Classification, Localisation, and Tracking) initialism, the four main functions of Sonar (Hughes *et al.*, 2010). Stages were concurrent but could occur in parallel for multiple contacts; operators could detect signals from a contact, classify these signals to understand the emitter, localise the emitter to a specific location, and then track its movements. The order was not prescriptive, and operators would structure their work according to their goals. Stages of DCLT were present as Purpose-Related Functions, with nodes either being a direct equivalent ('*Detection of Contacts*' for Detection), or components '*Interpret Contact Actions*' and '*Predict Contact Actions*' for Tracking.

Thus, it is evident that DCLT had been used to subset functionality within the system. Such a design choice is logical, although it does require operators to memorise information about contacts and switch between multiple screens of information to complete their activities, which could increase cognitive load (Michailovs *et al.*, 2022) and saturate limited working memory (Cravens, 2021). More screens have been added with the aim of improving the information available to operators for decision making (Chalmers, Easter and Potter, 2000), although they might not have been integrated in the most optimal fashion (Hall, 2012). While the effect of multiple screens is partially offset by information being carried between screens and trackers being used to alias information across the screens, there is still opportunity for improvement. Therefore, it is proposed that a redesign could better support operators in achieving their goals by orienting the redesign around contacts, and collocating information to do so, as opposed to separating the information along fixed lines of functionality. However, it might be pertinent to still provide some degree of focus for operators, as Michailovs *et al.* (2022), citing Posner (1980), argue that this could form a natural 'attentional spotlight'.

5.3.4 Physical Objects and Object-Related Processes: How does the waterfall work, and how could be improved?

The DCLT initialism manifested itself as five separate screens, of which three were used: broadband, narrowband, and DEMON (Detection Envelope Modulation On Noise (de Moura, de Seixas and Ramos, 2011)/Demodulation of Noise (Mill and Brown, 2005)). The last two screens, active and sound speed profile, were not used as they were beyond the scope of experiments being conducted in this research. Active sonar is seldom used by submarines due to their need to remain covert, and was therefore not used in the conducted experiments. The sound speed profile screen is used for more advanced sonar operations, such as understanding the paths of sound through the water to optimise detections. This capability was not selected for use in experiments as these competencies were not being assessed (i.e., whether operators could optimise the sonar system for a specific area). Instead, it was assumed that the sonar was working optimally, with the focus on how participants would construct a tactical picture, which could be readily achieved using the first three screens. Operators could switch between each screen using on-screen buttons. Screens maintained their state, allowing operators to switch between them as necessary without losing data. This allowed them to work across various screens without noting information when switching, saving time and cognitive capacity. However, this still required the capacity to remember information between screens. Screens shared core component types, facilitating generalizable knowledge of other screens when one is understood. Of these common components, the waterfall was the most prevalent, appearing in all screens.

As a reminder from Section 4.1.1.4, a waterfall displays aural information visually over time, see Figure 24a. The horizontal axis is typically bearing, although it can be frequency, and the vertical axis is time. When Sonar sensor data from available bearings is received, the interface displays it in a single horizontal line at the waterfalls' top. Marks are drawn at their Direction of Arrival (DOA) to show received noise. This received noise includes background noise and noise being emitted by objects. New data is drawn above previous data, moving it down (see Figure 24b and Figure 24c), creating the appearance of a waterfall, hence their name. Background noise appears as specks, the concentration of which is directly correlated with the background noise. Over time, vertical lines can form, allowing operators to identify nearby sources of sound.

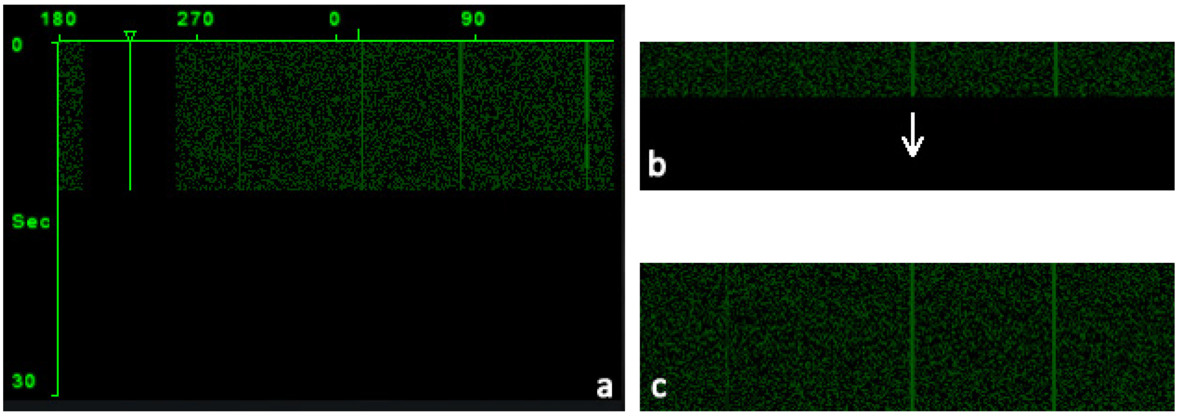


Figure 24 – a) A Sonar waterfall, b) Aural data moving down the waterfall c) Aural data filling the entire waterfall, showing current and historical data

To aid understanding of where improvements could be made and how, components of the waterfall were isolated from the main abstraction hierarchy to understand what functionality they afforded and how they did so. Isolation of the nodes was performed to make the nodes and links more apparent as compared to the entire abstraction hierarchy. Following means-ends links in the whole abstraction hierarchy could have revealed this information, however, due to the size of the full hierarchy, they have been presented separately. By only extracting the relevant nodes, the nodes, as well as links, were much more accessible and understandable. The Physical Objects being examined are the bearing tape, time tape, contact trace, tracker indicator, and audio. The broadband waterfall was chosen to present for the waterfall examples, as it is more accessible to demonstrate without a detailed understanding of Sonar. For each Physical Object, an abstraction hierarchy specific to the object is presented, along with a description of current work, and what design ideas are yielded for a redesign of Sonar as a whole (i.e., not simply redesigning the Physical Objects in a 1:1 manner). The definitions for all nodes are provided in Table 15 to avoid repetition. After isolating the relevant nodes for each of the five main components of the waterfall, a holistic overview of the waterfall in-situ will be provided, alongside further ideas for improvement in a redesign.

Table 15 – Definitions of nodes at each level (Functional Purposes → Physical Objects) contained in the isolated waterfall abstraction hierarchies presented in the sections below.

Level	**Node**	**Definition**
Functional Purpose	Facilitate Tactical Picture Generation	Support the creation of a picture that depicts what is currently around the submarine
	Supporting Higher command activities	Ensure that senior posts can carry out their duty
Value & Priority Measures	Maximise Contact Detection	Ensure all possible contacts are detected
	Maximise Known Contact Information	Provide as much information as possible about contacts
	Maximise Mission Objective Completeness	Contribute towards completing the assigned mission
	Maximise Safety of Operations	Ensure that the submarine is always safe
	Maximise Signal Clarity	Ensure that signals received can be interpreted by operators
	Minimise Counter-Detection	Reduce the ways the submarine can be detected
Purpose-Related Functions	Classify Contacts	Functionality that enables operators to classify contacts
	Detection of Contacts	Functionality that enables operators to detect contacts
	Gain Speed Estimates on Contacts	Functionality that enables operators to make predictions about a contact’s speed
	Interpret Contact Actions	Functionality that enables operators to understand a contact’s behaviour
	Predict Contact Actions	Functionality that enables operators to predict a contact’s actions
	Transfer Information About Contacts	Functionality that enables operators to communicate information about contacts
Object-Related Processes	Available Bearings	Which bearings are currently accessible to be analysed
	Create Tracker	Assigns a tracker to a detection, so an operator does not have to manually follow it

	Determine CPA	Allows for the calculation of a contact's Closest Point of Approach (CPA)
	Determine on-board gear (Cranes, Chains, etc.)	Allows an operator to determine what equipment a contact has
	Determine Rough Range	Can be used by an operator to calculate the rough range of a contact
	Determine Sonar Background Noise	Shows an operator how noisy the background is, facilitating calculation of a Signal to Noise (SNR) ratio
	Drop Tracker	Stops the automatic tracking of a contact
	Identify Boat Load	Allows an operator to understand how loaded a contact is
	Listen to Audio from Sensor on Bearing	Plays aural data directly from the array (sensor)
	Manual Target Assignment	Allows an operator to explicitly assign a tracker/send a cut
	Name Shown Data	Aliases a specific detection pattern so it can be referred to by the command team
	Predict Contact Course	What course a contact will take, based upon current data
	Show Bearing of Contact	The current bearing of a contact
	Show Currently Selected Bearing	Shows which bearing the operator is analysing
	Show Scale of Historical Data	Enables an operator to understand the timespan associated with historical data
	Show Tracker ID Associated with Current Bearing	Shows the tracker ID related to a given bearing
	Visual Representation of Data from Array	Data from the current array (sensor) that is represented visually
	Bearing Tape	A waterfall axis, showing possible bearings
Physical Objects	Bearing Tape Position Indicator	A small line that appears on the 'Bearing Tape', showing which bearing is being inspected
	Broadband Audio	Current aural data from the current bearing being played
	Broadband Contact Trace	A line formed from the ongoing representation of sound above background noise. All noise is represented as dots on the waterfall, with a coalescence of dots indicating a source of sound
	Time Tape	A waterfall axis, showing the scale of historical data
	Tracker Indicator	Text on the 'Bearing Tape', showing the presence of a tracker. For Sonar the text starts with 'S' and is followed by an incremented number

5.3.4.1 **Bearing Tape**

Bearing tapes were one possible horizontal axis of waterfalls, representing all bearings 0° - 360° as a flat strip. These angles are true bearings, using true north as 0°. This means that the tape did not move with ownship's course. The '∇' icon on the tape represented the stern of ownship, allowing operators to understand ownship's course, which in turn allowed them to determine the relative location of contacts. Operators could move their mouse over each bearing to listen to the audio. The '|' on the bearing tape represented where the operator was currently listening, inspecting, or analysing. Figure 25 shows an example bearing tape, with associated AH nodes. Using the '*Bearing Tape*' an operator could identify the '*Available Bearings*' on a waterfall. In combination with the '*Bearing Tape Position Indicator*', which '*Showed Bearing of Contact*' they could determine a contact's exact bearing. Together, this allowed an operator to '*Transfer Information about Contacts*', therefore '*Maximising the Safety of Operations*'. By ensuring the safety of ownship, the operator was '*Supporting Higher Command Activities*'.

While the representation of bearing data was compact, being represented in a straight line, this does not represent the actual bearings, which could be represented as a circle extending from ownship as an origin on a map display, similar to the schematic representation provided by in Shar and Li (2000; Figure 1). Adopting this representation of bearing information may increase the space required to display the information, but could make the information more accessible to operators by capitalising on a skeuomorphic representation of the bearings, not requiring a mental transition from the flat bearing strip to their internal view of the tactical picture and where contacts are. This could further support adherence to the Value & Priority Measures by reducing the amount of cognitive workload that operators require to interpret the current tactical picture. Such a representation is likely to require a map-based view, suggesting a direction for the organisation of any redesigned interface to reveal the functional aspects of the work domain.

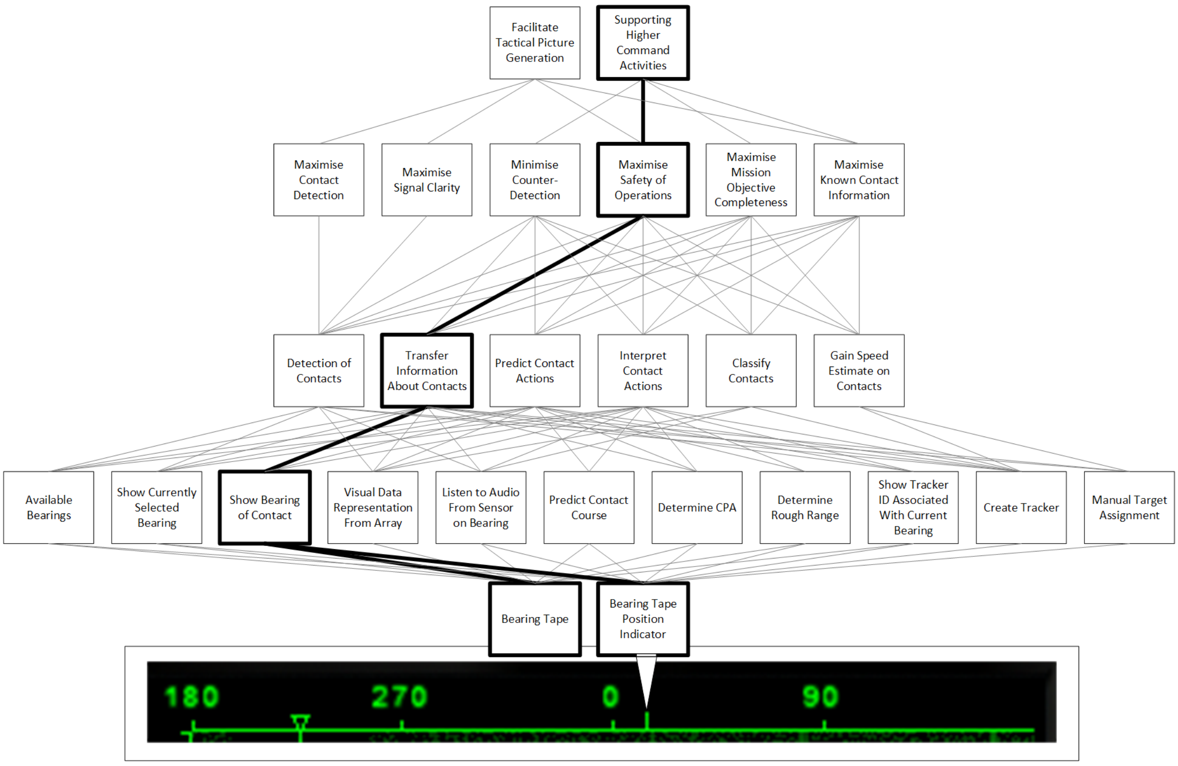


Figure 25 – Bearing Tape Abstraction Hierarchy

5.3.4.2 Time Tape

Time tapes were the vertical axis of waterfalls, representing a given historical timeframe. Present time was shown at the top, and the maximum time displayed was shown at the bottom. The unit of time was marked in the middle to provide scale. Data associated with a point in time was horizontally aligned with a given point on this axis, allowing operators to understand the position of contacts at that time. Over time, the history formed a trace on the waterfall, which could be used to understand more about a contact. Figure 26 shows an example time tape with associated AH nodes. The '*Time Tape*' '*Showed the Scale of Historical Data*' to an operator. They could use this to '*Transfer Information about Contacts*'. Combined with other data, this allowed the command team to '*Maximise Known Contact Information*', therefore '*Facilitating Tactical Picture Generation*'.

The time tape appears to be a sensible reflection of history, although its representation is tied to that of the bearing tape. If the bearing tape changes to become a skeuomorphic representation, then that will remove the possibility of a time tape, potentially losing functional information about the data in the work domain. Consequently, this will require different Physical Objects to represent the data. Careful thought should be given to this, as the operator must be able to differentiate between past, current, and future data (Bennett, Payne and Walters, 2005). This is especially important as the abstraction hierarchy reveals that the time tape contributes to transferring information about contacts. If the information is not temporally suitable, such as communicating an old position, then it could endanger ownship, violating the goal of *'Maximising Safety of Operations'*, which is the top goal.

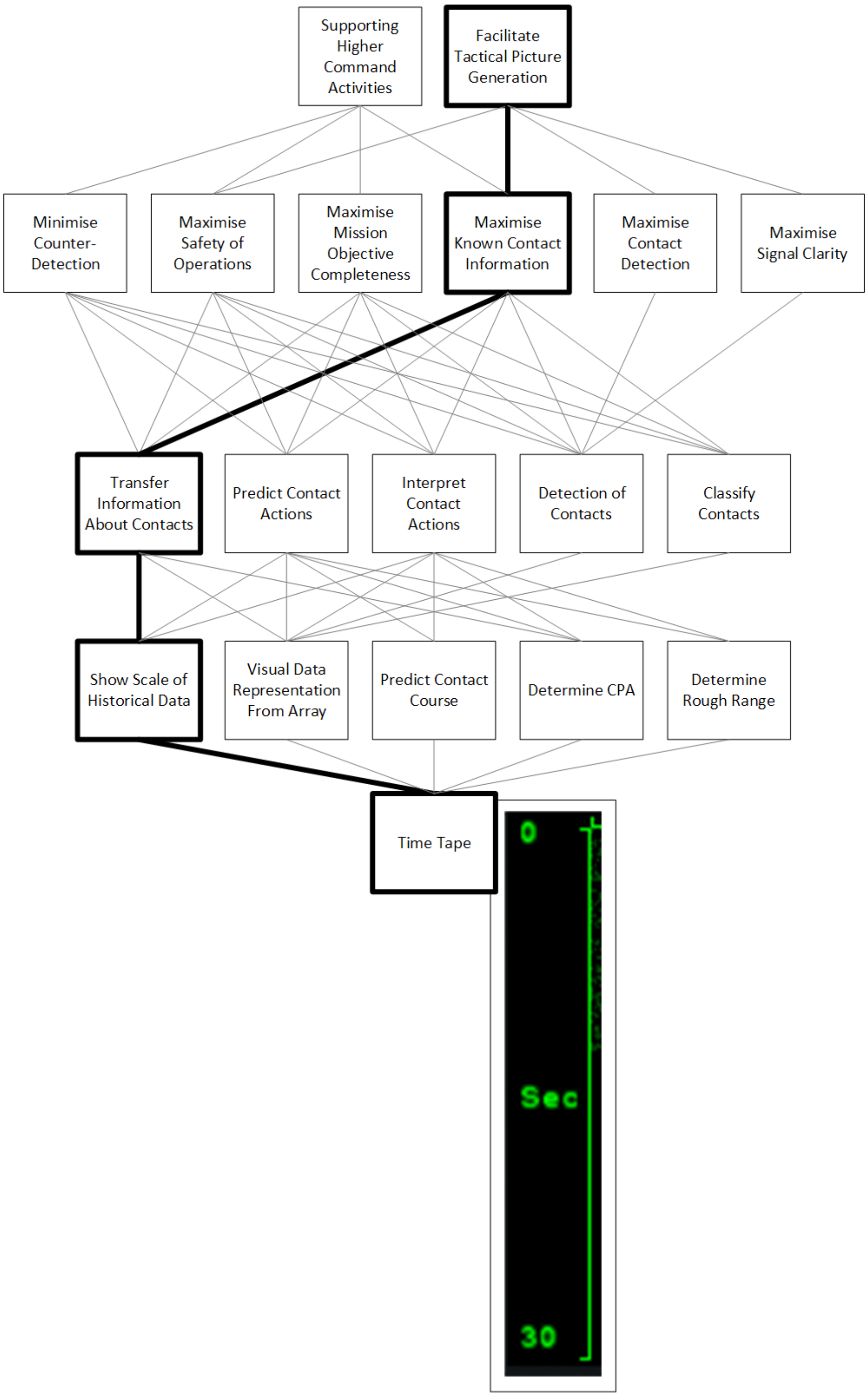


Figure 26 – Time Tape Abstraction Hierarchy

5.3.4.3 **Broadband Contact Trace**

Solid lines, called traces, could appear on a waterfall display, indicating the presence of sound-emitting objects, which could be a ship, submarine, or biological entity. For a broadband waterfall, these occurred because of a concentrated visual representation of sound over time in the waterfall. An increased level of sound compared to background noise was shown by a higher concentration of polygons, which over time drew a line across the waterfall. Figure 27 shows an example trace from a broadband waterfall with associated AH nodes. A '*Broadband Contact Trace*' was the visual representation of underlying data, for which operators could '*Create a Tracker*'. Trackers could be used to '*Interpret Contact Actions'* and *'Predict Contact Actions'*. In doing so, the command team could ensure that they maintained a safe distance in order to '*Maximise Safety of Operations*'. This '*Supported Higher Command Activities*' as the Captain is ultimately responsible for a submarine's safety, and as such any information that supports this is vital.

The traces are well designed, as they can display a lot of information that supports information at higher levels of the abstraction hierarchy. In addition to being plotted against the bearing and time tapes to show those variables, they also utilise thickness and colour intensity to represent signal strength and how clear that signal is. This provides functional information to the operator, as it allows them to deduce information about the current status of the environment, such as how close entities are, or how fast they are travelling. This suggests that any redesign should seek to maintain, or even improve information provided by traces, permitting the operator to reason about the higher levels of the abstraction hierarchy. Furthermore, there is evidence that the colour and luminosity of visual sonar information can improve detection capability (Dawe and Galbreath, 1997). This should be capitalised on in any redesign to ensure that information is as salient as possible for operators.

Additionally, when displayed in the rectangular format of the waterfall, it might be easier to interpret contact behaviour as opposed to a spatially accurate representation (a circle extending from ownship). As the data is normalised onto the waterfall against the provided axes (bearing and time tapes in this instance), the behaviour could be predicted in a linear fashion, such as using a ruler to extrapolate the direction of a contact trace to identify its future bearing. Therefore, the redesign should seek to keep this level of utility in any redesign. This is especially important considering that the operator might focus on the traces as a representation of their work environment, so as much information as possible should be available to ensure that pertinent information is not missed.

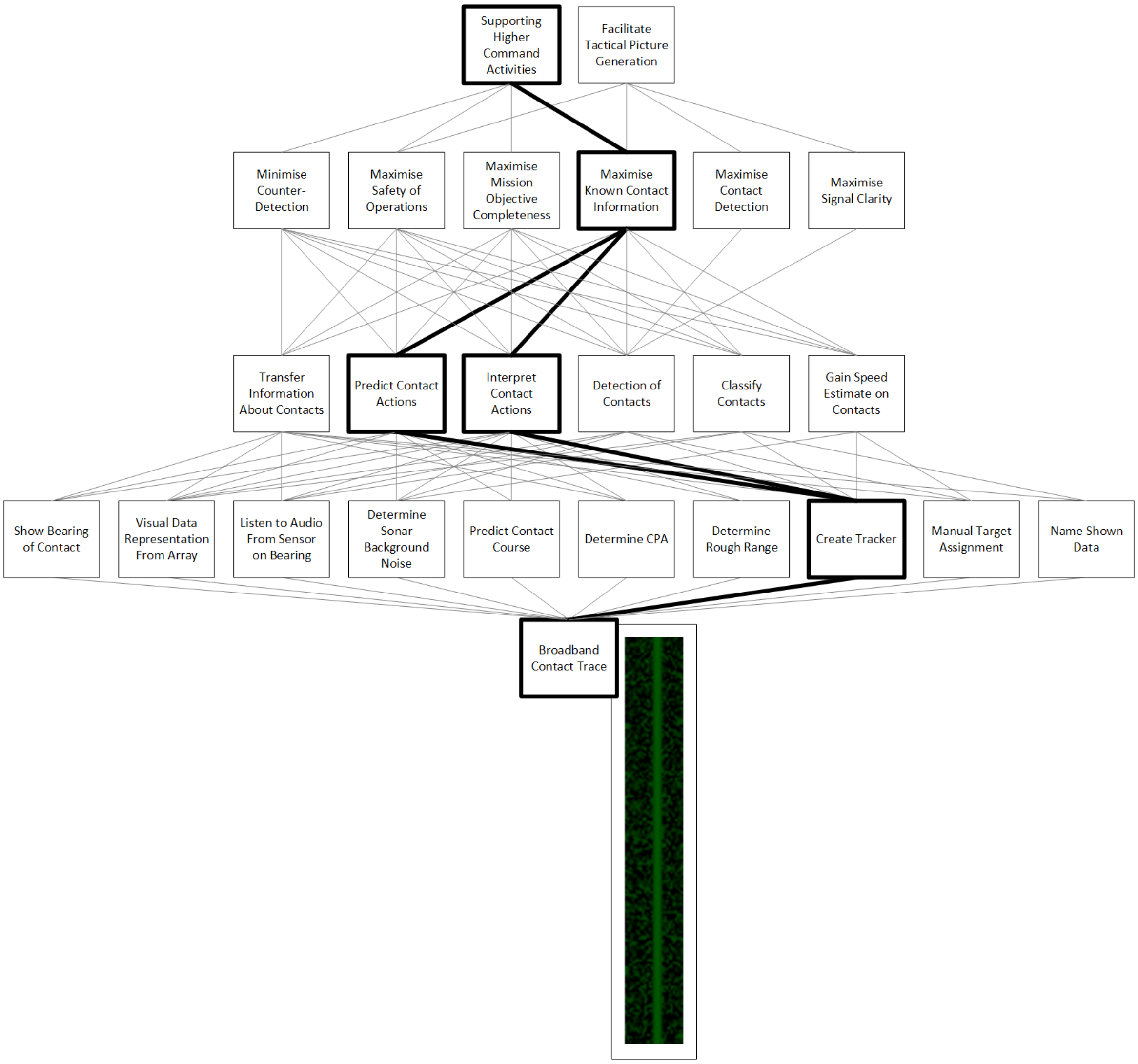


Figure 27 – Broadband Contact Trace Abstraction Hierarchy

5.3.4.4 **Broadband Audio**

In addition to the visual representation of aural data, waterfalls played audio directly from the array (sensor) to operators. For broadband audio, it would sound like listening to sounds underwater. Operators could detect vessels that have not made a trace yet or listen for more information on a vessel. For example, if clanking chains could be heard, this could indicate the contact is a fishing vessel. This is important as a trace alone would not reveal this information. Figure 28 shows associated nodes for broadband audio. An operator could '*Listen to Audio from a Sensor on a Bearing*' using '*Broadband Audio*' to assist with the '*Detection of Contacts*'. The combination of listening for contacts and visually searching for traces '*Maximised Contact Detection*', providing information that '*Facilitated Tactical Picture Generation*'.

As with the visual trace, the audio is an important aspect of Sonar. It reveals counterpart functional information about the contact, utilising the different modalities to provide information that corroborates and builds on visual information displayed in the visual trace. For example, the audio trace could allow the operator to classify the contact. This is useful information by itself, although can reveal the clues about the constraints currently acting on the submarine's environment, such as staying a safe distance from fishing vessels, or expecting the presence of an escort, revealing that more contacts might be present. A bimodal representation of sonar data can also produce a reliable advantage over one modality alone (Doll and Hanna, 1989; Lewandowski and Kobus, 1989). This suggests that the redesign should capitalise on this, maintaining synergistic links between the information represented by a trace and its counterpart audio. Furthermore, Watson and Sanderson (2007) demonstrated that auditory data may be better suited to supporting skill-based behaviour over visual data, owing to its transitory nature. As the audio is a live feed, it represents present states in a natural manner, allowing operators to naturally interpret it over time.

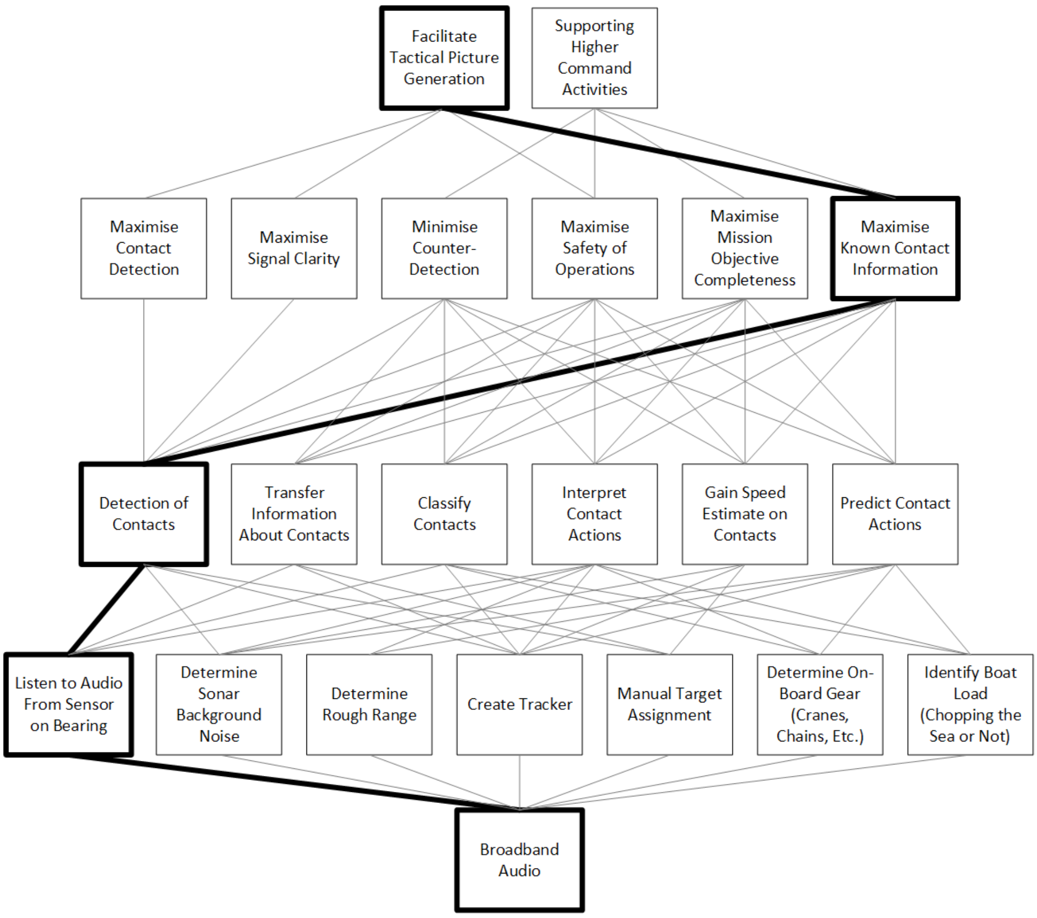


Figure 28 – Broadband Audio Abstraction Hierarchy

5.3.4.5 **Tracker Indicator**

When a trace or sound was identified on the waterfall, the operator could assign a tracker. The tracker appeared as text in the bearing tape, showing an incremental number, prefixed with '*S*' for

Sonar. For example, the first and second trackers assigned would be 'S01' and 'S02' respectively. All trackers followed the bearing of their underlying signal so long as they could detect it. This freed the operator to perform analysis without having to manually collect data. However, the number of active trackers was limited to eight per sensor. If more trackers were required, the operator either had to wait for one to become available or drop an unused tracker. Trackers could be dropped by dragging and dropping them onto the waterfall area. This was not done often and occurred for contacts no longer of interest. At a pre-specified interval, each tracker would automatically send a cut (the exact bearing) through to TMA, who would perform further analysis to determine the contacts' bearing, course, range, and speed. A signal with a tracker was termed a contact, an external entity of which the command team was aware. Figure 29 shows an example tracker indicator with associated AH nodes. A '*Tracker Indicator*' allowed an operator to '*Name Shown Data*'. Using this alias allowed operators to '*Transfer Information about Contacts*' by providing a descriptor for a specific contact. By associating all contact information with a tracker alias, the command team could '*Maximise Known Contact Information*', which '*Facilitated Tactical Picture Generation*'.

The tracker indicator is a Physical Object representation of the concept of trackers that is adhered to around the control room, passing information using the assigned alias, as opposed to 'the contact at bearing …' or a similar alternative. This means that information becomes 'attached' to the tracker, increasing an operator's cognitive workload to maintain situational awareness of this information. However, the interface does not currently support this, with information that operators require being displayed across multiple screens, as will be shown in Section 5.3.5. Furthermore, this also places requirements on an operator's limited working memory, requiring them to remember information from across the interface if they are seeking to understand higher-level information about a contact, such as *'transferring information'* to *'maximise known contact information'*. This suggests that operators could be better supported in achieving these goals by representing summary information about contacts in proximity to their tracker, reducing the cognitive workload and working memory requirements, in addition to making the underlying data structure of information for contacts more visible. The latter is important as this is not readily achieved in current submarine control rooms, with information being distributed across the control room.

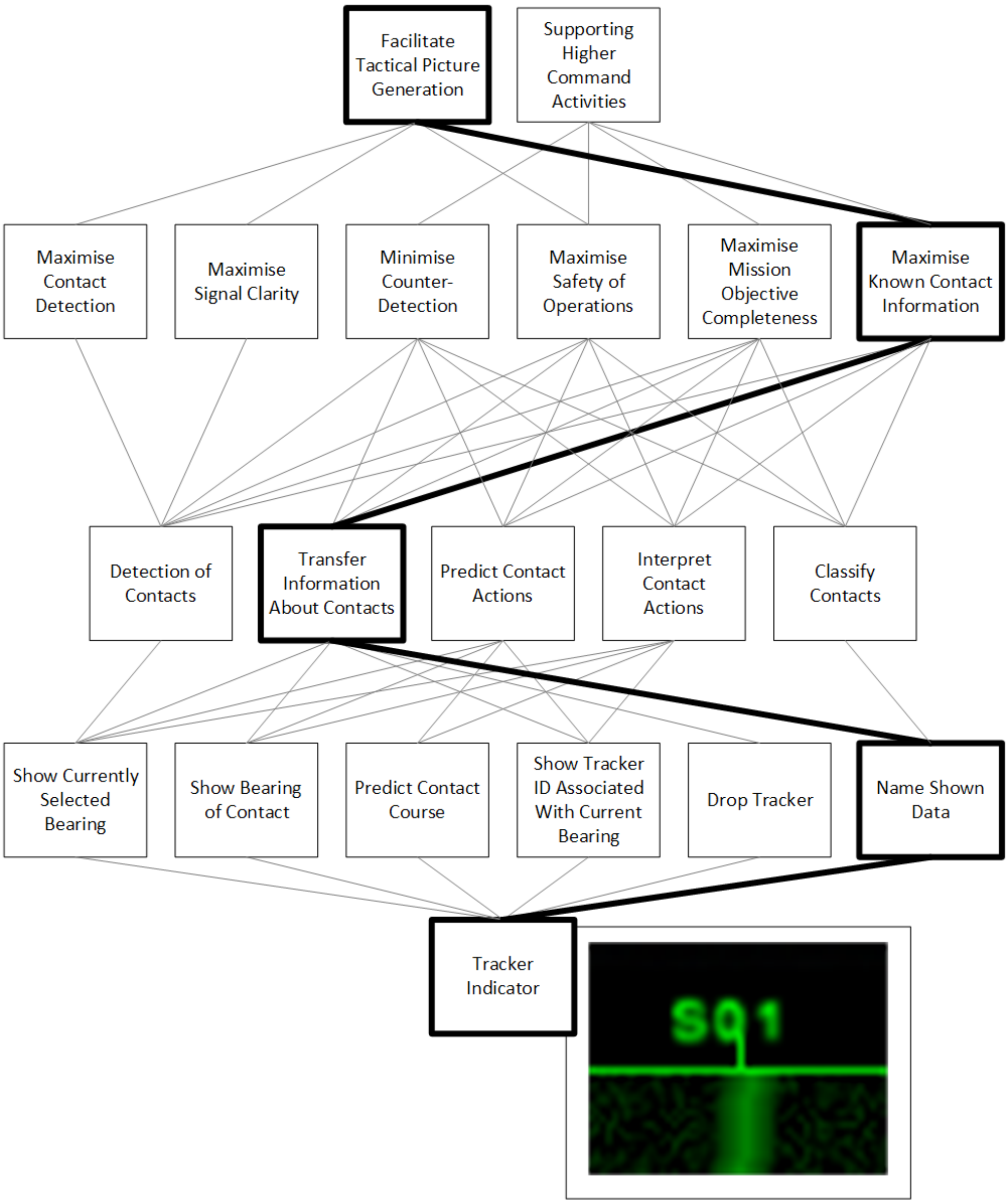


Figure 29 – Tracker Indicator Abstraction Hierarchy

5.3.5 Operation of the Sonar screens

It can be seen in Figure 30 that there are four broadband contact traces visible, on bearings ranging from around 290 to 170. Each trace has been visible for at least thirty seconds. The operator is currently looking and listening to a bearing of ~170° (under 'S01'). The trace at this bearing has been assigned the tracker 'S01'. While polymorphic waterfall implementations exist, their core objects and functionality, expressed by each isolated abstraction hierarchy above, remain consistent. An understanding of waterfall functionality facilitates a greater understanding of the three Sonar screens they appear (see Figure 31 – Figure 33), thus providing operators and system designers with an understanding of the Sonar operation and where improvements could be made.

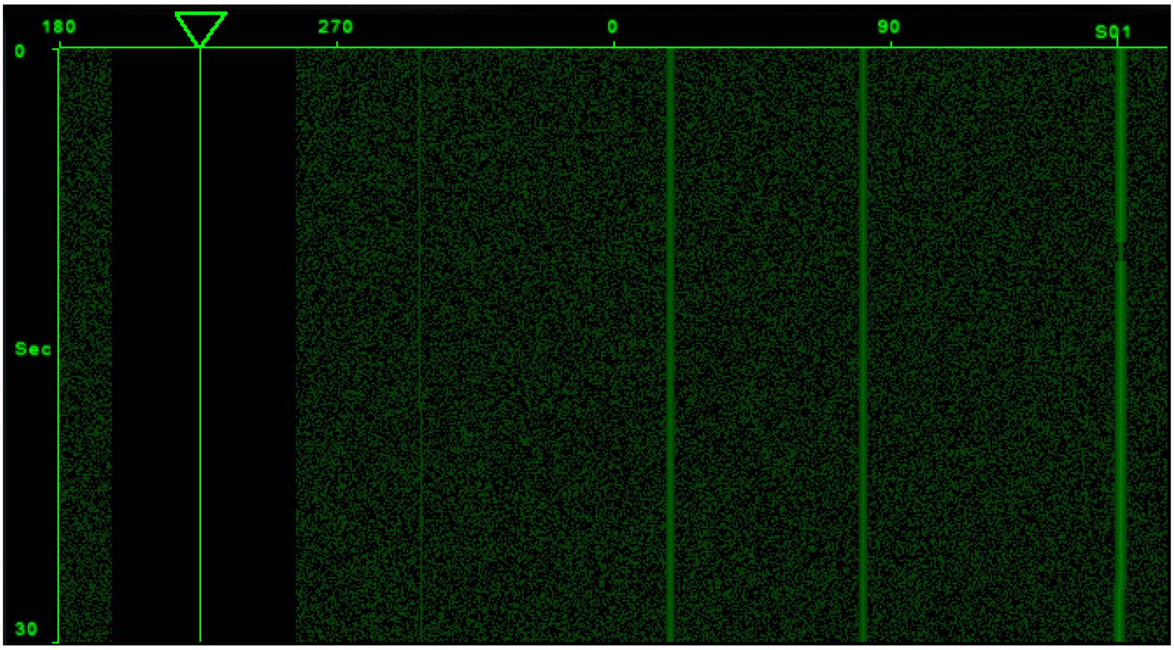


Figure 30 – The Waterfall as a display, with four traces shown. Note that 0° is at the centre of the display.

5.3.5.1 **Broadband Screen**

Figure 31 shows the ComTET Sonar Broadband screen, bounded with areas of interest. Area "a" highlights two discrete waterfalls. The direction of ownship's stern is marked by '∇' and the line below it. Black space on either side of the line shows where the submarine is and hence a signal cannot be received. These voids can also occur when a sensor's field of view does not include a range of bearings. In both cases, they are called baffles. An exact bearing for the bearing tape position indicator is shown in the '*Cursor Position*' textbox in Area "d". Area "b" enables the operator to toggle playing aural data and changes whether 0° or 180° is at the bearing tape's centre. This does not affect the functionality of the interface but can make it easier to interpret data. Areas c and e change the time tape's scale for the top and bottom. Area "f" changes the array that populates the waterfalls. Area "d" handles contact management. The '*Designate Target*' button allows operators to assign a tracker. Operators can manually send data at any time for a tracker using the '*Manual Cut*' button. If an operator is inspecting a bearing with a tracker, it will show in the '*Track I.D*' text box. '*Tracker Review*' allows an operator to jump between trackers. '*Clear Waterfalls*' removes all history from both waterfalls, and is sometimes used in conjunction with course changes, or bearing centre changes to prevent confusion arising from old data being displayed.

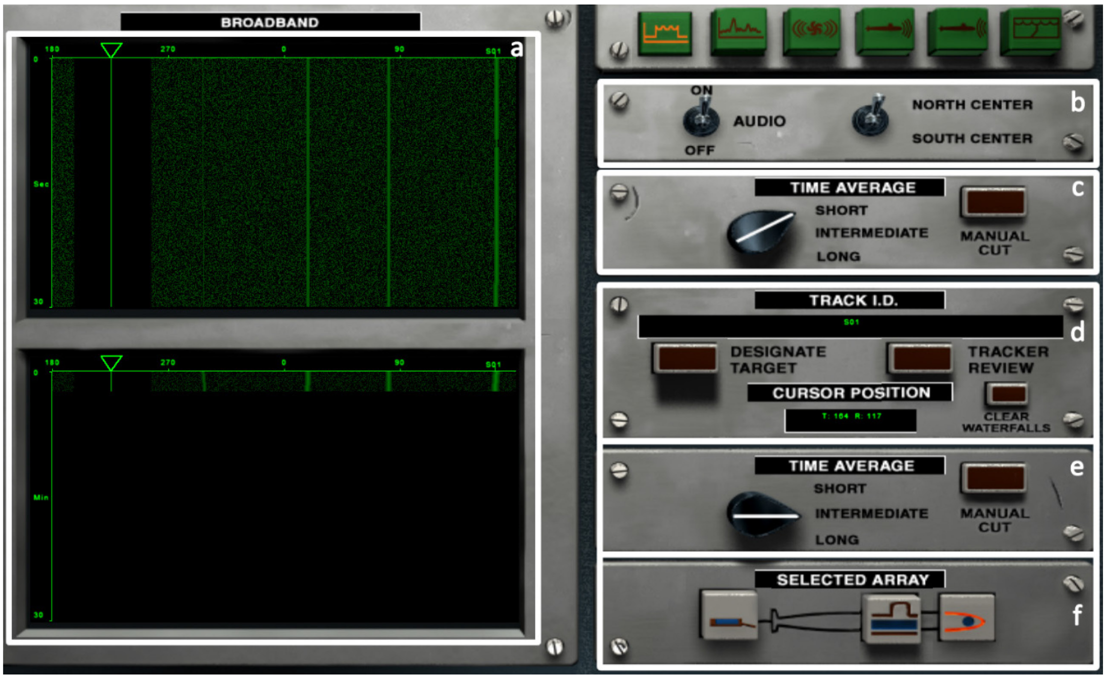


Figure 31 – Sonar Broadband User Interface, with ownship stern marker ('∇') enlarged for clarity

5.3.5.2 **Narrowband Screen**

Figure 32 shows the ComTET Sonar Narrowband screen. Area "c" is a Low-Frequency Analysis and Recording 'gram' (lofagram; de Moura, de Seixas and Ramos, 2011). '*Gram*' as a suffix signifies something being written or recorded in a certain way. This resembles a broadband waterfall, with a frequency tape replacing the bearing tape. For a specified direction of arrival for sound, the audio is split into its constituent frequencies and plotted on the waterfall. These frequencies can be unique to known contacts, being emitted from known models of machinery used only for a specific class of vessel. This allows comparison to the gram in area "b", which shows the known frequencies of the selected contact from area "e". Area "d" changes the scale of frequency tapes for areas '*a*' and '*b*', permitting an operator to see small frequency ranges clearly, and to fit large frequency ranges. The currently selected frequency is also displayed. It is possible to filter the profiles that can be selected using an algorithm, although this may be inaccurate, so it is possible to cycle through all profiles. Area "a" resembles the broadband waterfall. However, it does not have a time tape, displaying all information in real-time with no history. Larger peaks than surrounding areas correlate to traces on broadband and selecting that bearing will display the signal breakdown in area "a". It is possible to assign a tracker in narrowband, but trackers are usually assigned in broadband, as this is where a contact will first be detected.

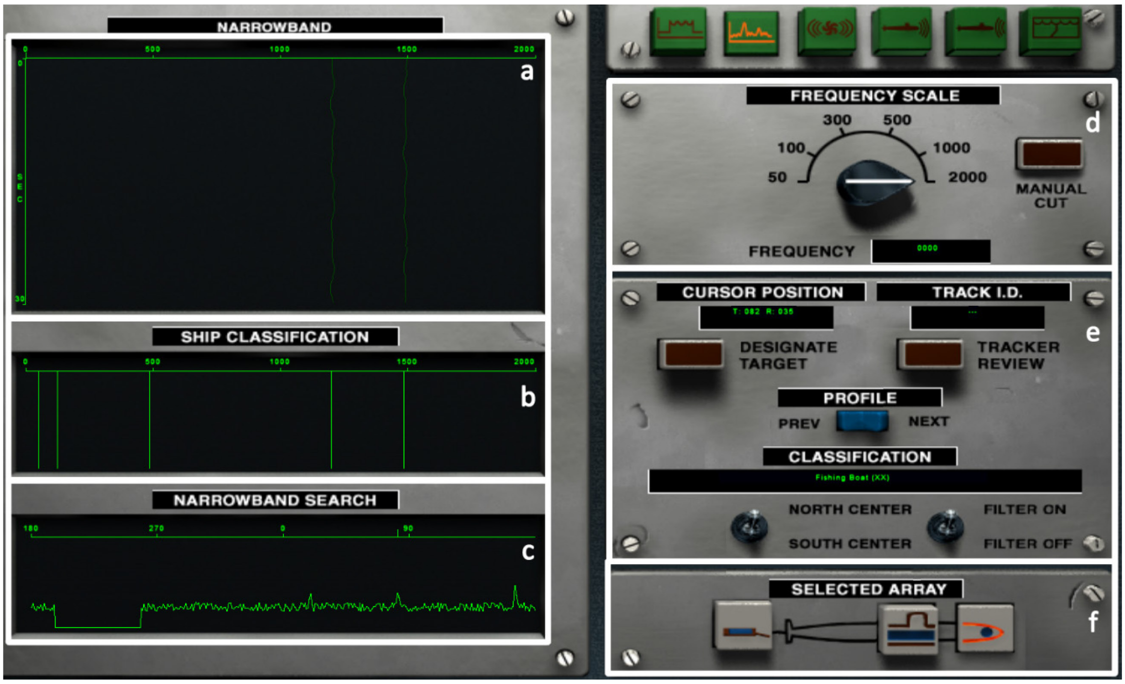


Figure 32 – Sonar Narrowband User Interface

5.3.5.3 **DEMON Screen**

Figure 33 shows the ComTET Sonar DEMON screen. Area “b” splits a signal into its components as in area “a” in the narrowband screen but concentrates on propulsion. In the figure, the leftmost line represents a shaft, and the other three represent attached blades. Multiple shafts are represented with a horizontal gap between groups of signals. The signal being inspected is changed from area “a”. Area “d” calculates speed using the frequency of a shaft, combined with an entered Turns per Knot (TPK) value. A TPK value is the ratio of propeller rotations to speed for a given vessel, or type of vessel (Dryer, 2002). This figure is derived by taking a classification from narrowband or another source, retrieving a stored value from a dataset, and entering it using the rocker. The speed textbox will automatically update as the TPK value changes.

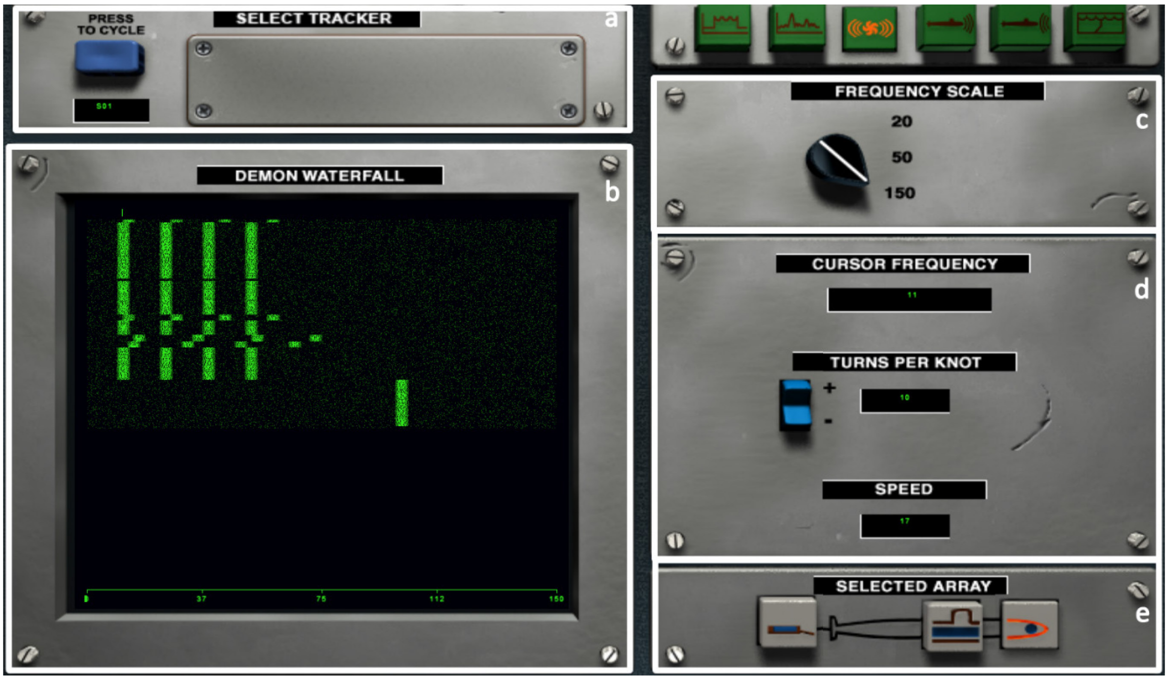


Figure 33 – Sonar DEMON User Interface

5.3.5.4 **Issues with Operating the Screens**

From examining how the operator could use the available Sonar screens, the issue of information being spread out across the interfaces, alluded to in Section 5.3.4.5, becomes apparent. Operator cognitive workload will be increased from managing data across different screens, which also places requirements on their working memory. A common strategy that operators use to address this is to create a table of information on a whiteboard or paper, and then record information to this as it becomes available. While this can reduce working memory utilisation, it places a requirement on the operator to maintain their records so that they are synchronised with the current data. This introduces two related potential points of failure. The first is that they might use stale information from their record, potentially negatively affecting attainment of Value & Priority Measures such as *'Maximise Known Contact Information'* and *'Maximise Safety of Operations'*. The importance of information being passed correctly is highlighted by Stanton and Roberts (2018) finding that 'bearing', 'course', 'contact', 'speed', and/or 'knots' are consistently passed between operators in scenarios. Each piece of information can be determined or inferred from sonar, creating a central point of information degradation that could propagate around the control room. As incorrect information about contacts can lead to accidents (National Transportation Safety Board, 2001; Marine Accident Investigation Branch, 2016), it is vital that any redesign seeks to support operators in maintaining information about contacts and situational awareness by reducing the need to maintain their own record of data. The second issue is that there is that the recorded information requires cognitive processing to translate the written format into an understanding of the real-world representation to *'Interpret Contact Actions'* and *'Predict Contact Actions'*. Therefore, in addition to supporting operators in maintaining their information, the new interface could order the data such that it requires cognitive processing. This is congruent with the notion of providing functional information in an ecological interface design, adding the physical information typically displayed (Dinadis and Vicente, 1996).

5.4 Worker Competencies Analysis

WDA has furnished an understanding of how Sonar could be improved in a redesign by examining the waterfall as a technical leverage point, although it did not reveal the social aspects of accommodating operators of different skill levels. As a sociotechnical system, formed by the operator and console, it is vital to ensure both for a complete understanding of Sonar operation. A WCA, see Table 16, shows how operators of different skill levels interact with Sonar. From this, it is possible to understand what functionality is useful to different skill levels of operators, and to inform design requirements (McIlroy and Stanton, 2012).

The WCA shows that there is support for all operators in the interface. However, the interface does not differentiate between the operators, with limited configurability available. For example, expert operators rely on experience to recognise patterns and extract information (e.g., classification, speed) from this, whereas novice operators would use tools to extract the same information. The options available in Figure 31b-e appear to exist to assist SBB for operators when '*Detecting Contacts*'. Novice operators exhibiting KBB may only rely on strong visual and aural signals and thus may not employ these settings designed to detect subtle traces. Similarly, expert operators may not have any use for the classification information in Figure 32e when '*Classifying Contacts*', although novice operators may find this information useful. This suggests while capability is provided to support different skill levels of operators, a redesign could offer more configurability to support this. This is reinforced when considering that information is currently spread out across multiple screens, which may disproportionately affect novice operators, who require access to more supporting tools across these screens.

Table 16 – WCA for a Sonar Interface

	Skill-based Behaviour	**Rule-based Behaviour**	**Knowledge-based Behaviour**
	Observed behaviour of experts	If-Then rules to identify if a task should be done	Prerequisite knowledge / capability for novices to complete tasks
Detection of Contacts	Experience in identifying subtle signals that indicate a faint, or evasive, contact Changing settings to maximise detection capability	Establish if detection occurs on any other arrays when detected on Sonar Find Sonar counterpart to detections from other sensors if not already detected	Listen for aural noise from Sonar arrays Visually identify traces on the waterfall
Transfer Information About Contacts	Assess command team requirements, providing information proactively	Inform command immediately of changes to contact information	Assign a tracker Send a manual cut Verbally communicate details
Classify Contacts	Classify unknown contacts using behavioural/signal traits	Consult reference material for known vessels if classification is not immediately known	Compare narrowband to database signatures Identify sound patterns
Gain Speed Estimate on Contacts	Calculate speed directly from a signal Use traits of speed to check speed, such as high revs for high-speed	Use priors for default speed based on contact type if not known	Analyse DEMON signal Identify sound patterns
Predict Contact Actions	Use experience to predict contact actions based on their type, disposition, and previous actions	Use intelligence provided to identify actions that a contact may take (e.g.: follow a shipping route)	Make a forecast based on trace history Make forecast based on contact known information Make forecast based on contact known classification
Interpret Contact Actions	Use experience to interpret contact actions based on their type, disposition, and prior behaviour	Use intelligence to identify causal factors, such as shipping lanes, weather conditions and/or other vessels	Assess trace history Make judgement based on contact known information

	Make judgement based on contact known classification

5.5 Insights Gained

5.5.1 Cognitive Workload and Working Memory Are Key

The key areas for generating design directions for a new sonar interface have been reducing an operator's cognitive workload and reducing demands on their working memory. While separate concepts, their joint induction is linked to current design choices, namely information and tools being displayed across screens (Cravens, 2021). This requires operators to switch screens to gather all information on a contact, increasing cognitive workload, and to store information in working memory between screens. The effect of screen switching was noted by Michailovs *et al.* (2022), although they provided an alternative viewpoint in that the separation could create a natural "attentional spotlight" and reduce cognitive burden, referring to the work of Posner (1980). This work proposed that people can only monitor one spatial region at a time (Tong, 2004). However, Tong (2004) presented evidence that the spotlight could be divided in certain circumstances. This divided attention could be used to focus on multiple pieces of information in an integrated display, with the benefits of reduced cognitive workload from the "attentional spotlight".

A move to a representation with greater ecological validity as part of the ecological interface design process may also address the issue of multiple screens. Firstly, it could reduce cognitive workload by facilitating direct perceptual processing without intermediary steps where possible (Gibson, 1979; Rasmussen and Vicente, 1989; McIlroy, 2016). Secondly, by arranging the interface according to the functional information identified from the abstraction hierarchy, information could be sufficiently collocated to derive an appropriate understanding of the current state and structure of information in the operating environment (Vicente and Rasmussen, 1992; Bennett and Flach, 2019). This could be extended to providing sonar operators with information from other roles that they currently do not have. This was shown to be beneficial by Michailovs *et al.* (2021), who found that teams constructed a more accurate tactical picture when they had access to pertinent shared information.

Colocation of information based on identified functional requirements may also reduce the constraining effect of limited working memory, as operators will not be required to transition information between screens to gather information on a single contact. Conversely, this approach might increase the amount of working memory required to construct information formed from data from multiple contacts, such as refreshing all speeds. Another consideration for working memory is factoring in the communication of the command team, which would utilise the phonological loop. This would require an operator to split their working memory between processing current sonar data, and handling communications. A reduction in capability for either is not optimal. Therefore, it would be pertinent to consider furnishing operators with communication using a different modality, such as textually (visual), with the aim of freeing capacity. This could increase capacity as is has been suggested that there are separate limits for different modalities, with increased demand occurring when modalities make joint demands on a common representational format (Wang *et al.*, 2003; Fougnie *et al.*, 2015; Tamber-Rosenau and Marois, 2015). Increased capacity could also help to address the issue of production blocking that Roberts *et al.* (2019) identified, where synchronous information communication (i.e., one person may speak at once) can degrade verbal information (Roberts and Cole, 2018). As only one person can speak at a time, individuals may not be able to communicate when required, and may subsequently forget to communicate, or consciously choose not to as their information seems less original or relevant (Stanton *et al.*, 2003), citing from (Diehl and Stroebe, 1987). Roberts and Cole (2018) also discussed that there is evidence demonstrating that workload affects verbal communication capacity. Adding an alternative mode of communication such as textual chat might also help to address these issues, providing operators with more outlets to communicate as quickly as possible, and using a suitable medium to do so. For example, sonar operators who are listening for a quiet contact might prefer to share potential bearings via text chat to keep the audio modality free, and to avoid blocking other operators' communications (multiple simultaneous textual messages would get stored in the conversation history, instead of being hard to distinguish overlapping voice communications).

5.5.2 Data Should be Displayed in a more Representative Manner

WDA showed that Sonar exists to enable tactical picture generation, in addition to supporting higher command activities. The successes of these purposes were identified by the fulfilment of local and submarine goals. Local goals appear to have driven Sonar's design, splitting functionality along the DCLT process, using screens to perform different operations. The abstraction hierarchy revealed the waterfall as a core concept across these screens. Implementations were polymorphic, with comprehension of one waterfall being translatable to others. It is hypothesised that the waterfall continues to be a pervasive Sonar interaction mechanism due to its ability to display time series data in a readily interpretable and data-dense fashion, minimising the chance that the amount of data presented would exceed the operator's ability to interpret it (Woods, Patterson and Roth, 2002) or store it in their working memory (Mason *et al.*, 1989), when compared to other representations of the data, such a table of numbers representing the signal parameters at each bearing. Both current and historical data can be presented simultaneously, facilitating an operator's goal of understanding contact actions, maintaining their situational awareness. Traces further enhance presented information, allowing operators to understand signal intensity and patterns. However, despite the effective display of information, a waterfall does require cognitive processing by the operator, which may negatively impact their performance (Hanisch, Kramer and Hulin, 1991; Masakowski and Hardinge, 2000). This cognitive capacity could potentially be better utilised on other tasks. For example, the bearing tape represents circular data but is a straight line. If bearing data was represented as a polygon, operators could potentially understand a signal's bearing more intuitively. Alternate designs should be sought for physical objects that require additional processing, ensuring operators have as much cognitive capacity as possible for more complex tasks, which could be safety critical. The waterfall presents significant amounts of information in a compact area, utilising co-located components to enhance presented information. Should a redesign occur, it would be pertinent to place a focus on ensuring that information is represented sufficiently in the new design and that functionality gained from co-located components is not lost. For example, the aural and trace representations of a contact provide operators with a lot of information separately, and their colocation tightly integrates this information readily. Any redesign should seek to keep this symbiotic relationship or ameliorate functionality deficits that may result from their separation.

5.5.3 Configurability and Support for Different Skill-Levels is important

The WCA revealed that extensive training allows operators to use Sonar at a skill-based level, with experience facilitating rule-based and knowledge-based completion of tasks. However, the Sonar interface is the same for all operators, regardless of experience. This poses two potential problems: overwhelming junior operators and inhibiting experienced operators.

When performing tasks from Table 16, working memory will be utilised by operators to process what they perceive, commit experience to long-term memory, and inform future actions. For novice operators utilising KBB, perceiving task-irrelevant, but still present, interface aspects may saturate their working memory (Mason *et al.*, 1989). Consequently, relevant information may be overlooked, or an improper outcome may occur from valid information, potentially adversely affecting ownship safety. Additionally, these operators may expend cognitive effort on placing only relevant information into working memory, delaying decisions, and potentially becoming overwhelmed. In both circumstances, this could affect their situational awareness, which may affect the command team's distributed situational awareness. For example, if an operator utilising KBB was using the broadband screen (Figure 31), they may only use the top waterfall from area "a" and contact management in area "d". However, as Sonar cannot be configured to support this, the extra waterfall and available settings may reduce clarity. Potential options to mitigate this would be an abridged interface, or the ability to hide irrelevant components temporarily. As operators gain experience, they will be able to work at the Rules as well as Knowledge levels of the SRK Taxonomy for certain complex tasks, enabling the performance of these tasks in a comparatively autonomous manner. The extra cognitive capacity created can help operators to perform more and more complex tasks. However, this may require more switching between different levels of cognitive control as they may have differing skill levels for tasks. Thus, transitions between the levels should actively be supported by the interface (Vicente, 1999b; Kilgore and St-Cyr, 2006) to avoid hindering operators. Operators should be able to perform tasks at any level of the SRK Taxonomy, without the Sonar interface requiring changes that could disrupt their cognition. This could be achieved using a combination of simple features, such as using a dropdown menu to select waterfall features or advanced features that change the system to suit the operator. The operators' skill level should be accounted for, exploiting their innate cognitive capabilities. These can be simple, such as knowing the colour red indicates danger, or complex, such as finding an object of interest in a perceptually rich environment. By capitalising on abilities that do not require conscious processing, spare working memory capacity can be retained, potentially enabling operators to maintain better SA.

Expert users will work with Sonar using SBB. If the interface does not support this fully, then conscious cognitive effort will be required. While expert users are still likely to generate correct outcomes despite being inhibited by the interface, task timeframes may be extended. For example, consider that from the broadband screen an expert user cannot see a contacts classification, DEMON traces, or speed. These factors may be used by operators to predict contact behaviour. If all variables were immediately accessible, these predictions could be made readily; however, as they must be found in the interface, time and working memory penalties are introduced. These penalties could lead operators to take too long to identify a hostile contact or miscalculate a contact's speed, leading to adverse ownship safety consequences that could be avoided by fully supporting the operator's cognitive requirements.

5.5.4 Summary of Insights

A summary of design insights proposed throughout this chapter is presented in Table 17 for convenience.

Table 17 – A summary of insights gained from conducting WDA and WCA on Sonar

Section	Insight
5.3.2 5.5.3	The interface should be configurable and support effective data management and provide tools to process data into information for operators of all skill levels.
5.3.2	Use information representations that reduce cognitive workload.
5.3.3 5.3.4.5 5.4	Orient the interface around contacts and their information, as opposed to specific information gathering processes (DCLT) and provide configurability.
5.3.4.1 5.5.2	Adopt a map-based display to show bearings and other related information with greater ecological validity.
5.3.4.2	Represent time effectively on the map-based display.
5.3.4.3	Maintain affordances provided by traces.
5.3.4.3	Maintain utility of traces on waterfall for interpreting contact behaviour.
5.3.4.4	Maintain bimodal processing of sonar data.
5.3.5.4	Move as much information management as possible into the interface so that separate sources do not need to be maintained.
5.5.1	Provide operators with information from other stations if useful.
5.5.1	Textual chat might be useful for operators.

5.6 Conclusions

This chapter has explored the waterfall as a leverage point for deriving insights for redesigning Sonar. The concept of a leverage point is to make changes to a system aspect that can affect change throughout the system (Read *et al.*, 2018). The waterfall was focused on as it is a common representation of sonar data (Chen and Burns, 2007; Ericsson, 2009; de Moura, de Seixas and Ramos, 2011).

It presented the method used for completing the WDA and WCA for doing so, mostly in line with that of Stanton *et al.* (2017a), Naikar, Hopcroft and Moylan (2005), and Burns and Hajdukiewicz (2004). However, a different approach was proposed for WDA. The notion of making contextually driven changes appears throughout the literature, including in seminal work (Burns and Hajdukiewicz, 2004; Jenkins *et al.*, 2009; Stanton *et al.*, 2017a). The difference was to scope the analysis to the existing interface, populating the Physical Objects level with interface components. This approach was taken as it was realised that the interface was already utilising representations of Physical Objects within the work domain. Thus, it was posited that it was possible to use these items as proxies for the actual Physical Objects in the domain. This built on the concept of methods of representing Physical Objects in an EID (Talcott, Martinez and Stansifer, 2007; Bennett and Flach, 2011). If these perceptual forms had already been included in the interface implicitly, then it could be possible to utilise them. It is proposed that this approach has merit to align the process with conceptual work, as presented in this thesis, by appropriately limiting the analysis size and scope with the planned experiment in mind, as advocated for by Kortschot *et al.* (2017). Additionally, the approach was designed to capitalise on available subject matter expert resources, requiring fewer experts to provide detail at the scope required. Finally, the change was made to be congruent with the cyclic nature of WDA, cognisant of the requirement to iterate to provide more detail where required. In this instance, this would involve mainly changing the bottom two levels to account for using the actual Physical Objects within the work domain. This means that the abstraction hierarchies can be taken forward for more traditional applications.

The resultant WDA and WCA were presented, along with the design insights that they yielded, which are summarised in Table 17. For the WDA, the waterfall was broken down into its component Physical Objects to understand how it operates and where changes could be made. This was done as the operation of Sonar is not a readily understandable concept, such as driving. Therefore, a degree of normative information was required to contextualise the formative recommendations. The insights gained focused on the reduction of cognitive workload and working memory utilisation, both of which arose from the way that information is structured and processed in the current interface. Overall, the first difficulty was how aural data is represented. The waterfall is a common representation of sonar data for a good reason, owing to the amount of data it can encode, although it takes a different form than the environment it represents, which would degrade situational awareness and require additional cognitive effort to process. Providing a representation with greater ecological validity might address these challenges. It is also apparent that Sonar remains the same for operators of all skill levels, potentially degrading their efficacy due to working memory limitations or lack of support for an operator's skill level. It was suggested that the redesigned interface could address this by providing configurability and tools appropriate to operators at each skill level. The next chapter will focus on detailing recommendations for TMA, utilising the method proposed in this chapter.

Chapter 6 **Cognitive Work Analysis of Target Motion Analysis: What could be improved about solution entry?**

6.1 Introduction

Chapter 5 presented the results of generating design directions from Work Domain Analysis (WDA) and Worker Competencies Analysis (WCA) for Sonar. This chapter follows the same approach (see Section 5.2 for the method) to generate design directions for Target Motion Analysis (TMA), and as such does not contain a methods section. The leverage point identified for TMA was the Local Operations Plot (LOP), which is a top-down map view of a submarine's operating environment that contact solutions are plotted on. It was identified as a predominant method of entering solutions and evaluating their accuracy (Clarke, 1999). As with Chapter 5, this chapter aims to build on the understanding of submarine control room operation from Chapter 4 using Cognitive Work Analysis (CWA) to understand where improvements might be made when creating the redesigned TMA interface.

Like Sonar, TMA is working with larger quantities of more complex data. This data is cumulatively generated from sensors getting more advanced and providing more data, and also by newer combat systems, such as Submarine Combat System – Next Generation (SMCS-NG; BAE Systems, 2015) or Aegis Weapons System (AWS; Threston, 2009), becoming more capable. These advances present an opportunity to introduce more automation. Automation can be defined as applying technology to reduce the need for, or replace, human workers in a process or system (Boy, 2014; IBM, 2021b). It could reduce human error by ensuring cognitive requirement is the minimum possible (Breton and Bossé, 2003). However, this introduces a newer challenge of ensuring that operators are still capable of stepping in for the automation in case they have not been reviewing their knowledge of practising skills (Bainbridge, 1983). Therefore, it would be pertinent to frame design directions within the automation capabilities that newer systems could afford, while taking care not to degrade operator capability as a result. Additionally, as will become clear, the current method of solution entry can be cumbersome, requiring extra effort from the operator to work with solutions. Consequently, recommendations in this chapter also focus on removing cumbersome interaction as a pain point in the current design. Read *et al.* (2018) define a pain point as problems or issues that represent user frustration, conflicting goals, or information bottlenecks.

6.2 Work Domain Analysis

The abstraction hierarchy for TMA is shown in Figure 34. There is not an initial version as with Sonar, because Sonar was the first abstraction hierarchy to be completed and was iterated based upon the proposed updated method in Section 5.2.1, whereas this was created directly the approach.

Figure 34 – Full Abstraction Hierarchy of TMA

6.2.1 Functional Purpose

The TMA interface exists to *'facilitate tactical picture generation'* and *'support higher command activities'*. A tactical picture is a map of a submarine's current environment, with all known details plotted onto it, providing support for command decisions. TMA takes time series data collected for detected vessels from other stations, such as Sonar and Periscope, to create and calculate a likely location and movement profile. The location is plotted directly onto the tactical picture, directly informing decisions by the Officer of the Watch (OOW). As these decisions affected the submarine's mission and safety, TMA is an integral part of a submarine control room. These goals are common to all aspects of the command team, although there is a greater emphasis on facilitating the tactical picture, as TMA combines the provided data to synthesise solutions. While other stations do contribute to the tactical picture, a requirement to directly input information onto it means that TMA is a high priority goal and tactical picture compilation should be supported and/or enhanced by as much functionality as possible.

6.2.2 Value & Priority Measures

As with Sonar, there is a mix of global goals and local goals present. All three global submarine goals (Mack, 2003), '*Maximise Safety*', '*Minimise Counter-Detection*', and '*Maximise Mission Objective Completeness*', are present. To '*Maximise Safety*' operators plot the location of vessels onto the tactical picture, allowing the command team to navigate safely around them, preventing collisions. Plotting vessel locations allows the submarine to *'Minimise Counter-Detection'*, either by maintaining distance or facing the quietest part of the boat towards the specific contacts. By contributing information to strategic decisions made by the OOW, TMA can '*Maximise Mission Objective Completeness*'.

Measures of success specific to TMA are '*Maximise Knowledge of Contacts*', '*Minimise Threats from Threat Priority*', and '*Maximise Solution Integrity*'. The measure of success '*Maximising Knowledge of Contacts*' has two meanings. Firstly, as much information as possible about a contact's location and behaviour must always be available. Second, to achieve this, operators must have all the information known by other members of the control room, such as Sonar. This collection of knowledge is closely related to the measure of success *'Maximising Solution Integrity,'* as operators should ensure that all information provided about a vessel is as accurate and complete as possible. By maintaining a high level of information integrity, the OOW can '*Minimise Threats from Threat Priority*' by effecting suitable strategies for the current tactical picture.

All measures of success for TMA are synergistic, aimed at maintaining the most accurate tactical picture possible to ensure that ownship remains safe as a top priority. With more data being generated from control room operations across more screens (Dominguez *et al.*, 2006; Stillion and Clark, 2015), the 'how' of meeting these goals should focus on the proper integration of data to reduce screens, as well as provision of appropriate information and tools. This could also support the OOW further by making information more accessible for them to make decisions, as knowledge integration can be challenging without the correct support (Dominguez *et al.*, 2006).

6.2.3 Purpose-Related Functions

The Purpose-Related Functions were different stages of processing contact information, including '*Choose Contact*', '*Create Solution*', and '*Refine Solutions*'. Each function has connections to multiple measures of success, indicating that all functionality contributes generally to achieving station measures of success. Functionality is not subset into groups for global and local measures of success, reinforcing this indication. Having multiple links between each Purpose-Related Functions and Value & Priority Measures indicates the system is designed to provide a toolkit of functionality that can be used at the operator's discretion to achieve their goals in a contextually appropriate manner. For example, if the OOW asked TMA to '*Maximise Knowledge of Contacts*', but they did not have enough information to '*Refine Solutions*', they could still achieve this by '*Creating Solutions*' for contacts they had not yet processed.

This manifests as having only one screen for TMA, which should reduce the increased cognitive workload and working memory requirements (Cravens, 2021; Michailovs *et al.*, 2022) associated with switching between screens. However, these are likely to be replaced by the need to manage the data and information for contacts. While the information is displayed on screen, the operator must still take information from other operators and enter it onto their display. An example of this is when Sonar operators verbally communicate a calculated speed, requiring them to remember this speed until they can navigate to contact and enter the information into the solution. As with Sonar operators, they can record information externally, but this has the implications discussed previously (see Section 5.3.5.4). This is further exacerbated for TMA operators, as they handle data from other stations, meaning that they must communicate with others to update their data. This can be subject to information bottlenecks (Roberts, Stanton and Fay, 2017b), production blocking (Stanton *et al.*, 2003; Stanton and Roberts, 2020), and mistakes. Given that the information concerned is easily passed digitally, a design suggestion would be to implement mechanisms for sharing information between operators in this manner where possible. Another challenge for operators is that they must interpret contact solutions to determine a future state, a key aspect of situational awareness (Stanton *et al.*, 2017b). This is in part provided by the functionality to evaluate a current solution, which contains mostly physical information about its current parameters visually and using number readouts. This could be expanded to include functional information, such as making the future state explicit (Chalmers, Easter and Potter, 2000).

6.2.4 Physical Objects and Object-Related Processes: How does solution entry work, and how could it be improved?

The LOP is the main component of the TMA screen, and most stages for processing a contact can be performed using it. As a reminder from Section 4.1.1.5, a LOP displays time series data of a contact's detections on a geographic plot, see Figure 35. The background represents a given map area. Data is represented as cuts ('a'), which are lines drawn along the bearing from which a contact was detected from ownship ('b'). Speedstrips ('c'), also known as rulers, represent a contact's course as straight lines, with marks on to represent when cuts occurred. Operators translate, rotate, and scale a speedstrip until the marks match the cut lines for a contact. Speedstrips can be directly manipulated visually or known information can be entered ('d', which is not shown as not part of the LOP, but labelled for clarity; the entire screen can be seen in Figure 46). If a speedstrip matches the cut lines and is accurate in the accuracy display ('e') this path can be shared, which plots it as an icon on the global geographic tactical picture. Speedstrips automatically dead-reckon (extrapolate a new position from historical trajectory), allowing the tracking of a contact's movement.

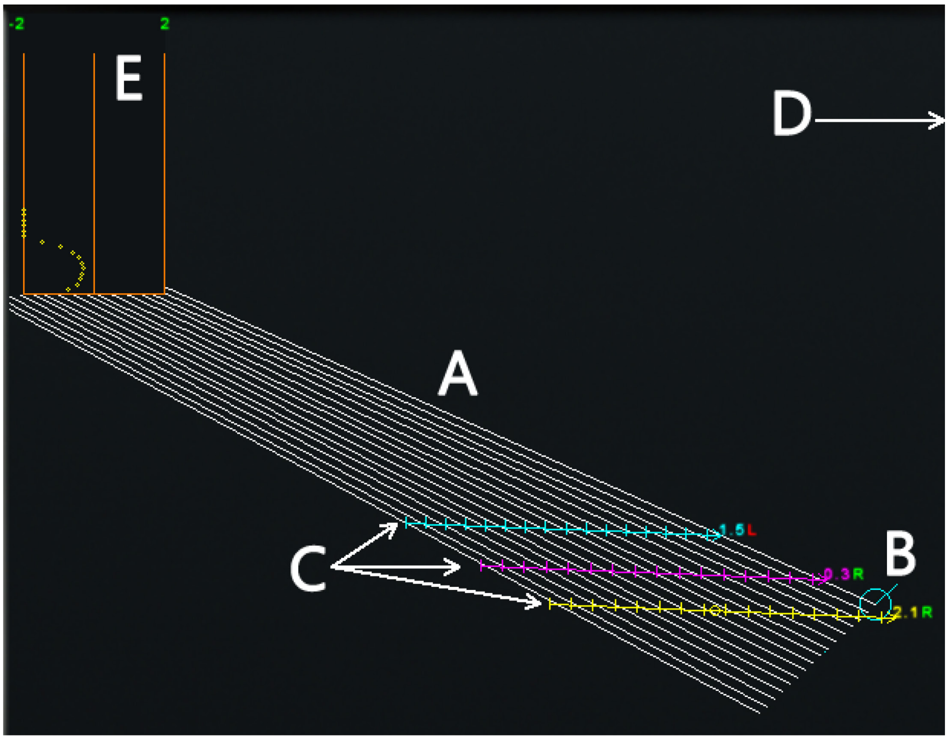


Figure 35 – A screenshot of a Local Operations Plot, demonstrating speedstrips being overlaid onto historical cuts for a contact

The LOP does not display all contacts at once and displays a maximum of two contacts at a time, one that is being worked on, and another for comparison of data as a merge candidate. The operator is visually shown the cuts and can decide, although they are not provided with functional information to verify that the merge is fundamentally viable, such as the contacts having similar bearings or being from different sensors. This could be improved in a redesigned interface by alerting the operator if a merge appears to be incorrect. The contact shown can be changed using drop-down menus, which list all available contacts to work on. However, operators are not alerted by the interface if a new contact is cut to them, or a contact otherwise needs attention. Thus, another design suggestion would be to always make the available contacts visible to operators, and to use the contact displays to show functional information about the contact (e.g., has not been seen in __ minutes, and therefore needs to be investigated to fulfil higher-level goals of *'maximise solution integrity'* and *'maximise safety of operations'*). The state of the LOP is saved when changing contacts and restored when returning to a contact. However, there is no overview that allows operators to see all contacts at once, which could provide them with more functional information to understand their environment, such as limiting solution possibilities for a contact based on where other contacts are. There is a tactical picture screen that could be used, although this would require the operator to split their attention between the two screens, whereas it could be better to provide the information appropriately integrated into the LOP (Hall, 2012; Michailovs *et al.*, 2021).

To aid understanding of where improvements could be made and why, components of the LOP were isolated from the main abstraction hierarchy to understand what functionality they afforded and how they did so. This was done to make the nodes and links more apparent when compared to the entire abstraction hierarchy. The means-ends links in the abstraction hierarchy could have revealed this information, however, they have been presented separately to make them more accessible and understandable. The Physical Objects being examined are the map space, ownship icon, cuts, and speedstrips. For each Physical Object, an abstraction hierarchy specific to the object is presented, along with what could be achieved using it and why. The definitions for all nodes are provided in Table 18 to avoid repetition. After isolating the relevant nodes for each of the four main components of the LOP, a holistic overview of the TMA screen operation will be provided, alongside further ideas for improvement in a redesign.

Table 18 – Definitions of nodes at each level (Functional Purposes → Physical Objects) contained in the isolated LOP abstraction hierarchies presented in the sections below.

Level	Node	Definition
Functional Purpose	Facilitate Tactical Picture Generation	Support the creation of a picture that depicts what is currently around the submarine
	Supporting Higher Command Activities	Ensure that senior posts can carry out their duty
Value & Priority Measures	Maximise Known Contact Information	Provide as much information as possible about contacts
	Maximise Safety of Operations	Ensure that the submarine is always safe
	Minimise Counter-Detection	Reduce the ways the submarine can be detected
	Maximise Mission Objective Completeness	Contribute towards completing the assigned mission
	Minimise Threats from Threat Priority	Reduce threats to ownship, whether accidental or deliberate
	Maximise Solution Integrity	Ensure that solutions are valid, accurate, and reliable
Purpose-Related Functions	Create Solution	Provide time and space positional information for a contact
	Assess Solutions	Determine a solution’s suitability, validity, accuracy, or reliability
	Ensure Solution Accuracy	Specifically ensure a solution’s accuracy
	Merge Solutions	Combine two separate solutions into one solution
	Refine Solutions	Make changes to a solution
	Remove Solutions	Delete a solution
	Split Solutions	Split a combined solution into its component parts
	View Solutions	Look at available solutions
Object-Related Processes	Choose Contact	Select a contact to work with
	Gauge Solution Accuracy	Determine how accurate a solution is
	Representation of Contact Information	Represents information pertaining to a contact
	Compare Possible Solutions	Comparison of differing solutions for a contact
	Predict Contact Actions	Predicts the actions that a contact may take in the future
	Analyse Solution History	Assessment of data in a solution
	Create Possible Solution for Contact	Synthesise a new solution for a contact
	Determine Course Direction	Ascertain the course of a solution
	Assess Contact Information	Analyse the available information for a contact
	Change Solution Drastically	Make considerable changes to a solution
	Manipulate Solution	Make changes to a solution
	Locate Solution on Display	Find the location of a solution
	Determine Bearing Rate	Ascertain the bearing rate of a solution
	Display Parameters of Solution Speedstrips	Shows the component parts of the solution
	Analyse Previously Entered Solutions	Assessment of prior contact solutions
	Choose Solution Speedstrip	Select a speedstrip to work with
	Commit Solution	Saves a solution, either sharing it or keeping it locally

	Representation of Current Environment	Displays the submarine's current environment
	Change Accuracy Required	Vary how accurate a solution is
Physical Objects	Line	The central line of a speedstrip that other components are attached to
	Distance Marks	Small perpendicular marks on the main speedstrip that denote where cuts would have occurred on the solution
	Directional Arrowhead	An arrow at the leading end of the speedstrip showing direction, and allowing manipulation
	Movement Circle	A circle in the middle of a speedstrip, allowing manipulation of its location
	Bearing Rate Indicator(s)	Displays the bearing rate of a given solution speedstrip
	Solution Display - Working	Numerical parameters of the solution that is being worked on by the operator
	Solution Display - Shared	Numerical parameters of the solution that has been shared to the tactical picture
	Solution Display - Private	Numerical parameters of the solution that has been saved locally
	Solution Accuracy Display	A component for determining how accurate a solution is

6.2.4.1 **Map Space**

The map space is the black background to the LOP, which represents the environment in which the submarine is currently operating. For security reasons, the actual map is not usually displayed and is replaced by a blank area. This is so that operators do not know where the submarine is at any given time, although adding some form of a map might provide additional information useful to solution construction. Operators can pan and zoom the area using either the mouse, keyboard, or buttons below the LOP. While no geographic information is available, the blank space is representative of a two-dimensional top-down map view, so the pan and zoom capability enables operators to select an appropriate working area. Figure 36 shows an example map space, with associated AH nodes. The '*Map Space Display*' is the background for '*Representation of Current Environment*', allowing operators to '*View Solutions*'. This allows them to '*Maximise Known Contact Information*', which can be used to '*Support Higher Command Activities*'.

The LOP being based around a map view suggests a level of synergy that could be exploited by a redesigned interface. If both Sonar and TMA information could be represented on the same display, the integration could offer better information assimilation (Dominguez *et al.*, 2006; Hunter, Hazen and Randall, 2014). While Sonar information is currently displayed using the waterfall display, the design recommendation to shift to a skeuomorphic representation of bearings makes this a possibility if both interfaces are designed simultaneously, or even as one interface that combines both sets of functionality.

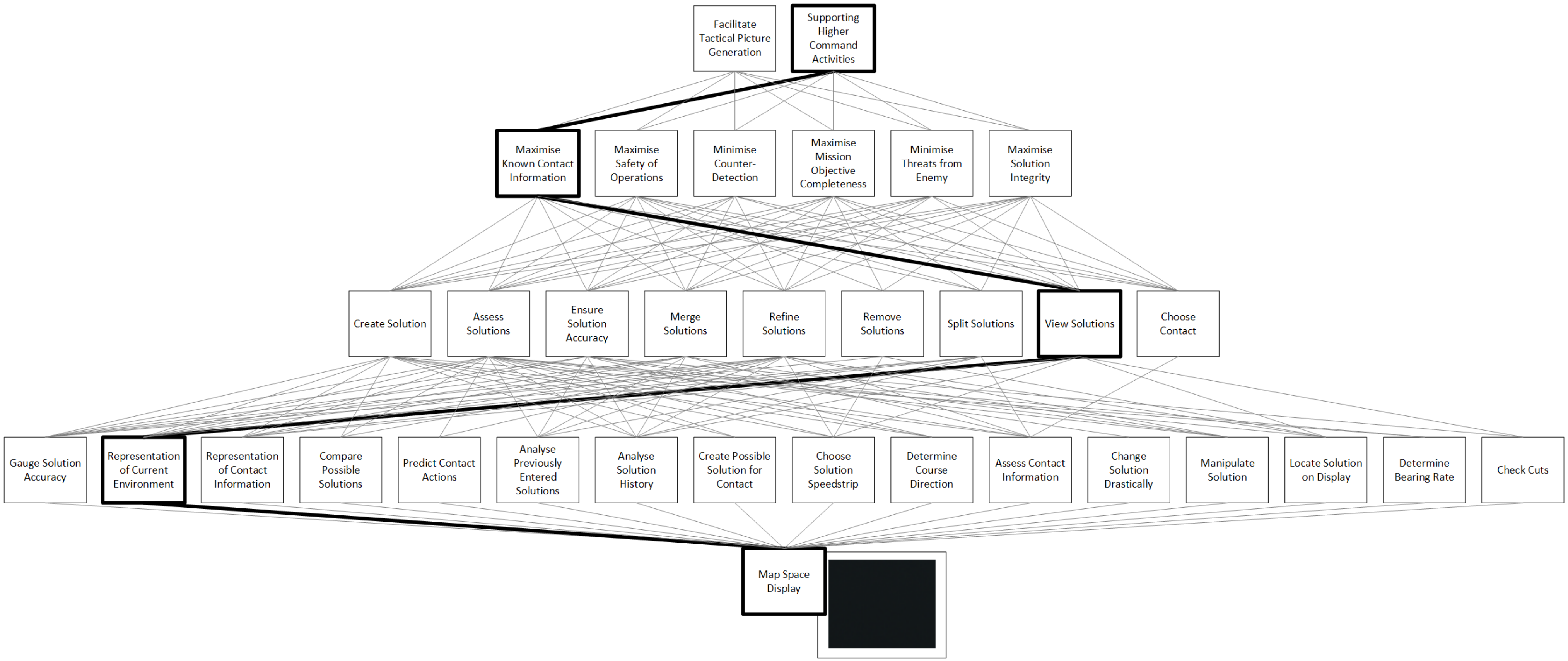


Figure 36 – Map Space Abstraction Hierarchy. Note that the map space is deliberately blank for operators.

6.2.4.2 **Ownship Icon**

The geographic position of ownship is represented as an icon on the map view. In addition to showing location, it also shows ownship's location, course, and speed. Course is represented by an intersecting line with an angle of the current course, and the length of the line is proportional to the submarine's speed. The icon's position is updated automatically, so operators are not required to manually track or otherwise update the icon's location. Figure 37 shows an example ownship icon, with associated nodes. The *'Own-Ship Icon'* helps to place the submarine spatially in a '*Representation of the Current Environment*'. This can help during '*Solution Creation*' as it provides a point of reference. By enabling accurate solutions to be entered the submarine can *'Minimise Counter-Detection'* by maintaining separation from contacts that could counter-detect, therefore completing the purpose of '*Remain Undetected*' to '*Support Higher Command Activities*'.

While the current interface provides a representation of the current environment, it does not provide functional information associated with ownship that could be beneficial for operators to complete their work. For example, sensor ranges are not displayed from ownship, meaning that operators cannot directly perceive the sensor geometry, which could help narrow down solution possibilities; contacts cannot be detected outside of a sensor's capability, so solutions should not be placed outside of this range. This information should be represented to TMA operators in a redesigned interface, allowing them to combine knowledge from the position of ownship's icon and the sensor ranges.

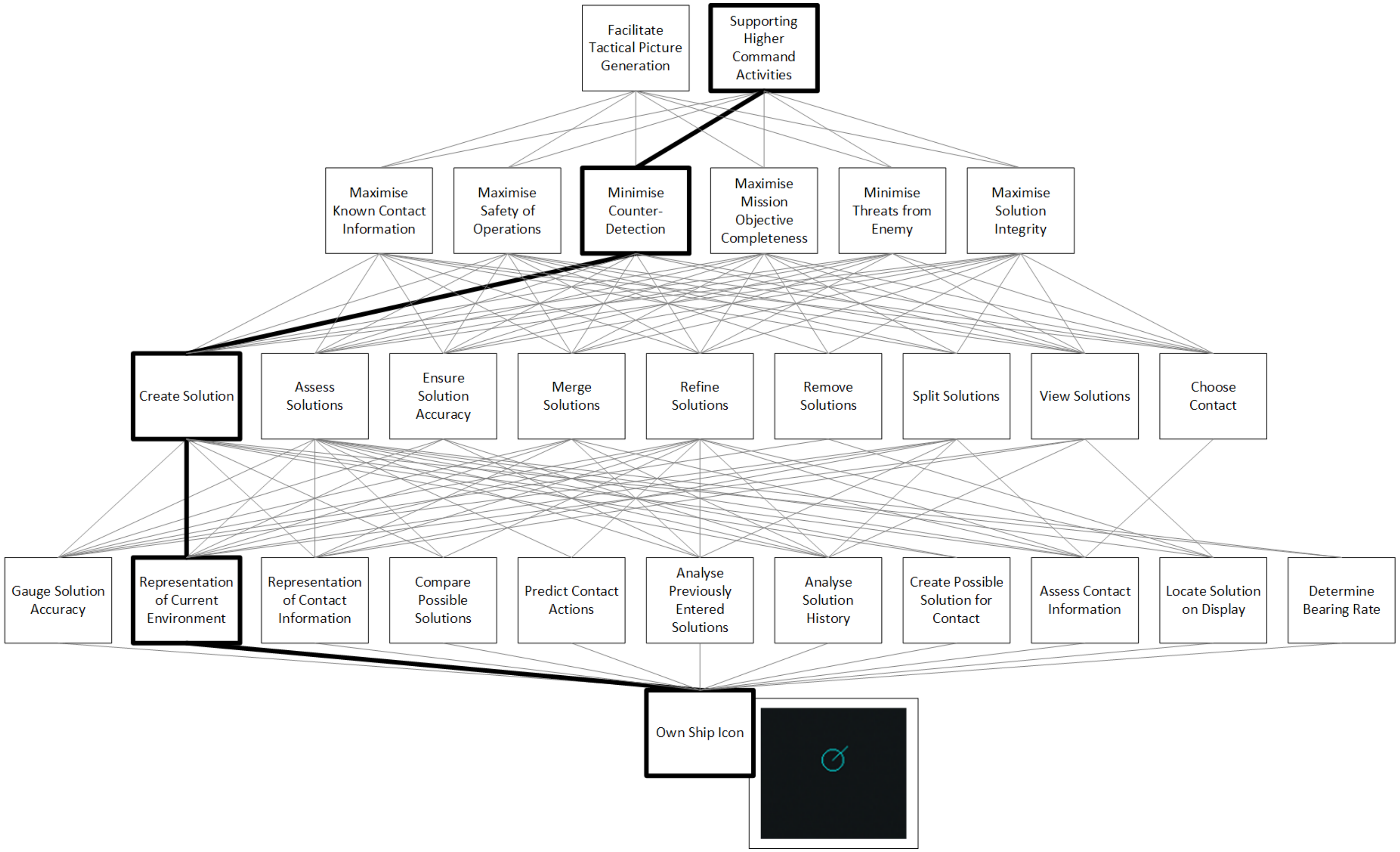


Figure 37 – Own-ship Icon Abstraction Hierarchy

6.2.4.3 **Cuts**

Cuts are lines drawn to represent a detection on sensors. They are drawn as straight lines originating from ownship's location to the maximum range of the detecting sensor, with an angle equivalent to the detection's bearing. A different colour is used to represent cuts from each sensor. A new cut is provided either when a tracker is assigned, when information is manually sent, or when a tracker automatically sends data at a specified interval. Cuts automatically being sent from trackers is useful as it removes the need for operators to do so, although it does mean that information can lag for TMA operators if an action occurs a significant amount of time before the next cut (i.e., immediately after a prior cut), which might be important if a contact is manoeuvring. A redesigned TMA interface could address this by providing a mechanism for TMA operators to receive continuous information about contacts where possible. As with the Sonar broadband audio, this could benefit operators by allowing natural interpretation over time, including the use of updated visualisations (Cunningham and Thomas, 2005). A related issue is that the last cut is not differentiated for the operator, which can make it unclear what the most recent information is, which should be addressed in any redesign.

As each sensor can assign its own contacts, one physical entity may have had several contacts, at least one on each sensor, and therefore several groups of cuts. When a sensor creates a contact, it assigns a number based on how many contacts it had detected previously and prefixes it with an identifier for the sensor type. For example, if the periscope found three contacts previously, the new contact would be designated V04 (V stands for Visual). Alternatively, the first Sonar contact would be designated S01 (S stands for Sonar). These cuts can be combined by merging contacts and treating them as one, unifying the cuts. Figure 38 shows an example series of cuts, with associated abstraction hierarchy nodes. When the first '*Cut*' is sent, an operator should '*Check Cuts*' to ensure that it was received and that subsequent '*Cuts*' can be. When a solution is created the '*Cuts*' allow an operator to '*View Solutions*', '*Maximising Known Contact Information*'. This information can be used to assist with '*Tactical Picture Generation*'. Currently, the interface does not provide feedback on the merges, even if they are obviously incorrect, which could be addressed by having automation validate the merge choice to determine if it was suitable and alerting the operator if not. This could be bolstered by having the interface arrange contact information in such a way that makes merge possibilities evident.

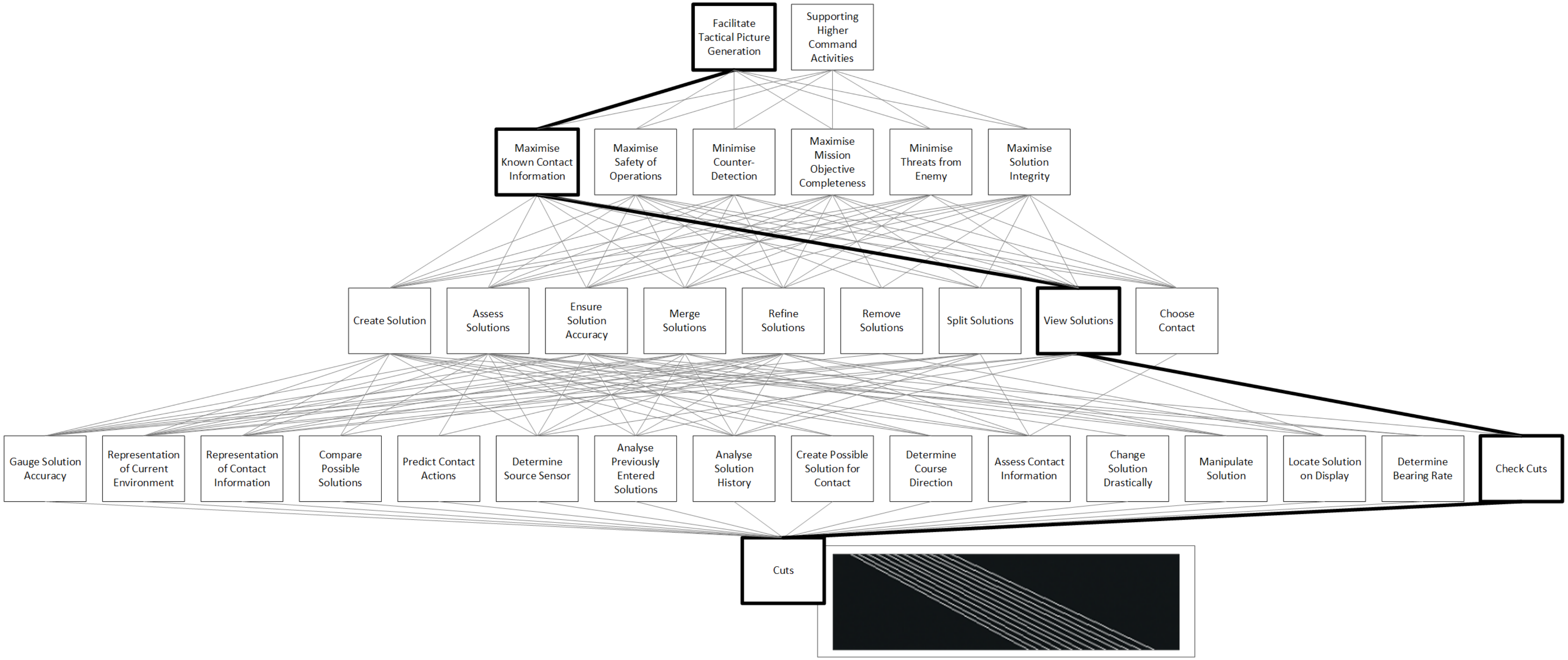


Figure 38 – Cuts Abstraction Hierarchy

6.2.4.4 **Speedstrips**

Speedstrips are visual representations of a solution and its parameters. Figure 39 shows an example set of speedstrips with an associated abstraction hierarchy. Each '*Solution Speedstrip*' can be used to '*Create a Possible Solution for a Contact*', which is part of the '*Create Solution Process*'. By ensuring that solutions have been entered for each contact, the submarine can steer away from collisions to '*Maximise the Safety of Operations*'. By assisting with this process, TMA operators are '*Supporting Higher Command Activities*'. Speedstrips are manipulated by operators working with the LOP to enter, assess, or modify a solution. There are three types of solution: working, shared, and private. The working solution can be manipulated by the operator to create an accurate solution. Once this is achieved, they can share the solution and a copy of the working speedstrip is marked onto the LOP. This copy is the shared speedstrip and cannot be directly interacted with but can be updated by updating it from the working solution again. The private speedstrip is generated when the operator saves a solution locally as a personal 'backup'. Like the shared speedstrip, it cannot be directly edited, but can be updated from the working speedstrip. The lack of direction manipulation possibilities for the speedstrips is not optimal, as changes will require shifting speedstrips around if the operator is currently working on a solution. This is cumbersome for the operator and could result in lost work if they do not save their working speedstrip to the private speedstrip before overwriting it with the shared speedstrip to edit it. Consequently, direct manipulation should be added for all speedstrips, allowing the operator to work on the separately, making the process less cumbersome.

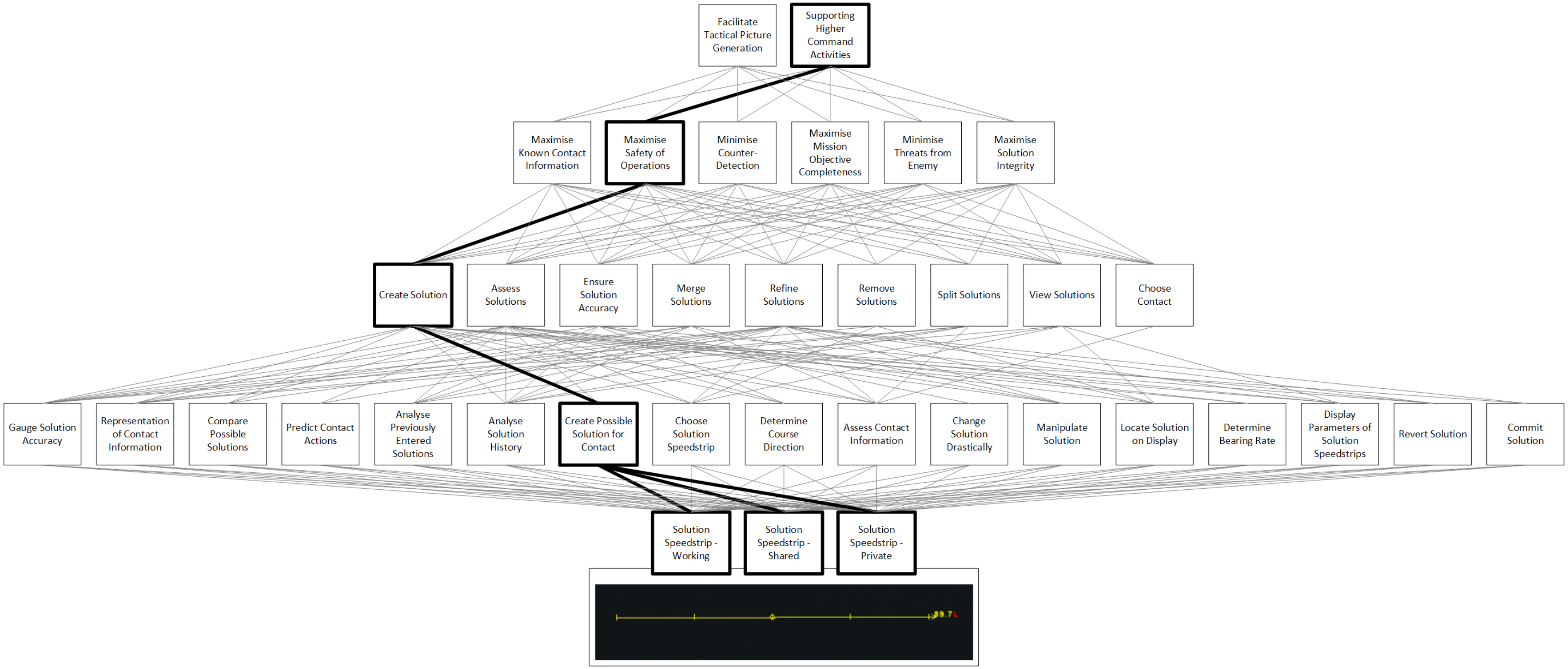


Figure 39 – Speedstrip Abstraction Hierarchy

A speedstrip as a holistic Physical Object has different properties to its component Physical Objects, which could reveal additional requirements for an updated interface. These components are the line, marks, arrowhead, movement circle, solution parameters, and solution accuracy display. They will be presented below to understand the design recommendations that they yielded.

6.2.4.4.1 Line

The speedstrip line visually represents a contact's path. Figure 40 shows an example line with associated AH nodes. The '*Solution Speedstrip Line*' helps operators to '*Assess Solutions*' by '*Analysing the Solution History*' to determine its accuracy. By understanding how accurate each solution was, operators can '*Maximise Known Contact Information*', which '*Facilitates Tactical Picture Generation*'. As the speedstrip is manoeuvred, the line visually reflects the solution's course and distance covered. The line's length is not fixed and is directly proportional to speed, which dictates its length. As the line is straight, any change in course by the contact requires the speedstrip to be positioned such that the line matches the new course. This is not representative of the work domain, as contacts can change direction and do not always travel in a straight line. This limits direct perception of a contact's history as an operator will have to recall or redetermine previous directional information for the contact, which increases cognitive workload and might be inaccurate. Therefore, it would be pertinent to introduce capability to map multiple legs of contact behaviour so that historical data can be readily viewed. Related to this is displaying future data to maximise situational awareness, which is also not currently shown. While the speedstrip dead-reckons, it does not show the operator functional information in the form of predictions, which could be improved in future interfaces.

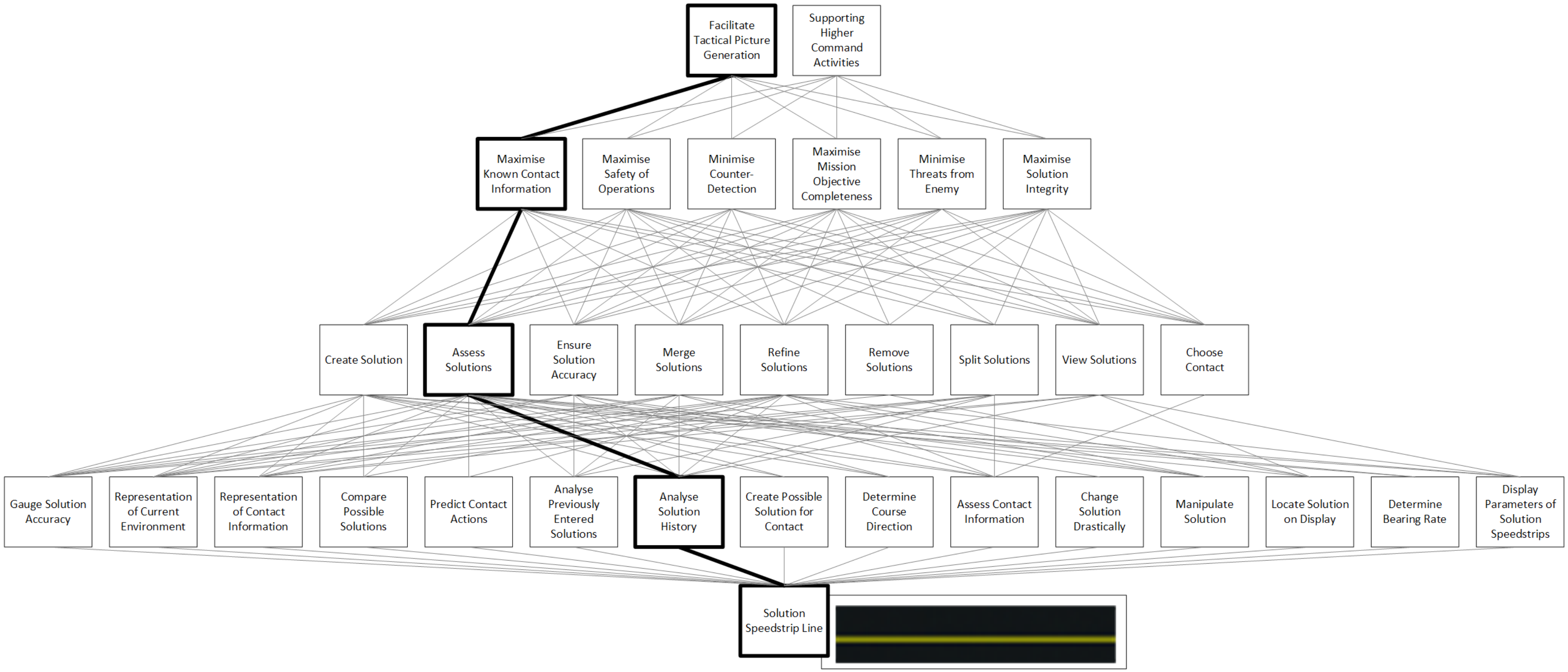


Figure 40 – Line Abstraction Hierarchy

6.2.4.4.2 Marks & Solution Accuracy Display

Marks are perpendicular lines to a speedstrip's main line that represent cuts. An analogy would be that the main line was a timeline, and each mark was an event marker. Marks are not always equidistant, as a contact could change speed, resulting in different distances being covered in the interval between automatic measurements being taken by a tracker. When creating a solution, they should be aligned with the cut that they represent. If cuts and marks are aligned it means that the solution could be accurate, as it places the events at the same place as the cuts. However, if the line were moved to accommodate a course change, cuts from the old trajectory are not expected to match as they were on a different path. The cut nearest the arrow allows the speedstrip to be manipulated by dragging its position, as does the first mark. Figure 42 shows an example mark with associated abstraction hierarchy nodes. '*Solution Speedstrip Distance Marks*' allows operators to '*Analyse Solution History*' when they are '*Refining Solutions*' to ensure they get the most accurate solution possible. This '*Maximises Solution Integrity*', and allows command to understand the environment, '*Supporting Higher Current Activities*'.

The solution accuracy display allows operators to determine how close their cut marks are to the cuts. The middle line represents the most accuracy, and the lines on either side represent a lower accuracy. Each dot represents a mark on the speedstrip. If the dots align with the central line, then that cut mark is aligned with a cut. If not, then it was far from a cut. Aligning all dots is not a necessity, as contacts could change direction, meaning that a straight speedstrip would never align. However, it is important to align the final few dots representing a steady course and speed. Alignment does not mean complete accuracy, as without narrowing it down, there are many configurations where the speedstrip could match the underlying cuts. Figure 41 shows an example solution accuracy display, with associated nodes. The '*Solution Accuracy Display*' allows an operator to '*Gauge Solution Accuracy*' when '*Assessing Solutions*'. This assessment is part of a process to '*Maximise Solution Integrity*', and therefore '*Facilitating Tactical Picture Generation*'.

Despite the matching of marks to cuts being a key part of a TMA operator's workload, there is little capability to support this, such as the solution accuracy display. This causes an increase in cognitive workload for operators as they attempt to manoeuvre the speedstrip to align the marks and can lead to frustration as they can visualise what solution they would like but cannot easily enter it into the computer and get the marks to align. Therefore, the capability to 'draw' solutions directly, instead of manoeuvring an existing one should be added. Additionally, given that it should be a trivial issue for modern computers to automate the alignment, the relevant intersections could be computed automatically from the speedstrip, adding marks at the intersections of the two most recent cuts that are overlaid by the speedstrip. Two intersections instead of one to determine velocity between the two intersections. In general, it should be considered a priority to ensure that the interface completes tasks that can be automated wherever possible, especially if there is little, or no, benefit to the operator completing them manually.

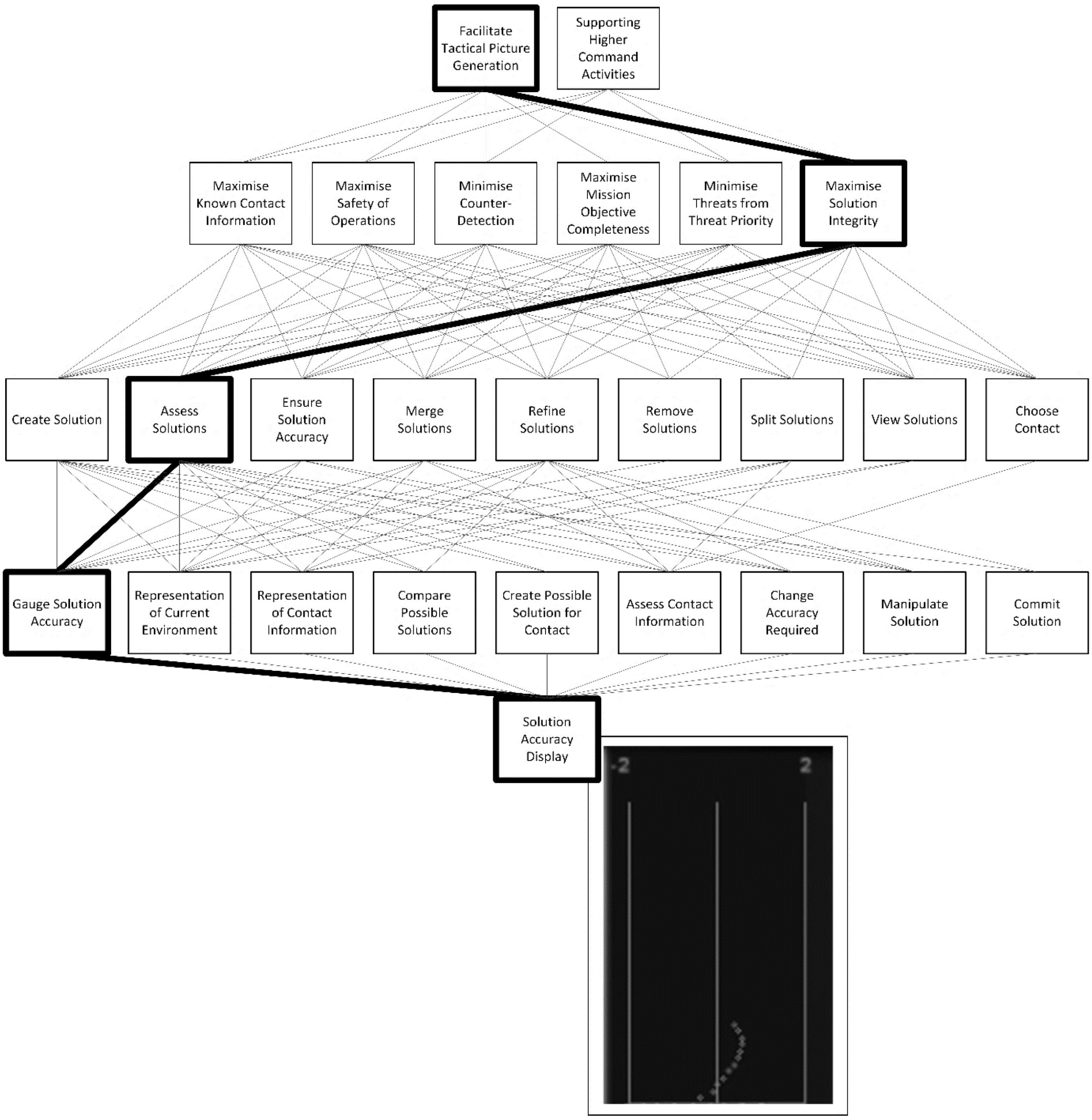


Figure 41 – Solution Accuracy Display Abstraction Hierarchy

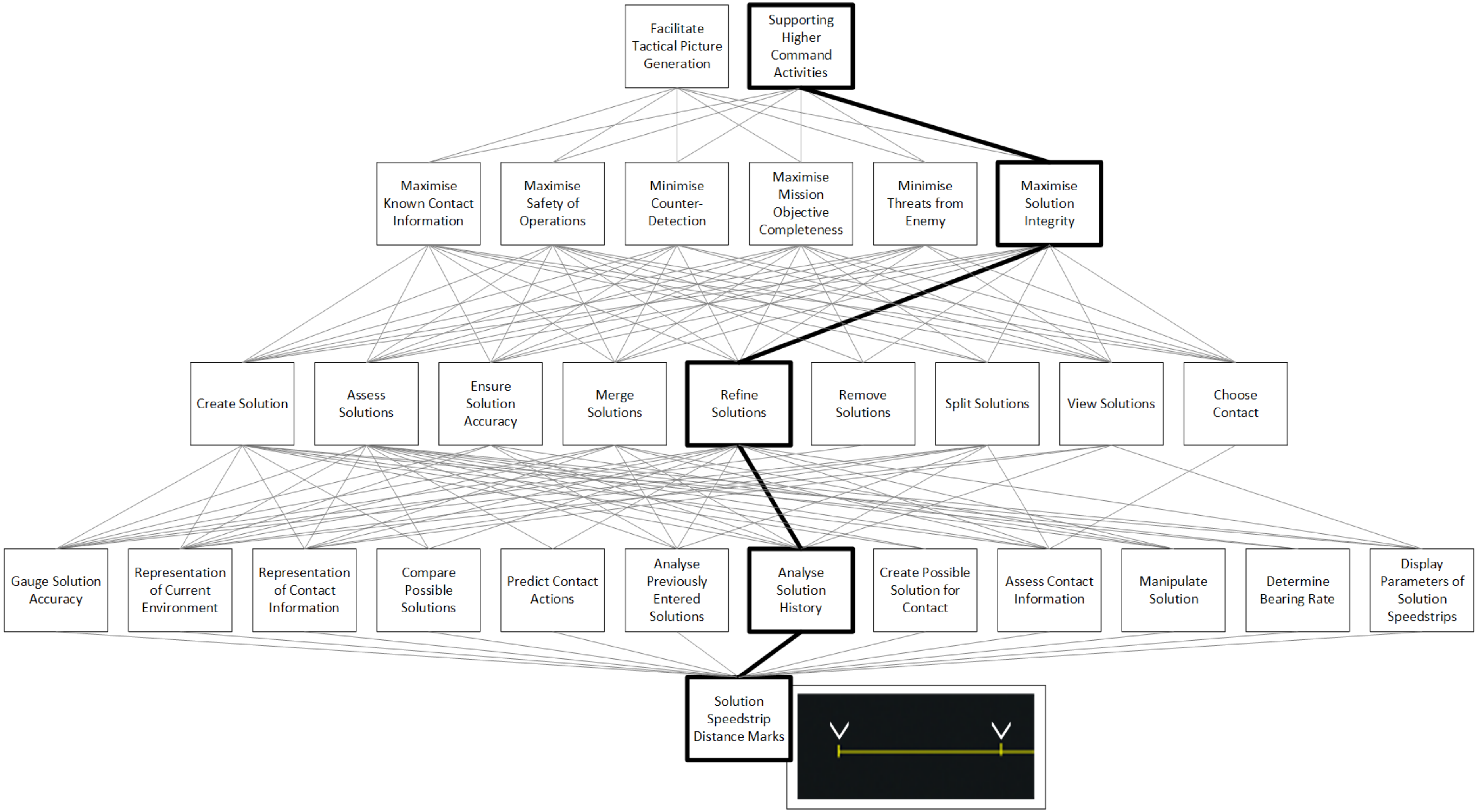


Figure 42 – Marks Abstraction Hierarchy. The white chevrons indicate where the marks are located and are not part of the Physical Object.

6.2.4.4.3 Arrowhead

The solution arrowhead is the leading end of the speedstrip, indicating a solution's course. It does not indicate the last detection. It is used to manipulate the solution direction by dragging it around the map to align cuts and marks visually. Figure 43 shows an example arrowhead, with associated abstraction hierarchy nodes. The *'Arrowhead'* allows an operator to *'Determine the Course Direction'* of a solution. As a vessel's direction could generally be determined by the order of cuts, the *'Arrowhead'* can be used when *'Creating a Solution'* to ensure it is heading in roughly the correct direction. By plotting the correct directions for contacts, their movements can be determined, and they can be avoided. This can 'Maximise the Safety of Operations', which 'Supports Higher Command Activities' by ensuring the safety of ownship.

Like other components providing information to the operator, they arrowhead provides physical information to the operator showing the direction of a solution, but stops short of providing associated functional information, such as whether the direction is possible. As cuts have a time associated with them, the interface could warn operators if their speedstrip does not align with the data. For example, if operators placed the arrowhead at an old cut, the solution would automatically be incorrect, as they are using stale data, but the interface does not warn them. Generally, there is an opportunity for automation to be implemented to support operators in validating logical constraints affecting solution creation.

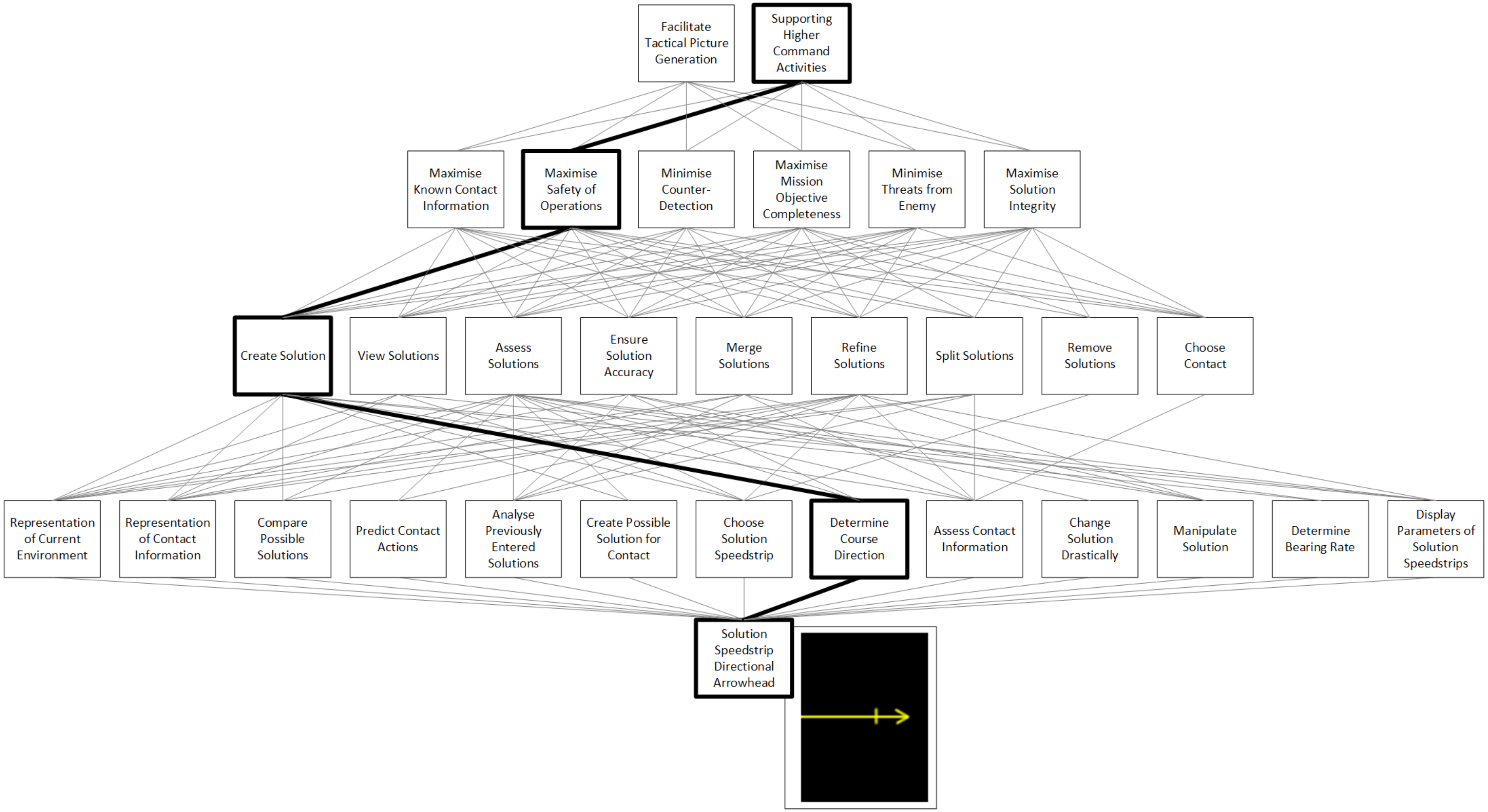


Figure 43 – Arrowhead Abstraction Hierarchy

6.2.4.4.4 Movement Circle

The movement circle is in the middle of a speedstrip and facilitates speedstrip translation through dragging and dropping. Sometimes a cut appears behind the circle, changing its appearance from '*o*' to '*ϕ*', however, this does not affect functionality for either. Figure 44 shows an example movement circle, with associated abstraction hierarchy nodes. The '*Movement Circle*' allows an operator to '*Manipulate a Solution*' to '*Ensure Solution Accuracy*'. A vessel's position can be known with an accurate solution, allowing it to be avoided to *'Minimise Counter-Detection'*, and therefore '*Supporting Higher Command Activities*' by evading detection. While operators can manipulate solutions using the movement circle and arrowhead, this interaction can be cumbersome and incur additional cognitive workload. An example of this would be if an operator wanted to move the solution to match a proposed solution they had visualised in their head. They would need to utilise multiple controls to translate and transform the speedstrip, an operation which might be better supported by allowing operators to 'draw' a new speedstrip in the desired location.

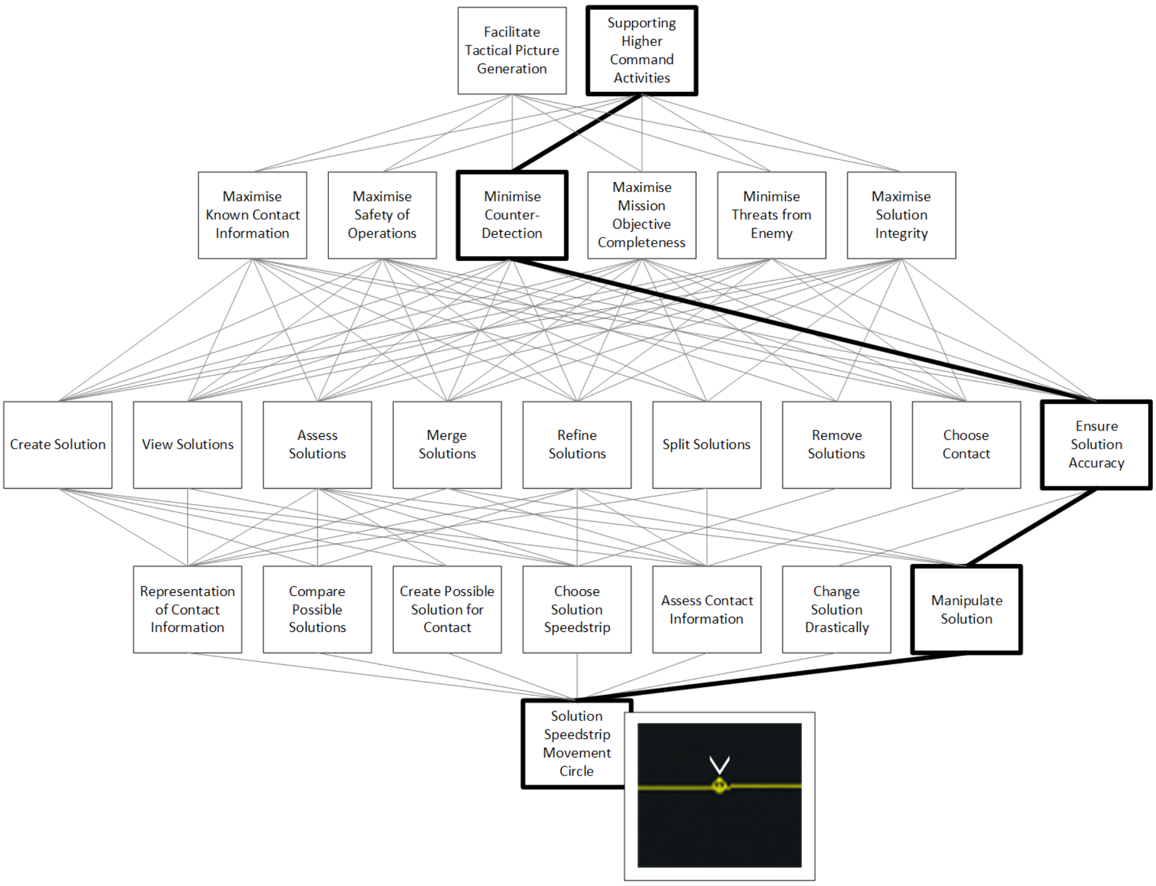


Figure 44 – Movement Circle Abstraction Hierarchy. The white chevron indicates where the circle is located and is not part of the Physical Object.

6.2.4.4.5 Solution Parameters

The four components of a solution (bearing, course, range, and speed) and the associated bearing rate are shown in the solution parameters area. Bearing rate is the magnitude of bearing change for a solution in a given direction, either port (left) or starboard (right). Figure 45 shows an example working solution parameter display, with associated AH nodes. The '*Solution Displays*' provide numeric data from each speedstrip, allowing an operator to '*Assess Contact Information*'. Based on this information, they may '*Refine a Solution*' to '*Maximise a Solutions Integrity*'. This '*Supports Higher Command Activities*' by ensuring the most up-to-date and accurate solution is always available. Each component is displayed in a text box. For the working solution display, these values can be changed to known information, such as speed from sonar analysis. However, despite the speed being generated digitally on a Sonar operator's interface, it must be communicated to the TMA operator. This requires the use of working memory if the operator must switch the relevant contact and enter the data provided, especially if data on multiple contacts is provided at once due to communication bottlenecks identified by Stanton and Roberts (2019).

Component solution values can also be locked if the information is known. This restricts a speedstrip's degrees of freedom, which in turn reduces the number of available speedstrip positions, enabling a solution to be found easier. For example, the speedstrip would have a fixed length if the speed were locked. Operators can then match the fixed space marks to the cuts at the most appropriate location. However, the interface does not provide any support for operators to do so, despite the problem being especially suitable for algorithms that could show possible solutions, or possible areas for solutions based on the current restrictions.

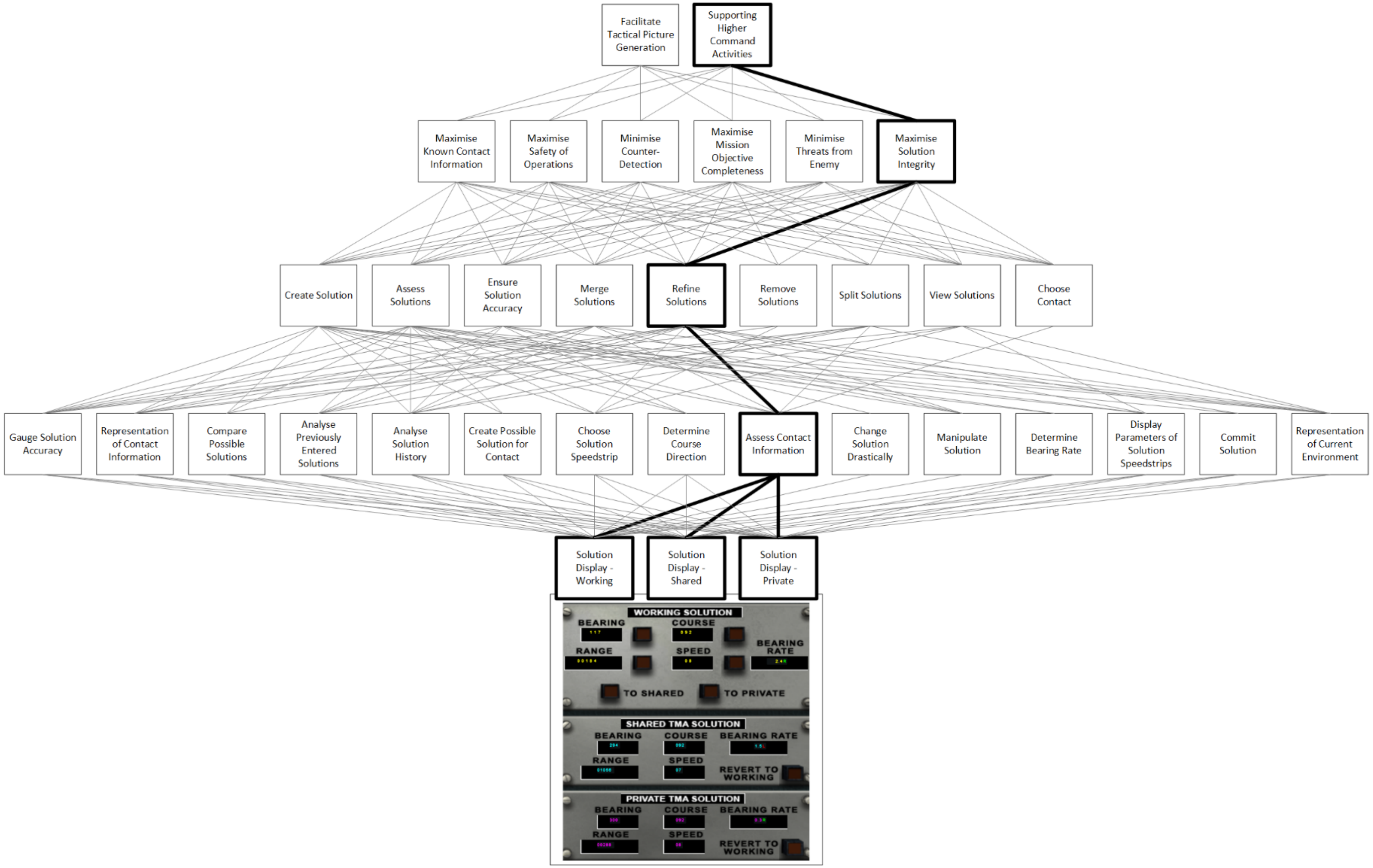


Figure 45 – Solution Parameters Abstraction Hierarchy

6.2.5 The TMA screen: Operation and Issues

Figure 46 shows a complete LOP screen, with areas of interest separated by bounding boxes. The LOP for a contact is in Area "A", with three speedstrips present. This means that the operator is working on a solution, and has saved a previous version, in addition to sharing a previous version. The working speedstrip can be moved around the map space to determine if there are any better matches, using the accuracy display to help guide this process. However as detailed above, this process is still completed manually, instead of the interface supporting the operator using the complex data it has access to. This unnecessarily increases operator cognitive workload, which could detrimentally affect their performance and negatively affect ownship safety.

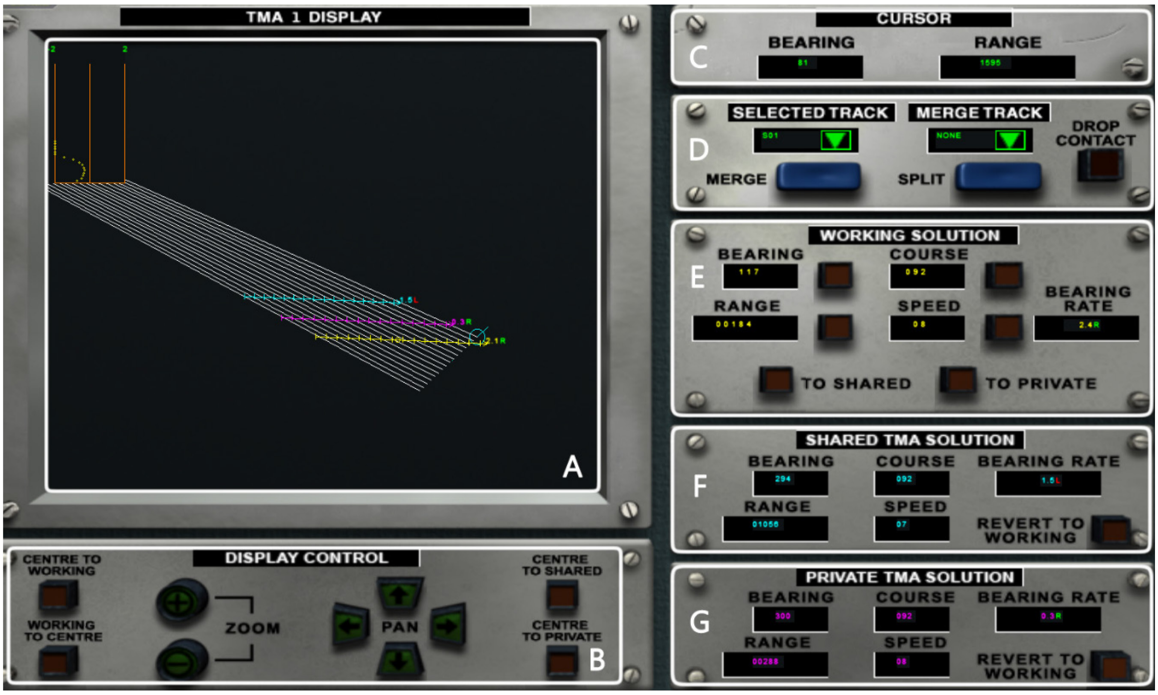


Figure 46 – TMA LOP User Interface

The controls for interacting with the map are shown in Area "B". Pan and zoom controls are available to change the map perspective. The "Centre to" button moves the map so that the specified speedstrip is in the centre, and the "Working to Centre" button moves the working speedstrip to the centre of the current display. Area "C" displays the bearing and range of the LOP cursor location. This can be used when manipulating the speedstrip to determine geospatial properties. While there is comprehensive capability to navigate around the map, it could be replaced by direct manipulation of the map. The removal of the buttons would also provide more space for the map view, allowing the operator to perceive more information pertinent to creating the solution.

Track management is located in Area “D” and enables the selection of a track to generate solutions for. A second drop box is available to nominate a second contact for merging. This will display both sets of cuts in the LOP, allowing the operator to determine if the cuts are similar, a sign that the contacts refer to the same physical contact. These sets of cuts should be from different sensors, as one sensor cannot detect a contact twice. Thus, it would be illogical to merge contacts from one sensor. If the operator identifies a match then they will press the “Merge” button to fuse the two contacts together, into a new master contact. The master contacts sensor identifier will change to “M” to denote that it is a master record. This process is reversible, using the ‘Split” button. As described above, there is opportunity to introduce more automation to support the operator by validating merge decisions. This would be supplemented by displaying information about each contact readily to operators so that they can better perform the task of contact management, instead of the list of contacts, and the associated information of how many there are, being effectively hidden in dropdown menus.

Areas “E – G” display the parameters of one speedstrip each. Each component is represented as a label. The working display, Area “E” allows for editing solutions. These values are bound to the speedstrip, so changes to one will reflect in the other. Additionally, each component can be fixed using the button to the right, narrowing down speedstrip movements to permit more effective solution finding. The “To Shared” and “To Private” buttons of Area “E” will respectively share or locally commit the working speedstrip. The working speedstrip can be reverted to the positions held by either the shared or private speedstrip using the “Revert To” buttons of Area “F” and “G”. This process is manual and increases cognitive workload to manipulate the speedstrips. It can also introduce an element of frustration for operators if they have a location visualised, but cannot easily enact this in the interface. As with the merge support, automation should be integrated into the interface to assist operators with these tasks as they are data-based. This would free up the operator to concentrate on SBB, such as determining why a contact is behaving as such, and predicting its future movements based on this, which in turn could maximise adherence to the higher-level goals of *‘maximising safety of operations’* and *‘maximising known contact information’*.

6.3 Worker Competencies Analysis

The WDA has furnished an understanding of how TMA could be improved, based on examining how work is completed using the LOP to contextualise improvements. However, WDA does not reveal the skills required to use the system so that they could be factored into the design process. A WCA (see Table 19) revealed these aspects, and how operators of different skill levels will interact with TMA. For example, when ‘*Creating a Solution*’ novice operators might not use precise adjustments of solution parameters using the inputs available in Figure 46e based on operational experience, instead preferring to manipulate the speedstrip directly in the LOP to experiment with what parameters match (Figure 46a). Both types of interaction should be supported by a redesigned interface. An understanding of the full spectrum of competencies required is especially important when considering automation, as WCA can highlight competencies that operators are expected to utilise, but have a computational basis that might be better suited to automation by a computer, informing further design requirements (McIlroy and Stanton, 2012).

The WCA shows that there is support for operators at all skill levels. It confirms that there are behaviours required from operators that might be better suited to computational processing and automation. These behaviours are mainly from the Rule-based Behaviour (RBB) column, which represents responses based on specific triggers, but automation could provide support at all levels. The suitability of these behaviours for automation arises from the data-driven aspect of TMA, meaning that most triggers for the RBB would be based on conditions that the interface could monitor for, and could provide a response that was quicker and more accurate than an operator completing the same task. This does not suggest that operators are not capable, but rather that they should not have to complete tasks that their interface could complete. Thus, a design suggestion is to look at implementing RBB where appropriate into the redesigned interface to support operators operating at that level. In doing so, this could free up capacity for operators to utilise Knowledge-based Behaviour to handle unanticipated and/or novel events where required, shifting from routine Skill-based Behaviour arising from extensive training.

Table 19 – WCA for a TMA Interface

	Skill-based Behaviour	**Rule-based Behaviour**	**Knowledge-based Behaviour**
	Observed behaviour of experts	If-Then rules to identify if a task should be done	Prerequisite knowledge/capability for novices to complete tasks
Choose Contact	Experience of knowing which contacts to work on, and when	Choose priority contacts when they are identified Choose contacts that may be dangerous (either by proximity or hostility)	Select a contact to use Select a contact to request information on
Create Solution	Experience of typical solutions for contact types Experience of patterns in cuts, which suggest solution components	Revert solution if no better solution can be found Utilise information from other operators	Manipulate Working Speedstrip Share working solution Save a working solution

		to restrict speedstrip degrees of freedom	
View Solutions	Make an informed decision on a solution to work on, based on command team activity	Look at solutions to determine which to work on	Look at a solution when ordered to
Assess Solutions	Know how different changes will affect solutions to make changes quicker and more accurately	Check if a solution is no longer accurate Check accuracy if a solution has not been updated for an extended period Check solution after changes in course or speed	Check solution suitability
Ensure solution accuracy	Use experience of contact types to determine if solutions are reasonable	Utilise information from other operators to restrict speedstrip degrees of freedom Correlate information from other operators to ensure that the solution is within the bounds provided by their sensors or intelligence	Check marks match cuts Check accuracy display
Merge Solutions	Ensure that solutions are merged properly	Recommend contacts to merge based on similar cut profiles Recommend contacts to merge based on similar solutions	Combine contacts when ordered
Refine Solutions	Experience of how changes to the environment, contacts, and own ship would affect solutions	Manipulate speedstrip for a better fit when it is no longer suitable Manipulate speedstrip for a better fit when changes to contacts are detected by other stations Do not merge contacts on the same sensor	Manipulate speedstrip for a better fit when ordered
Remove Solutions	Experience of what contacts are not required	Drop contacts that are no longer required	Drop contacts when ordered

		Drop contacts that should not be tracked, such as schools of fish, or pods of whales	
Split Solutions	Experience of when a split is incorrect	Identify erroneously merged contacts to split	Separate contacts when ordered

6.4 Insights Gained

6.4.1 Automation is Key

A key aspect of creating design directions for a new TMA interface has been a focus on the introduction of automation so that operators do not have to manually complete tasks which could be completed using automation. These tasks are especially suited to automation as TMA is largely a computational process (Punchihewa *et al.*, 2022). This has led to multiple algorithms and processes for solution generation in TMA being proposed (Aidala, 1979; DeAngelis and Green, 1992; Lee *et al.*, 2008; Genç, 2010; Punchihewa *et al.*, 2022). While these algorithms have been shown to be beneficial, it would not be appropriate to completely remove the human operator from the process. This is because while manual TMA is operator-intensive, it is useful when the efficacy of algorithms is degraded (Huf, Arulampalam and Manning, 2006). Additionally, TMA incorporates 'soft' data that must be processed by an operator before being utilised (Punchihewa *et al.*, 2022). While contemporary automation is very capable, it might not have the same capabilities that operators naturally have to process this data. For example, a Sonar operator might be able to determine a rough speed of a fishing vessel from intuition, which would be passed to the TMA operators as a guideline for them to experiment with in their solution. Hou *et al.* (2015) also noted that while automation can be introduced to mitigate a limited attention span arising from limited working memory capacity, removing operators from decision-making tasks can degrade efficiency (Clare *et al.*, 2012). This suggests that operators should be provided with automation but should retain involvement in the processes being carried out or utilise the automation to support them when workload increases.

This approach is congruent with the suggestions by Huf, Arulampalam and Manning (2006) and Huf *et al.* (2009) that operators could perform their tasks under a normal workload but might rely on the assistance of automation to cope with a more hectic picture compilation process. Hou *et al.* (2015) built on this by also proposing that automation could detect when an operator was attempting to complete a task and complete it for them if they were under a high workload. The utilisation of automation for track management has been shown to be beneficial to task performance and workload, although impaired non-automated task performance (Tatasciore *et al.*, 2020). However, it has been shown that increasing the degree of automation does not further detrimentally impact situational awareness, non-automated task performance, or return-to-manual performance (Tatasciore *et al.*, 2018; Tatasciore *et al.*, 2020). Tatasciore *et al.* (2022) posited that the slow-moving nature of tasks and contract tracks was an enabler for maintaining situational awareness and return-to-normal performance, when compared to the faster evolving tasks such as air traffic control, unmanned vehicle control, and driving. Alone, this would suggest that as much as possible should be automated, as the detrimental effects would remain constant. However, Tatasciore *et al.* (2022) found that more comprehensive automation for a task produced the expected benefits but caused poorer automation failure detection. Given the direct role of TMA in maintaining the tactical picture, an unnoticed failure in automation designed to yield information pertinent to this is undesirable. Therefore, the automation might be better implemented as small chunks of capability that the operator can immediately tell if they have failed, such as generating a solution for a contact given the current data. While this work structure (manual, supported by automated tools) might incur a workload cost for the operator, this might be preferrable for operators so they can maintain their situational awareness for quick reactions where required. This is important, as while contacts can move slowly (Tatasciore *et al.*, 2022), transient factors in the environment such as sudden starts (the abrupt appearance of a previously undetected contact) and the temporal window for accident avoidance through problem recognition and remediation can be fast-paced events. Another benefit of implementing automation that responds to an operator request could be that the operator will be focused on the output, allowing the automation to include any doubt, which could mitigate the automation bias (Wickens *et al.*, 2015; Man *et al.*, 2018). Automation bias can be defined as the tendency to use automated cues in lieu of vigilant attention seeking and processing, resulting in errors arising from a failure to detect errors (Mosier and Skitka, 1996).

6.4.2 Increase TMA Automation & Enforce Solution Parameter Verification

WDA showed that TMA exists to facilitate tactical picture generation and to support higher command activities. TMA had global measures of success, which were to maintain a submarine's three tenets, and local measures of success to ensure the provision of an accurate tactical picture. The speedstrips, core components of the LOP, contributed heavily towards this goal by enabling solution creation. They functioned as tools that enabled information from cuts to be transformed into a full solution, which contributed heavily to submarine goals. Presently, it is possible to manipulate the speedstrip to express a solution. However, this allows inaccurate, or plainly wrong solutions to be entered, as no verification is performed. Given the safety-critical nature of TMA, this poses a risk. Suppose that a speedstrip does not intersect a lead cut, which shows the last known bearing of a vessel. The solution generated is likely to be inaccurate, as the most recent information has not been considered. Should the contact be manoeuvring, a negative situation could develop, violating the goal of keeping ownship safe. Furthermore, this could mitigate instances where an operator's situational awareness does not include the newest cut information. This could occur if they were zoomed into a specific map area and did not see new cuts appearing. Despite possible negative consequences, the operator is not prompted to make changes, despite the system being able to measure a solution's accuracy. A deviation of solution accuracy over time may indicate an incorrect solution, which, again, poses a safety risk. Issues such as these suggest that while TMA has the capability, it does not validate operator input against known parameters. However, this may be by design, preventing the system from overriding what could be a correct solution and causing a potentially dangerous situation. Thus, additions should be made to the system that seeks to inform operators of potentially wrong solutions and to recommend a better solution that could be accepted or rejected by operators. This could be implemented as automation that seeks to validate operator solutions by processing the data, allowing the operator to choose from possible solutions. Operators would potentially not have to manage the underlying data but can still choose the solution that matches their desired input. Furthermore, this may improve their situational awareness by making them aware of potential solutions for each contact, enabling them to consider each variation.

6.4.3 Removing Cumbersome Interaction

The analysis revealed that there were pain points in entering a solution that could be addressed in a redesigned interface. A common issue was that the operator could visualise a solution in their mind but could not readily commit that to the interface as they would like. It was proposed that functionality be added to facilitate operators in entering solutions. The approach of aiming for optimal interaction throughout is pertinent throughout the interface, not just for solution entry, however, and improvements should be pre-emptively made where possible. While operators are successful at using the interface despite pain points, having an "I can make it work" mentality as military users (Grier, 2013), the onus should not be on them to work around issues, and any redesign should work towards minimising the time operators spend on cumbersome interaction, freeing the capability to utilise their training and experience for goal-oriented behaviour.

The potential introduction of automation across TMA will require the HMI to be updated. Any automation introduced should be carefully evaluated to ensure that outcomes are positive, as this is not always the case. Care should also be taken to preserve the information processing capabilities of novices, ensuring that automation does not reduce their capability, nor detract from their ability to gain experience. Bainbridge (1983) argued that operators might not be able to sufficiently able to assume control if they have not been reviewing their knowledge or practising skills (known as skill fade). Thus, the introduction of any automation should be carefully balanced against operator readiness and training requirements.

Whilst the automation could be added to the existing TMA HMI, this may not be the best approach; a redesign that incorporates automation, as well as addressing existing usability issues, may be better in comparison. For example, the complexity of a speedstrip was revealed by the WDA, illustrating that there are multiple controls (circle, line, arrowhead) for manipulating it. However, this manipulation can be cumbersome, as both the end and start points must be moved to the appropriate locations. Furthermore, in some situations, the operator will have identified marks to align with cuts, but currently must do this manually. This manual manipulation could increase operator workload and increase the time taken to create solutions, even with automation. Given that solution manipulation is a large component of TMA, it is pertinent to address this usability issue through a holistic redesign, instead of simply adding more capability without full consideration of its integration (Hall, 2012). A more natural interaction could be for the operator to drag their mouse over the desired solution path and have the system automatically fit the speedstrip. If TMA is to be designed with automation in mind, it would be pertinent to assess current and proposed new interactions alike to ensure they are optimal for operators, minimising the required cognitive workload.

6.4.4 Ensure Operators Can Still Learn & Develop

The WCA further reveals how automation, and associated HMI changes, could support operators by assisting them at all stages of the SRK Taxonomy, ensuring no more cognitive demand than required is needed, helping to reduce human error (Breton and Bossé, 2003). These benefits could help to avoid dangerous incidents potentially arising from a lack of situational awareness (Danczyk *et al.*, 2015). Automation could benefit knowledge-based (novice) operators by checking their solutions to ensure that they are reasonable. For example, the system could inform them that their speed may be incorrect, as a speed of one hundred knots is improbable. This would facilitate operators gaining experience, as they would learn the checks the system performs, and incorporate them into their solution creation. This could be represented in the HMI by changing the relevant solution component red to illustrate a problem or changing the ruler's colour to green to suggest that no improbabilities were found. As operators gain experience with creating solutions, automation could support the operator with managing contacts by suggesting which contact to work on with tooltips, recommending merges based on similar contacts by highlighting them in the dropdown box, and warning the operator when a contact solution is no longer accurate with a message. These recommendations could allow the operator to concentrate exclusively on generating solutions without negatively affecting contact management. Additionally, automation providing contact information and recommendations may improve operator SA/DSA when compared to the operator acting alone. Contact management would be incorporated into the operator's experience, eventually allowing them to manage both with little to no help from the automation. As operators approach a skill-based (expert) level, they would be proficient in managing contacts and entering valid solutions themselves. At this stage, more advanced automation could be used, such as Artificial Intelligence (AI) which learns how the operator creates solutions and proposes solutions if the operator is under a heavy workload. The automation could draw a recommended speedstrip, which could be either accepted or dismissed by the operator. These suggestions are not exhaustive, nor fixed to operator capability, rather they are a suggestion of how to exploit capable computing systems to support operators and ensure that the HMI facilitates this. A highly experienced operator may not want to use AI but would like their solutions sanity checked during periods of high workload (Huf, Arulampalam and Manning, 2006; Huf *et al.*, 2009). By supporting operators at each level of the SRK Taxonomy, automation can also ensure that no more cognitive effort than required is expended. The provision of supportive measures, suitable to capability and workload, means that operators will be supported by the TMA HMI and underlying automation instead of being hindered by it.

6.4.5 Summary of Insights

A summary of design insights proposed throughout this chapter is presented in Table 20 for convenience.

Table 20 – A summary of insights gained from conducting WDA and WCA on TMA

Section	Insight
6.2.1 6.2.2	Add as much information and functionality as possible to support and/or enhance tactical picture compilation.
6.2.2	Information be better integrated to make it more accessible for analysis and decision making.
6.2.3	Implement digital information data sharing mechanisms for operators where possible.
6.2.3	Add functional information to solutions to show their future state.
6.2.4	Warn operators if an entered merge appears incorrect.

6.2.4.3	
6.2.4 6.2.5	Always make contacts and their associated information visible.
6.2.4 6.2.5	Allow the operator to see an overview of all contacts at once, instead of one contact at a time.
6.2.4.1	Consider merging the Sonar and TMA displays, as they will both use a map as a core component.
6.2.4.2	Visually represent sensor geometry in conjunction with ownship.
6.2.4.3	Arrange contact information so that merge candidates can be identified easier.
6.2.4.4	Add direct manipulation to speedstrips.
6.2.4.4.1	Allow multiple legs for speedstrips.
6.2.4.4.2 6.2.4.4.4 6.2.4.4.5 6.2.5	Add the capability to 'draw' solutions on the map and automatically enter intersections.
6.2.4.4.2 6.3	Automate tasks where there is little, or no, benefit to the operator completing them.
6.2.4.4.3	Warn operators that the speedstrip direction is not congruent with the cuts.
6.2.4.4.3	Validate entered solutions using logical constraints where possible.
6.2.4.4.5	Add the ability to trial possible solutions to see if they match the constraints created by the available cuts.
6.2.5	Add more direct manipulation to the map to remove control buttons, freeing up more space for the map.

6.5 Conclusions

This chapter has explored the LOP as a leverage point for synthesising insights on how TMA could be redesigned. It also looked to explore pain points within how LOP was used, as cumbersome interaction can frustrate users and increase their cognitive workload. These are both points of focus from the CWA Design Toolkit (Read *et al.*, 2018). The LOP was focused on as it is a key aspect of TMA (Clarke, 1999), meaning its redesign could offer benefits across the system.

Following the method proposed in Section 5.2, WDA and WCA were conducted, and their outputs were presented in this chapter to derive design insights. These insights are presented in prose and are summarised in Table 20. The WDA examined the key Physical Objects of the LOP to contextualise their operation, allowing for a formative understanding of where improvements could be made. As with Sonar, this was done to facilitate an understanding of why the improvements were pertinent, as TMA is not a widely understood concept like driving. From the analysis, it is apparent that TMA is primarily a data-driven process underpinned by advanced combat systems, which presents an opportunity to reduce operator workload by implementing automation to support operators. However, the automation should be carefully designed to ensure that its utilisation does not provide opportunity to negatively affect the tactical picture compilation process, and consequently ownship safety. The submarine control room literature, see Section 6.4.1, suggests that automation might be best provided as tools to support the operator, keeping them in the loop instead of automating as much as possible and relying on operators detecting an issue with the automation to take over. Another theme for change was that interaction issues could also be addressed to ensure optimum usability. The redesign of TMA offers an opportunity to consider the interaction mechanisms present and address any usability issues that might exist to improve operator experience and reduce their workload. In turn, this would free up the capacity to utilise their knowledge and experience. Automation was also proposed to address cumbersome interaction, removing the need for operators to manually complete tasks that their interface should be readily capable of. Finally, as modern combat systems continue to exist at the vanguard of function and innovation, it is vital to exploit their capability to further the potential of highly capable operators as well as to maintain ownship safety. However, these changes cannot be to the detriment of operator training and capability. Therefore, it was proposed that the HMI also assist operators with maintaining their skills and continuing to develop, by offering them tools and functionality appropriate across all skill levels. The next chapter will explore a comparison of the WDAs presented for Sonar and TMA to understand if the work domains are reasonably comparable to the real-world, ahead of using them for experimentation.

Chapter 7 **Validation Against the Talisman Trainer**

7.1 Introduction

Chapter 5 and Chapter 6 presented the results of the Cognitive Work Analysis (CWA) conducted on the Command Team Experimental Testbed (ComTET) versions of Sonar and Target Motion Analysis (TMA). As the facility was reasonably new at the point of creation, work was ongoing to validate its relative fidelity against the Talisman trainer at HMS Drake. It was decided to complete an additional WDA on sonar and TMA Human-Machine Interfaces (HMIs) in Talisman and compare them against their counterparts to determine if there were appropriate similarities in how the domains functioned (St-Maurice and Burns, 2018). The aim was to validate the fidelity, defined as the degree to which a simulated environment matches its corresponding real-world environment (Alessi, 1988; Gross and Freeman, 1997; Hancock *et al.*, 2008; Roberts *et al.*, 2020), of the ComTET Sonar and TMA implementation that this thesis investigates. This was important to verify that results could be generalised using relative differences, following the approach of using novice participants to help increase statistical robustness (Walker *et al.*, 2010c; Stanton and Roberts, 2019).

The term 'relative differences' is used in recognition of the different purposes of the facilities being compared. The Talisman trainer is a high-fidelity submarine training facility where real-world activities are conducted. This purpose was different to ComTET, which was an unclassified academic research facility for researching new ways of working using representative systems. An example of this would be the use of the TMA screen in ComTET, compared with the use of SMCS in Talisman (Stanton and Bessell, 2014), a vastly more complex product, acting as a whole combat system. These differences also influenced procedures used, which as with the interfaces, were designed to be abstracted sub-set representations of their real-world counterparts, designed appropriately for the experimental environment. Consequently, there was a difference in the precise nature of activity between the facilities, but the building blocks remained the same; therefore, relative deviations from baselines for each facility could reasonably be expected to be transferable. For example, if co-locating operators in ComTET increased productivity in terms of task completion (Roberts, Stanton and Fay, 2018; Roberts *et al.*, 2019), this could reasonably be expected to show increased productivity in Talisman (or other real-world environments).

Fidelity is not a monolithic construct, rather it is the interaction of different dimensions (Roberts *et al.*, 2020). Of specific interest for this chapter was assessing functional fidelity, defined as whether critical aspects of an environment are modelled (Roberts *et al.*, 2020), and how internal mental models correspond to a task's cognitive nature (Estock *et al.*, 2006). Specific tasks were not compared as WDA is not a task-oriented analysis (Naikar, Hopcroft and Moylan, 2005), but rather the work domains and the effective constraints within them. That is, determining whether both implementations provide comparable means and work domain structure. This can be illustrated using the "ant on the beach" metaphor (Rasmussen, 1974a; Waterson, Le Coze and Andersen, 2017; Simon, 2019). It puts forth that the behaviour of an ant on a beach is a result of complexity in the beach, as opposed to the ant itself. Comparing the implementations of Sonar and TMA in both simulator facilities would be akin to assessing the differences between a beach and desert environment; the aim is to compare two different environments with different specific tasks to ascertain if they provide the same environment to bound work.

Also of interest was the physical fidelity of the roles. Physical fidelity is defined as the degree to which the physical environment emulates the real environment (Hancock *et al.*, 2008), citing (Allen, Hays and Buffardi, 1986). It was important to achieve a level of physical fidelity to ensure that an operator's physical environment felt realistic, achieving buy-in from operators (Roberts *et al.*, 2020), citing Alexander *et al.* (2005). However, the degree of fidelity could be lower when compared to task fidelity because consideration of the environment was not a primary focus of the experiment, rather exploring ways of working. Physical fidelity was scoped to representing the control room as necessary for examining this, with only key elements of a control room considered. For example, the control room's layout affects how operators can communicate and what information they can see. Therefore, it was important to replicate the control room space and operator locations within it for ComTET studies. Similarly, it was also important that MFCs were used to replicate interaction as accurately as possible over simply providing them with an office computer setup. However, it was not required to exactly replicate other aspects, such as the chairs used or heat levels, within the context of this experiment. Of course, it is recognised that these factors could contribute towards operator performance, although examination of these factors is outside the scope of the current work.

Moreover, there is a high degree of physical variance between submarine control rooms (e.g., different classes of boats, and designs by individual navies), and this variance could be extended to the ComTET simulator design. An existing submarine control room did not have to be replicated; rather, participants had to reasonably believe that they were in a submarine control room environment and have their ways of working shaped as they would be in a real-world facility. In the context of the HMIs, this meant that they should be provided with interfaces and component controls that could be considered to be reasonable facsimilia. Ensuring core physical objects were present would improve the applicability of experimental results, avoiding testing something that does not exist in the real-world environment.

This chapter will detail a validation exercise conducted to compare and contrast the different Sonar and TMA implementations. To achieve this, analyses of comparable stations (ComTET Sonar to Talisman Sonar and ComTET TMA to Talisman SMCS) were performed at HMS Drake's Talisman trainer. The ComTET and Talisman WDAs were then compared to understand where similarities and differences existed.

7.2 Method

7.2.1 Interviews

Interviews were held with five Royal Navy Submariners at HMS Drake's Talisman trainer over three days. All had served at sea for an average of 856 days (SD = 412.25) and had experience operating Sonar or TMA. Some participants occupied supervisory roles, ensuring holistic analysis. Participation was voluntary. The study protocol received ethical approval from the University of Southampton Research Ethics Committee (10099) and the Ministry of Defence Research Ethics Committee (MoDREC; 551/MODREC/14). The work was completed by the author of this thesis, supported by their supervisor(s).

Minimal demographic information was collected from participants, consisting of their age, rank, command structure, length of service, and time at sea. These were collected only to understand each operator's experience and role, ensuring that interactions at all levels were captured, in direct support of the study objective. As with other ComTET studies (Stanton and Roberts, 2018; Stanton and Roberts, 2020), more in-depth demographic collection was limited by security considerations and recommendations from the Ministry of Defense Research Ethics Committee (MoDREC).

Participants received drawing equipment (pen, pencil, ruler, rubber, and paper) to utilise freely for communicating concepts. Printed copies of abstraction hierarchies for ComTET's Sonar and TMA stations were available. It should be noted that TMA functionality in ComTET is a Local Operations Plot (LOP; see Section 4.1.1.5 for details). As the LOP is a subset of the functionality of the Submarine Combat Management System (SMCS) in Talisman, SMCS was compared to ComTET's TMA. Interviews were recorded using an Olympus WS-831 digital voice recorder, with an additional external microphone attachment. Notes were recorded by interviewers using a notepad and pen.

For the interview, participants were asked to describe their job role, and how they used the available HMIs in their Multi-Function Consoles (MFCs, which are computer workspaces in submarine control rooms; Rhie *et al.*, 2017) to achieve their goals. They were encouraged to create supplementary material to aid interviewer understanding and construction of the abstraction hierarchy. One Abstraction Hierarchy (AH) each was constructed for Sonar and TMA, with all participants contributing to them, instead of starting a new abstraction hierarchy for each participant. Each interview was scheduled to last two hours, although this was fluid to allow more time for information elicitation, or less time for emergent requirements. The time was provisioned as per Table 21.

Table 21 – Talisman Interview Schedule

Activity	Minutes	Description
Introduction	10	Description of research and objectives
Role Scoping	30	Interviewee's role defined and scoped
Sketching	30	Sketching of console interface
Means-Ends Analysis	30	Creation of an abstraction hierarchy
Debrief	10	Quick debrief with the chance for questions
Slack Time	10	Extra time, for any activity above

Before starting, participants were provided with an information pack containing an information sheet, an interview schedule, a demographic questionnaire, and a consent form. Upon signing to signify informed consent, and filling out the questionnaire, the interview began. Interviewers introduced themselves to the participants, and the aims of the study were reiterated. Detail about ComTET was provided for context. Interviewer notes were taken throughout to assist with AH construction.

Next, participants were asked to describe their role within the command team and specify objectives that should be met. For each objective, criteria were elicited that would indicate successful, or otherwise, performance. Participants could answer freely, although were asked to prefix their objectives with 'maximise' or 'minimise' to assist with defining clear objectives.

After their job role and objectives had been documented, participants were asked to sketch various interfaces from their console. They were asked to add as much detail as they could remember, however, they were not expected to create exact screen replicas. They could use any drawing equipment available. During sketching they were encouraged to speak aloud to match drawings to explanations. The next stage, means-ends analysis, was performed either in parallel or subsequently, depending on participant preference. Participants were asked what functionality each individual aspect of their sketched interfaces provided, and why it did so. More sketches could be drawn if needed, and these were explored in the same manner. For pivotal processes, participants were asked to write the process down using lined paper whilst explaining each point for clarification.

Once participants had finished explaining how their station looks and works, their attention was drawn to the corresponding ComTET AH diagram. It was explained, but they were not required to absorb all shown information. They were asked to identify any differences between their role and its ComTET equivalent. They were asked to disregard the bottom two levels, as the interfaces were different hence generating different Physical Objectives and Object-Related Processes nodes. However, if they had feedback, it was noted and incorporated.

Finally, the participants were debriefed. The interview's purpose and what had been covered were reiterated. Participants were asked if they had any questions, and then thanked for their time, concluding the interview.

7.2.2 Observations

To ensure holistic coverage of stations and to mitigate unintentional omissions from interviews, a series of observations spanning multiple days in the trainer were also conducted, enabling experimenters to place their understanding of the system into the context of actual usage. The observations were ad-hoc and covered common tasks that operators were being trained on or were demonstrating to the experimental team. Their purpose was to physically observe each console, and to consolidate experimenter knowledge. Personnel also volunteered their free time during simulation runs to either provide a running commentary of interface usage, or to demonstrate their usage, answering questions in situ.

Experimenters used a notepad and pen to capture elicited information. Operators were not interrupted if they were training. However, for voluntary station demonstrations, an operator would talk through specific procedures of importance, navigating through screens so detailed sketches as well as notes could be made, and guide experimenters through practical use of the system. This also included the use of narration by the ComTET researchers, who vocalised key points of action for later review. This was done in addition to the written notes, as this would enable synchronisation of both sources during review. There were no fixed timings or topics for observations. Rather they were held ad-hoc in response to training demands on the submariners, personnel availability, and experimenter queries.

7.3 Talisman Abstraction Hierarchies

The resulting abstraction hierarchies from the visit to Talisman are shown in Figure 47 (Sonar) and Figure 48 (SMCS). The top three levels are presented in full. However, the bottom two levels are redacted as a security consideration, avoiding describing the exact makeup and capabilities of the real-word HMIs.

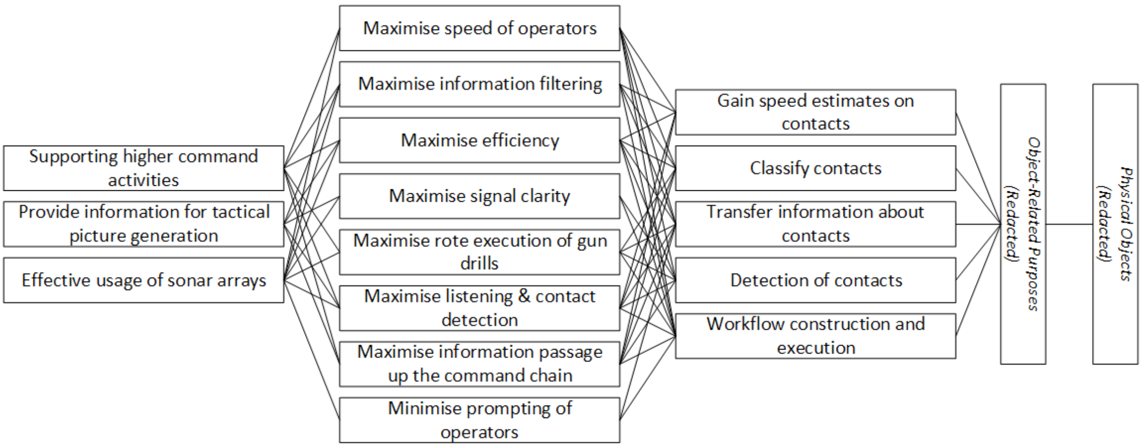


Figure 47 – Abstraction hierarchy of Talisman Sonar

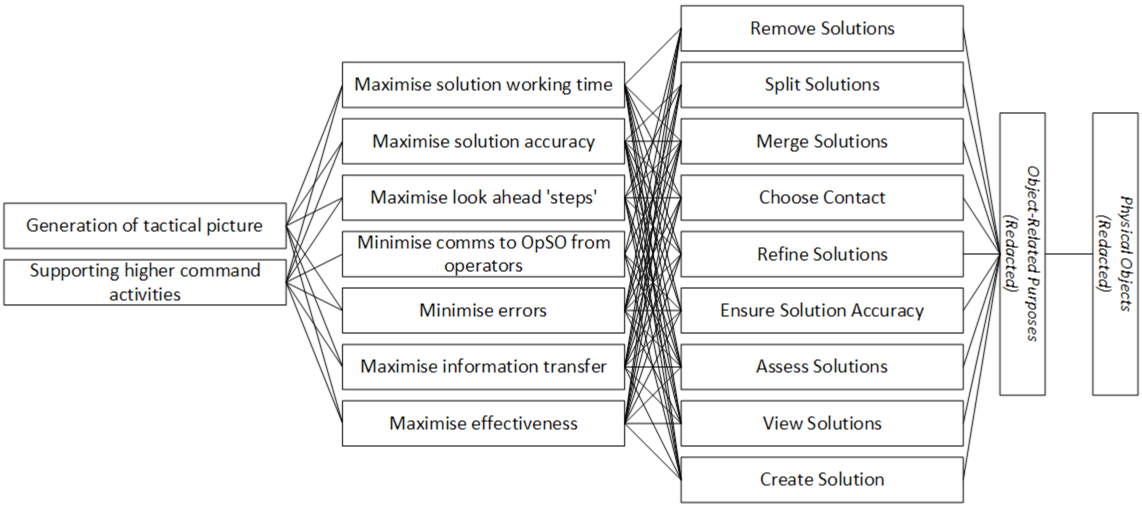


Figure 48 – Abstraction hierarchy of Talisman SMCS

7.4 Comparison of Abstraction Hierarchies

For each station, the nodes on each level were compared to their ComTET counterpart, similar to the approach taken by Burns, Bisantz and Roth (2004) and St-Maurice and Burns (2018).

7.4.1 Sonar

7.4.1.1 Functional Purpose

The overall goals of operators in both simulators were largely the same, although Talisman operators observed the additional goal of making effective usage of available Sonar arrays. In ComTET this requirement does not exist as the use of sensors is much simpler and is therefore not taken into consideration for the station.

7.4.1.2 **Values & Priority Measures**

There exists significant variability at this level. Only two measures are directly comparable, '*Maximise Contact Detection*' and '*Maximise Signal Clarity*'. The ComTET measure of '*Maximise Known Contact Information*' is split into two for Talisman: '*Maximise Information Passage up the Command Chain*' and '*Maximise Rote Execution of Gun Drills*'. The three remaining ComTET measures, '*Minimise Counter Detection*', '*Maximise Safety of Operations*', and '*Maximise Mission Objective Completeness*', were not included in Talisman's measures. This is because senior management take on these roles in Talisman, therefore absolving operators of a need to do so. However, this is not absolute, operators can feed information up the chain of command if they deem it necessary to directly influence meeting these measures. The remaining measures for Talisman all regard the efficacy of operators when using their station and communicating data. There are no direct comparisons for ComTET, however, efficacy is implicitly expected. It is hypothesised that the explicit inclusion of these measures for Talisman can be directly attributed to the values of the Royal Navy in being an effective defence organisation.

7.4.1.3 **Purpose-Related Functions**

The general function of Sonar is the Detection, Classification, Localisation, and Tracking (DCLT) of contacts. These are represented in both simulations by '*Detection of Contacts*', '*Classify Contacts*', '*Gain Speed Estimates on Contacts*', and '*Transfer Information About Contacts*'. In ComTET two additional general functions, '*Interpreting/Predicting Contact Actions*', exist. These are undertaken within Talisman, however, typically not by Sonar operators, and as such are not included. A General Function specific to Talisman is '*Workflow Construction and Execution*', which relates to measures pertaining to efficiency and accuracy. The emphasis on these measures requires that general functionalities be available to directly support this.

7.4.1.4 **Object-Related Processes / Physical Objects**

There were largely common physical objects and associated affordances for both systems. This is due to the prominence of waterfall displays. Generally, all waterfalls consisted of '*Bearing Tape'*, '*Time Tape*', '*Waterfall Area*', '*Traces*', and '*Contact Markers*'. While the representations of these Physical Objects differed, they were largely comparable. Similarly, their Object-Related Functions are comparable. Aspects outside of the waterfalls exhibited aesthetic variability, but afforded similar, if not the same, functionality for both simulations.

7.4.2 TMA

7.4.2.1 Functional Purpose

Both ComTET and Talisman had the same overall purposes, to '*Generate a Tactical Picture*' and '*Supporting Higher Command Activities*'. While there would have been other Functional Purposes for SMCS, a highly capable system, these were outside the scope of the focus on tactical picture compilation for this work.

7.4.2.2 Values & Priority Measures

Again, there exists significant variability at the measures level. However, comparisons are still present. The combined ComTET measures of '*Maximising Knowledge of Contacts*' and '*Maximise Information Gathered About Contact*' can be equated to Talismans '*Maximise Information Transfer*' and '*Maximise Look Ahead Steps*'. These measures are about collecting as much information as possible for a contact and ensuring that its future intentions are known. The ComTET measure '*Maximise Solution Integrity*' is present in Talisman as two separate measures, '*Maximise Solution Accuracy*', and '*Maximise Solution Working Time*'. As with Sonar, there were four measures included only in ComTET, as in Talisman they are performed by senior management: '*Maximise Safety of Operations*', '*Minimise Own-Ship Detection*', '*Maximise Mission Objective Completeness*', and '*Minimise Threats from the Enemy*'. Again, Talisman had explicit measures of effectiveness and accuracy. These were '*Maximise Effectiveness*', '*Minimise Errors*', and '*Minimise Communications from the Operations Officer*'.

7.4.2.3 Purpose-Related Functions

The general purpose of TMA is generating and maintaining the tactical picture. This is represented by the general functions '*Choose Contact*', '*Create Solution*', '*View Solutions*', '*Assess Solutions*', '*Ensure Solution Accuracy*', '*Refine Solutions*', '*Merge Solutions*', '*Split Solutions*', and '*Remove Solutions*' in both simulations. There were no simulation-specific general functions.

7.4.2.4 Object-Related Processes / Physical Objects

Base components were found to exist in both simulations, such as '*Contact Markers*', '*Cuts*', and '*Own-ship'*. However, due to the extended capability of Talismans' SMCS, compared to ComTETs' TMA, similarities existed on a more atomic level than Sonar. This means that while Physical Objects represented similar objects, with similar Object-Related Processes, how they were arranged led to highly divergent interfaces. As expected, the closest similarities were observed comparing LOP implementations, with different areas of Talisman's interface diverging to greater extents.

7.5 Discussion

In both simulations Sonar and TMA exist for the same reasons, but their measurements of success are different. It is thought this disparity is a result of their environment. Talisman is a Royal Navy facility, so its measures of success will be derived from the Royal Navy, whereas ComTET is an experimental laboratory and will use an experiment-orientated set of measures; the included measures in ComTET are ones that can be measured experimentally and can be achieved within the project's scope, whereas Talisman would be concerned with all aspects of performance. The equivalence of Functional Purposes, but with differing Value & Priority Measures, indicates that that they are appropriately included and evaluated in a contextually appropriate manner.

Congruent Purpose-Related Functions as well as overlap of Physical Objects and Object-Related Purposes for both interfaces suggest high functional fidelity, despite differing HMI designs. These created different mechanisms for task completion in the work domain using the available Physical Objects and their affordances but did not preclude the completion of similar work. This would not be dissimilar to the variation that exists in control rooms across the fleet; while the exact specifics of how submariners work would be different, they would create the same mental models of their work, relying on their training to process all information available to them. This is especially pertinent for roles such as the OOW, where they assimilate information in their mind (Dominguez *et al.*, 2006), structuring their mental model independently of available information.

Usage of different interfaces invariably led to different Physical Objects, and subsequently different Object-Related Processes. Consequently, physical fidelity is reduced. However, an essential degree of fidelity is maintained by common interface aspects, such as Sonar waterfalls or TMA cuts. These interface commonalities are mainly core objects required to carry out Sonar and TMA, suggesting that a reduction of physical fidelity can be attributed to non-core Physical Objects. Thus, while not exhibiting high physical fidelity overall due to different interface compositions, a common core for essential components makes a case for medium physical fidelity in the ComTET HMIs. This assertion is supported by the physical facility itself, which is representative of a submarine control room, including the use of MFCs for operators (Roberts, Stanton and Fay, 2015). Combined, participants could “buy” that they were in a submarine control room, a key aspect of physical fidelity (Alexander *et al.*, 2005; Roberts *et al.*, 2020). This can be further justified by considering that variation exists within the RN submarine fleet. Submariners could use entirely different physical objects if conducting their work in multiple control rooms, but the nature of their work and belief that they are in a submarine control room would remain the same.

7.6 Conclusion

This chapter has detailed a validation exercise that was conducted ahead of experimentation using the ComTET versions of Sonar and TMA. It was performed to determine whether the ComTET Sonar and TMA environments could be considered similar enough to their real-world counterparts in Talisman to generalise results gathered using the ComTET simulator. A study was conducted at Talisman with five currently serving submariners over three days. They were interviewed and observed in-situ to construct counterpart abstraction hierarchies to those already created for Sonar (Figure 23) and TMA (Figure 34). Comparison between the abstraction hierarchies suggested that ComTET exhibits a high degree of functional fidelity with a medium degree of physical fidelity for Sonar and TMA. This establishes confidence that findings and recommendations from experiments using ComTET Sonar and TMA could have validity in a real-world setting as well. Predicating real-world applicability is vital, ensuring a tangible realisation pathway for outputs of the ComTET project and this work. The next chapter starts this process of experimentation, detailing how the redesigned Sonar and TMA interfaces were developed.

Chapter 8 **Creation of Graphically Integrated Sonar and Target Motion Analysis**

8.1 Introduction

With the CWA outputs from Chapter 7, in addition to the design directions generated from them for Sonar (Chapter 5, see Table 17) and TMA (Chapter 6, see Table 20), it was possible to design and implement the new EID HMIs. This chapter details how this was achieved, concentrating on exploring how to align EID design processes with those of software engineering to support the eventual end-goal of EID – implementation.

This focus was chosen to address a scarcity of literature addressing the design and implementation of EID as software artefacts, but also to contribute to the literature from which practitioners can use to inform their process(es). Although there is no shortage of EID research, the exact nature of creating an interface design from CWA has long been without a concrete process (Vicente, 2002; Read, Salmon and Lenné, 2012; Read, Salmon and Lenne, 2015), and designs are often made on a case-by-case basis (Upton and Doherty, 2008). There is literature that aims to address this with generalisable processes (Upton and Doherty, 2008), visual thesauruses (Hajdukiewicz and Burns, 2004), good practice techniques (Burns and Hajdukiewicz, 2004), and design toolkits (Read *et al.*, 2018). These contributions made strides in addressing the issue, though all acknowledged that design cannot be constrained into a series of rigid steps and will always involve creativity, especially when applied to multiple different domains. Furthermore, the approaches are not mutually exclusive, and practitioners may opt to combine them in accordance with their needs and preferences. Thus, it would be illogical to attempt to create a unified, all-encompassing, method; much like the complex systems represented by the interfaces, it would be perhaps impossible to account for all use-case eventualities, resulting in modification during application. Conversely, while design choices are made on a case-by-case basis, there are commonalities and best practices that must be observed, instead of individualised processes for each design project. Literature addressing the design process for EID recognises both extremes (one size fits all, completely individual design process), and collectively advocates for a 'toolbox' approach where processes are proposed that are readily adaptable to project-specifics. Therefore, it is believed that there is merit in contributing an approach oriented around integration with modern software practices for consideration, utilisation, and modification as required by practitioners.

Such an approach was required as most EID literature concentrates on proposing interface designs and evaluating them, which is understandable given that is a design methodology. However, most HMIs are formed of a frontend and a backend, both of which need to be designed and subsequently implemented. Prototypes and proofs of concepts have been made to test, but they remain scant on details of the underlying software design and implementation. While modern software maintains separation where possible between the two, their utility without each other is limited and their relationship should be considered. Given the time and effort involved in creating CWA outputs (Vicente, 2002; Stanton *et al.*, 2013), all possible uses for them should be fully explored. This is pertinent as the analysis on which many EID designs are based is a systems analysis (i.e., CWA), a key step in the Software Development Lifecycle (SDLC), albeit in a different format than usually employed by the software industry. Using CWA for software design has been explored by Wells *et al.* (2011), Dhukaram and Baber (2016), Oosthuizen and Pretorius (2018), and McIlroy and Stanton (2012) who drew parallels between CWA and standard software engineering modelling outputs. They chose to use these outputs to facilitate communication with the software engineering discipline. This is a key challenge to address as organisations will have heavy investments in software tools and ways of working, and approaches should be compatible with them to gain traction (Baxter and Sommerville, 2010). As with the visual design methods, the proposed software design methods recognised that creativity and refinement must still be present.

It is clear from existing literature that there is no one size fits all design process, but rather a collection of guides and processes that can be used as needed, including modification for specific purposes if required. Thus, this chapter details how this was achieved for the creation of an EID HMI for Sonar and TMA, proposing an adopted method for EID with the aim of producing artefacts suitable for implementation. Both the frontend and backend design are accounted for, extending, and adapting existing processes to create an appropriate design process that considers them as separate, yet related concepts. The approach was designed to be compatible with modern software engineering practices, minimising barriers to transferring the analyses between the human factors and software engineering disciplines. With the role of human factors practitioners established and growing within many organisations, it is vital that their work is manifested across the organisation, especially in the creation of software, a bedrock of a modern, connected world. Proposing a method that can be incorporated into software engineering team practices could maximise integration between the disciplines and exploitation of analyses results.

8.2 Creation of Initial Interface Designs

8.2.1 An Object-Oriented Design Process

CWA has strong support and evidence for its role as its analysis tool, although it does not specify a subsequent design process, save for its connections to EID (Read, Salmon and Lenne, 2015). Read, Salmon and Lenne (2015) conclude that practitioners must create their own approach to design, which was manifested as a CWA Design Toolkit (Read *et al.*, 2018). This addressed the need to provide practitioners with a process, and while assistive to be generalisable, it was not prescriptive. There is work that aims to be more prescriptive in addressing the gap between CWA and EID though, such as the work by Burns and Hajdukiewicz (2004), Hajdukiewicz and Burns (2004), and Jamieson (2003). This interconnected work proposes a core concept of a 'visual thesaurus' that maps variable types from the CWA to proposed forms. For example, a univariate variable within specified limits, such as speed, could be represented in the context of its minimum and maximum possible values using a meter (e.g., car speedometer). Once this transformation was applied to all identified variables, and the resulting visual forms organised according to means-ends links, a practitioner would have a reasonable first design for an EID interface, which could be iterated. It was decided to use the approach of Burns and Hajdukiewicz (2004) to design the interface, over more recent design methods such as the toolkit proposed by Read, Salmon and Lenne (2015). This was because the ComTET project provided regular access to subject matter experts) over time, where they could be consulted on the designs and provide feedback, such as at quarterly progress meetings. Therefore, it was decided to utilise longitudinal input from subject matter experts to inform decision decisions, as opposed to the specific workshop approach proposed by Read, Salmon and Lenne (2015). However, other aspects of the design process, generated insights for Sonar (Chapter 5) and TMA (Chapter 6) in line with the CWA Design Toolkit.

The process of Burns and Hajdukiewicz (2004) is summarised as follows:

1. **Generate a list of variables:** Generate a list of variables present in the abstraction hierarchy, with each level providing a different type.
2. **Convert the variables to visual forms:** Each of the variables identified should be transformed into an appropriate visual form, utilising the visual thesaurus if useful. These forms are arbitrarily arranged at this point.
3. **Generate a constraint list:** Generate a list of constraints from the abstraction hierarchy. These can either be univariate (formed of one variable) or multivariate (formed of multiple variables). There should be two pass-throughs of the hierarchy:

   a. The first passthrough is to identify univariate constraints that would enrich the basic variables identified, such as providing their limits.
   b. The second passthrough is to identify multivariate constraints, which are formed from the relationships between variables. This also includes variables that need to be considered across time.
4. **Edit the visual forms to account for the constraints:** The visual forms from step 2 should be edited to include the identified constraints from step 3. Again, the visual thesaurus could be used to determine appropriate forms to use.
5. **Organise by means-end links:** The visual forms should now be arranged, taking cues from the means-end links in the abstraction hierarchy. Users should be guided to monitor the Functional Purposes of the domain.

The method proposed by Burns and Hajdukiewicz (2004) does not claim to be all-encompassing, and notes that flexibility is key for generalisability. Consequently, the process was modified in preparation for being aligned with the underlying software development activities.

The first modification was to adopt an Object-Oriented (OO) approach, using the classes present in the work domain as the starting point for designs. Classes are abstract blueprints for specific objects, which represent the class being made concrete and having information added to it. For example, ‘boat’ could be a class, which would store information such as the type of boat and identifier. A boat class would be made into an object for each boat present in the area surrounding the submarine, each containing the information of the entity they represent. This approach was chosen because the work domain was formed of a variable number of entities, with was subject to change during usage (e.g., new contacts being represented on a map as they are detected). Thus, an archetypical design would have to be created, which could be represented on the display as many times as required. This builds on the work of Dhukaram and Baber (2016) who explored extracting classes and variables from the abstraction hierarchy to design code for an EID. It extends their approach to also consider visual objects, and to look across all levels of the abstraction hierarchy, as opposed to just Object-Related Processes. The latter is important, as classes can be abstract and exist at different conceptual levels, so a holistic approach is more apt over only considering specific levels. The approach was also chosen in anticipation of linking with the design of the underlying software, which would use an object-oriented approach, a standard in modern software engineering. For an initial prototype, templates could be created that could be added to the main display, or the main display could be created directly, and the different objects manually replicated (e.g., copy and paste). Examples of entities included ownship, other vessels, land, and the ocean. As these objects were the building blocks of work domain being examined, it was natural to use them as such for the design. Furthermore, this created synergy with the approach for the backend, which also used an OO approach, and will be described in the next section.

The method that Burns and Hajdukiewicz (2004) proposed was variable oriented. This was reconciled by assigning identified variables to at least one object. Doing so added more nuance to each variable, as it could be considered differently in the context of its parent object, similar to the 'object worlds' concept in WDA, where different physical objects would have different connections based on their context. For example, part of ensuring ownship safety is to maintain a depth such that the submarine is not damaged by the sea floor or other environmental factors. Safe depth is the variable in this example and means different things when assigned to different objects; for the submarine, safe depth is the maximum depth it can dive to, whereas, for the water, safe depth is the maximum depth of the water take the submarines draft and a safety margin.

The second modification was to directly account for the uncertainty faced by the submarine command team when compiling the tactical picture. This is important for EID, as sensor uncertainty should not be perceived to represent the actual system state (Vicente, 2002; Burns and Hajdukiewicz, 2004), especially in the submarine domain, where there is a high level of uncertainty (Brolese, 2005; Dry *et al.*, 2005; Hunter, Hazen and Randall, 2014; Kirschenbaum *et al.*, 2014). Consider ranging a contact that has been detected on Sonar on a bearing of 45°. The direction is known, but the range could be anywhere between 0 meters and the range of the sensor (excluding sound propagation physics for simplicity). At this point, there is a constrained variable of range that could be represented by a spatial icon; if the operator enters a range for the contact, then its icon could be moved to show this. However, the icon must make the associated uncertainty clear, such as by changing its image to show this. If the information is presented as certain, then a decision might be made to steer the boat towards a supposedly safe area, creating a dangerous situation.

Furthermore, as the range could be wrong, a matter independent of operator confidence, accompanying representation is required to aid the operator in assessing this. Part of this representation is that which was used to derive the range in the first place, such as the cuts on a Local Operations Plot (LOP; see Section 4.1.1.5 and Section 6.2.4.3). Another part becomes possible through the operator entering the range, enabling representations to be based on this. For example, when the icon dead-reckons, it could change colour if it no longer aligns with the latest cut, which would aid with detecting faults with information entered by the command team. EID has been shown to improve fault detection and diagnosis (Borst *et al.*, 2017), and consequently, it is important to capitalise on this to ensure that operators are fully aware of uncertainty in the system by explicitly designing for it.

If a variable was known to be uncertain, then it received appropriate counterpart confidence and verification variables. This approach is aligned with the definitions of constraints found in section 3.2.3, where they were found to be relational and/or limit-based. It is possible that these variables could have been implicitly generated from other nodes in the WDA as in the original approach, although the explicit synthesis approach ensures their creation and subsequent inclusion in the resultant interface. Additionally, it forces consideration within the context of the parent object, which is important if the variable is being applied to multiple objects as described in the previous paragraph, as the counterpart variables might be different.

The result of these modifications to the interface design process is visualised in Figure 49. While the process is based on that of Burns and Hajdukiewicz (2004), it has been designed to serve as a generalised process capable of incorporating different process outputs present across the literature for the EID process, having a start point for each. The first is a *"Work Domain Analysis"*, which is used to inform *"Frontend Design"* through the work of Burns and Hajdukiewicz (2004) or/and (Read *et al.*, 2018 - the CWA Design Toolkit), but can also be interchanged with *"Classes for Code"* and other software engineering modelling outputs (Wells *et al.*, 2011; Dhukaram and Baber, 2016; Oosthuizen and Pretorius, 2018), as will be discussed specifically in Section 8.3. All starting points connect with *"Integration with software engineering method(s)"*, which represents an entry point to other software engineering processes that might be used, including the agile process in Section 8.3, Figure 52. To use the process, a practitioner would start with their current source material (*"Work Domain Analysis"*, *"Classes for Code"*, or an *"Interface Design"*) and proceed through the process to their target output. Assuming starting from *"Work Domain Analysis"*, the steps can be described as follows:

1. **Identify classes:** Generate a list of all objects present. In this context, an object is not strictly a physical object, but also abstract concepts and categories that could be a suitable name for a grouping of variables. For example, if a Value and Priority Measure was to 'Maximise Fault Detection', an object could be a 'Fault' as it would have information associated with it, such as the timestamp of its discovery, information on what it affects, and severity. These classes, along with associated data from steps 2 – 3.1 can be used to information the software engineering process in Section 8.3.
2. **Identify all variables and their constraints:** Generate a list of variables and constraints present. Variables can be defined as a name for a piece of data. This can follow the guidance of Burns and Hajdukiewicz (2004), or another method for systematic elicitation of all variables and associated constraints. Some variables might act as constraints for other variables, so they could be considered in tandem throughout the process.

3. **Assign variables to classes:** Each variable should be assigned to one or more classes as appropriate.
    3.1. **Consider uncertainty:** When a variable is assigned to a class, it should be evaluated whether it would be subject to uncertainty in the context of the class. If so, counterpart variables should be created that will be used to display the confidence in the information, and how the information can be verified. It is possible that other variables belonging to the class might address this information and could be used in lieu of the counterpart variables. This process should also be followed for the constraints affecting the variable.
4. **Create or modify visual forms:** Each class should now be visually designed. A canvas should be created for each identified class (or they could be added to a single diagram using an appropriate separation mechanism such as layers or grouping). The canvas should be populated with appropriate representations of the identified variables and their constraints.
5. **Identify and design the main display:** The visual forms from the previous step will provide the ecological representation for the interface but may not include the general look and feel. Examples of this would include the menu bar, status bar, and other mechanisms to arrange visual forms. A general frame of the interface should be designed, serving as the main canvas to arrange the class visual forms on. An alternative is to use a visual form as the base, such as a map view onto which everything else can be overlaid.
6. **Organise visual forms:** The visual forms should now be arranged, taking cues from the means-end links in the abstraction hierarchy. Users should be guided to monitor the Functional Purposes of the domain. Interaction and storyboarding can be considered at this stage, to map how the user would move through the application.

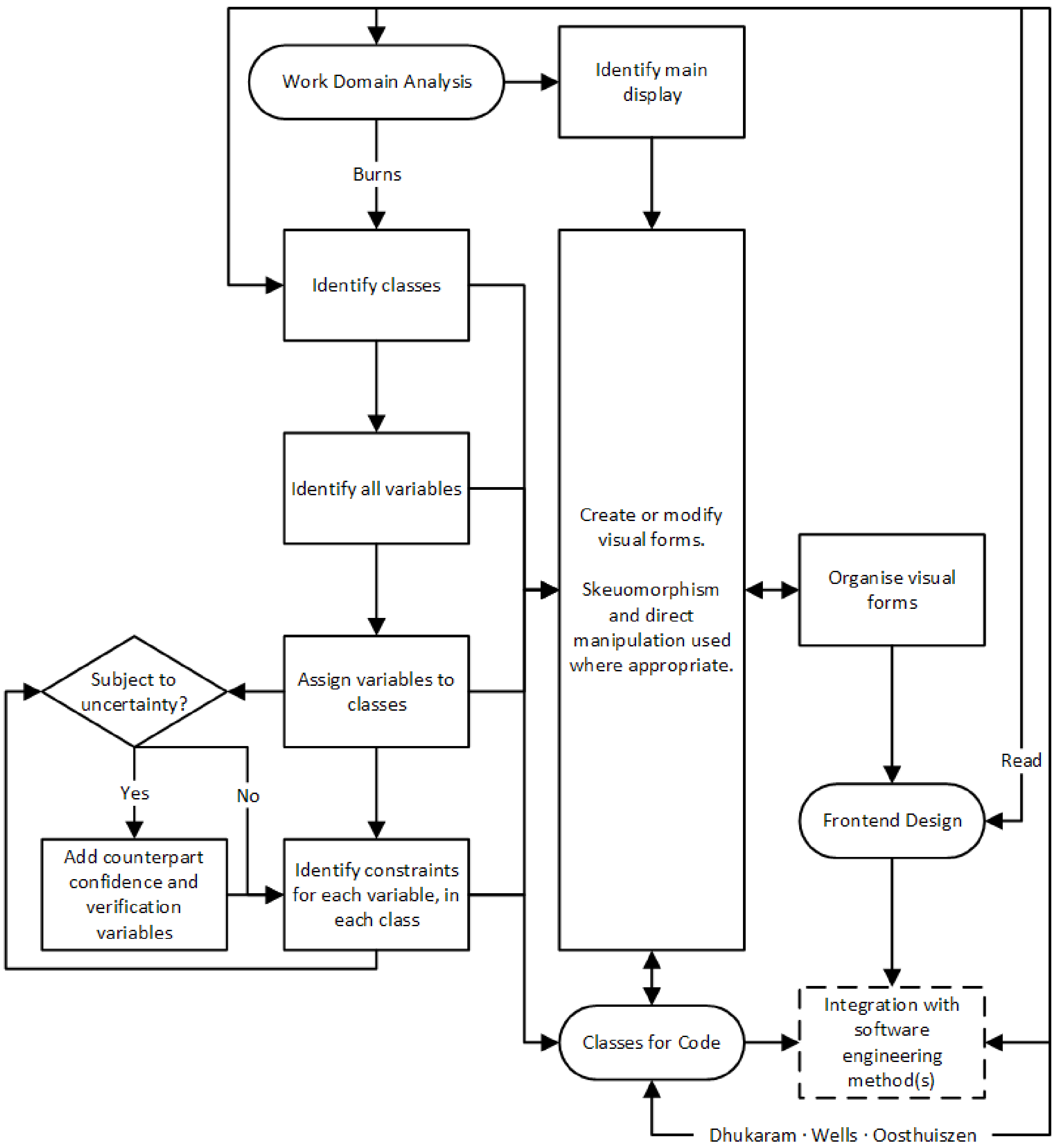


Figure 49 – A modified version of the EID process of Burns and Hajdukiewicz (2004), utilising an object-oriented approach and incorporating pathing to different EID outputs.

8.2.2 Resultant Initial Designs

A proposed interface for each role was created by following this modified process. The designs were made with the design directions from Chapter 5 and Chapter 6. Input from domain experts throughout the process was also sought at project review meetings, during the CWA comparison process in Chapter 7, and during the design workshop described in Section 10.2.1. The main feedback received was that the waterfall should be re-included as this was vital to Sonar operation. This feedback was taken onboard, and later iterations of GIST incorporated a waterfall in the Sonar tab of the information panel (see Figure 59). Not all data could be represented on the map.

Therefore, three common areas were added, which are the grey areas in each interface. The bottom area shows ownship parameters. While some information here could be represented on the map, it was relocated to minimise the occlusion of map data. The left area shows a list of all contacts that are currently held, represented as multi-coloured buttons. It also shows action panels for a contact, the content right of the white vertical line, allowing operators to carry out necessary analysis tasks on each contact. Additionally, appropriate interaction mechanisms were added, such as buttons to trigger actions, and textboxes for operators to enter information, as not all data would come from sensors. A single canvas approach was chosen for simplicity, with sample representations for each object identified added directly to the canvas and arranged as required. The resultant design for Sonar is shown in Figure 50 and Figure 51 shows TMA. The designs are separate images but were created with one interface in mind. This interface was named Graphically Integrated Sonar and TMA (GIST). A full description of GIST will be provided on the finished design (Section 8.5), as there was considerable change between these initial prototypes and the finished product. The designs were validated through review at quarterly project meetings, and contact with submariners, such as during the validation exercise in Chapter 7.

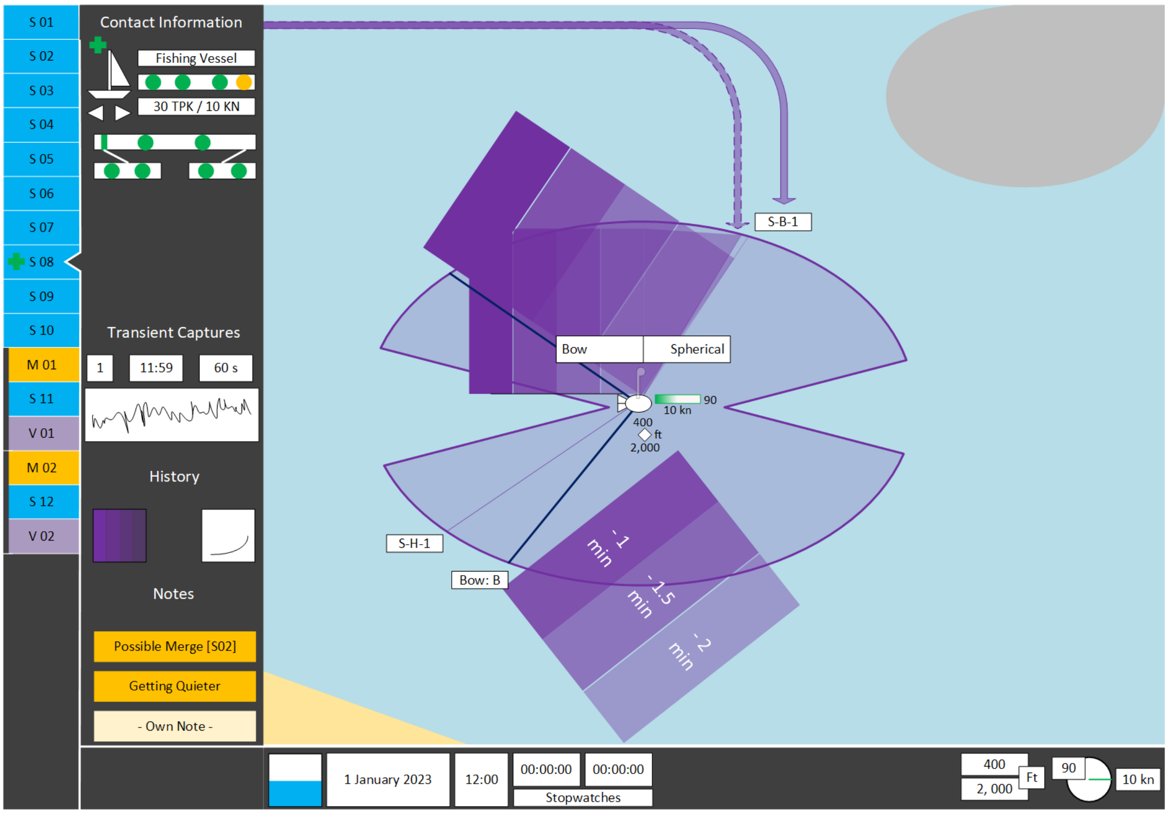


Figure 50 – Initial GIST Sonar Design

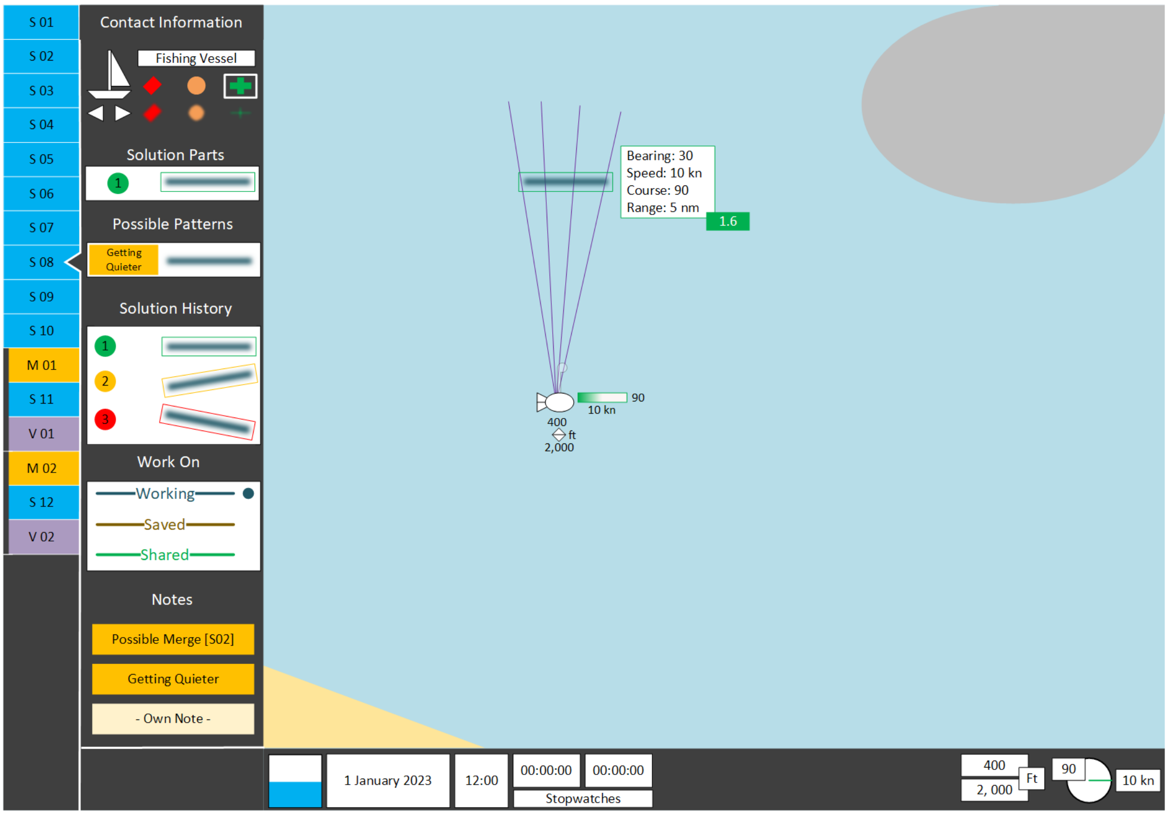


Figure 51 – Initial GIST TMA Design

8.3 Software Engineering Process

With the frontend design created, focus moved to planning the underlying code as the backend using the abstraction hierarchy. There is research on aligning CWA outputs with software engineering systems analysis methods. Wells *et al.* (2011) worked with Systems Modelling Language (SysML), creating a tutorial that aligned CWA terminology with SysML, and subsequently creating CWA equivalents in SysML using this. Dhukaram and Baber (2016) follow a similar approach, drawing parallels with Unified Modelling Language (UML), and proposing a translation process between it and CWA. Oosthuizen and Pretorius (2018) also used SysML but drew additional parallels with System Dynamics, which they defined as a technique to support modelling system behaviour at high levels of abstraction, citing from Sterman (2000).

These represent useful contributions to the literature, aimed at bridging the gap between human factors and software engineering systems analysis methods. There is a focus, and particularly so for the first two contributions, on translating CWA outputs to software engineering outputs for consumption by software engineers. The reasons for this are oriented around communication and suitability for use. Wells *et al.* (2011) suggest that cognitive engineers have difficulty communicating their findings in a format suitable for systems engineers, and that the former often do not have access to CWA information or training in a format that they would understand. Dhukaram and Baber (2016) also identify communication between human factors practitioners and software engineers, citing Bruseberg (2008), and a disparity between the aim of CWA in describing how a system could function and that of software engineering designs which describe how it must function. While this might be true in some circumstances, the assertion could be challenged where the software design is deliberately flexible, such as a code library or micro-service, which can be joined with other software components to create flexible behaviour.

The communication difficulties are acknowledged by Baxter and Sommerville (2010), who believed that it is not sufficient to explain the results of a sociotechnical systems analysis to engineers, rather that it must be suggested how to apply them to the system design. They also note that companies can be heavily invested in specific tools and methodologies, so sociotechnical approaches should be compatible with these to be successful. This sentiment is maintained at the individual level by Viller and Sommerville (2000), who explored the use of UML to communicate requirements to avoid software engineers having to learn an additional method.

The focus on translating to modelling languages for consumption by software engineers is appropriate as they are a key aspect of software engineering. However, there is a requirement for the CWA outputs to be translated into the modelling language, and the artefacts iterated, likely by a software engineer or architect. This poses three potential issues. The first is that the translation process is effectively replicating the CWA outputs in a different format. The second is that the richness of the analysis could be lost by specifying just requirements (Baxter and Sommerville, 2010; Dhukaram and Baber, 2016). The third is that a modelling language output might not be required, and software engineers might prefer to use other approaches for creating the software. For example, if the analysis has been completed on a current system, all changes could be made within the existing code base based on their knowledge of its operation, meaning that the design work was redundant.

This poses an additional related issue as CWA is scarcely applied to first-of-a-kind systems (Salmon *et al.*, 2010). An existing system would have established code, which must be accounted for to make a meaningful design. Unless the practitioner has access to and can understand the existing code, designs based on translating the CWA outputs might not be compatible with the implementation. This issue could partially be offset by the contemporary practice of developing software in small, self-contained, components. If the resultant design from the CWA outputs could replace the entire component, then less familiarisation would be required, but integration with other components would still have to be considered.

Combined, these issues suggest that it might be more appropriate for the software engineers to be provided with the CWA outputs and recommendations that were made as requirements. Again, there is literature that explores this. McIlroy and Stanton (2012) proposed that outputs from CWA could be used to inform software requirements. However, the resultant requirements are the level of the entire system and what it should do. These requirements would be taken by a software team and transformed into a series of smaller requirements, detailing how the software will be created. Consequently, there is an opportunity to further develop this step in the software development lifecycle.

Provision of the CWA outputs and associated recommendations would avoid duplicated effort, maintain the analysis richness, and allow software engineers to plan their work as desired. Doing so would require them to learn a different method, something that Viller and Sommerville (2000) advocated for avoiding, although as Wells *et al.* (2011) demonstrated, there are parallels which could aid this. If this is achieved successfully, then it could address the organisational concerns of Baxter and Sommerville (2010), with software engineers utilising the CWA in their software processes as they see fit; this is a distinct possibility with the rise of agile methods that allow teams to structure their work locally as required and decide on their processes as well as tools (Abrahamsson, 2002). Thus, an agile approach based on the CWA outputs and recommendations from Chapter 5 (Sonar) and Chapter 6 (TMA) was trialled for the creation of GIST. Agile was chosen over a traditional waterfall approach, where all design work is competed when a project starts, as it has been shown to improve the software engineering process and its outcomes. This is achieved by maintaining flexibility to respond to change and providing deliverables quicker.

The process was based on Kanban, which is an agile method where tasks are moved through different development stages, represented as columns on a board (Radigan, 2019). The board can be as simple or complex as required; this work used an Excel spreadsheet with a dropdown for the stage for ease of interoperability between parties involved in the development. The stages could be as basic as “To Do”, “In Progress”, and “Done”, with their exact nature depending on the team. All tasks to be completed are collected as a backlog. As the backlog is dynamic, agile methods can be particularly accommodating of iterative processes, as CWA and EID often are. Individual tasks from this backlog are moved across each column as capability becomes available, with a continuous cadence of tasks (i.e., work is not boxed into fixed-length sprints as with other methods, such as Scrum). The result of this process is that functionality is continuously added, and the software being created is rapidly updated with new functionality as the user stories are completed.

Cards are a visualization of these tasks and can contain a variety of content types (e.g., text, images, sound, videos; Rehkopf, 2019b). The cards can be specific requirements but can also be expressed as user stories, which often take the form of “as a [persona], I [want to], [so that]” and represent an end-goal to be implemented (Rehkopf, 2018). This has parallels with the abstraction hierarchy, and user stories could be described as “as a [sociotechnical system agent], I [what node], [connected why node]”. The format is not fixed and could be expressed as desired, such as describing it as a task to be completed (e.g., “Reduce overall fuel consumption” (Rehkopf, 2019a)). When expressed like this, the stories could come from the middle levels of the abstraction hierarchy, such as describing what a Physical Object should do from the Object-Related Processes, or how success is measured using the Value and Priority Measures. User stories could also be generated from a Worker Competencies Analysis, by creating stories for functionality that would be required by operators at each level of the Skills, Rules, and Knowledge Taxonomy. Using user stories is congruent with the CWA Design Toolkit (Read *et al.*, 2018), which proposed user stories as part of the “Communicating Findings” stage of “Concept Design”. In this application, the user stories were used to guide iterative improvements to the concept design but were also used as direct inputs to the software engineering process.

The user stories could be grouped into epics, which are overarching bodies of work that can be broken down into stories (Rehkopf, 2019a). They describe large bodies of work towards a specific goal. An example epic could be “Broadband Functionality”, as this could be broken down into several stories that detail exactly how it will be achieved. An alternative could be to use the Functional Purposes of the abstraction hierarchy to base epics on, as they describe a system’s purpose for existing. Depending on the size of the desired interface, epics might not be required, but any sizeable development should be grouped into epics to organise work into high-level threads.

As the user stories are being enacted, software engineers should make use of outputs from Section 8.2.1, which are lists of objects and their properties. They could be used to describe the data structures, constraints, and logic in code, known as the model. Described in CWA terms, a model would be a representation of the work domain. The model could be supplemented using the processes of Dhukaram and Baber (2016), Wells *et al.* (2011), or Oosthuizen and Pretorius (2018) to generate full blueprints for code that should be implemented to connect the model to the interface view. In software, this type of code is known as a controller, which is logic that takes user input from the interface view and uses it to manipulate the model. In turn, the interface will then reflect the information from the updated model.

Figure 52 provides a visualisation of the full process. The process can link to the frontend process in Figure 49 by incorporating the *“Classes for Code”* (or full software engineering modelling outputs using the work of Dhukaram and Baber (2016) or Oosthuizen and Pretorius (2018)) and/or *"Frontend Design"*. A example process would follow these steps:

1. **Generate user stories:** A series of user stories should be extracted from the CWA outputs present, which are commonly Work Domain Analysis and Worker Competencies Analysis for Ecological Interface Design (Burns and Hajdukiewicz, 2004; Jenkins *et al.*, 2009). A widespread template is "as a [persona], I [want to], [so that]", although this can be changed depending on requirements. User stories can also be created from identified design directions, such as those generated in Chapter 5 and Chapter 6.
    a. **Generate epics (optional):** If there is a large amount of work, then epics should be created for organisation. These can be based on the creation of key aspects of the software, or on Functional Purposes from the abstraction hierarchy.
2. **Create the Kanban backlog:** The user stories should be placed into the "To Do" column of the Kanban board. They should be prioritised as required as part of this process.
3. **Start the development process:**
    a. **Work through backlog:** Software engineers, user experience experts, and other practitioners can now choose a user story from the "To Do" column and move it to the "In Progress" column. This signifies that they are working on implementing the functionality required. They should use the outputs from the *"Frontend Design"* (e.g., images, storyboards, and software engineering modelling outputs, including *"Classes for Code"*) where possible to capitalise on prior work. Once the user story has been completed, it should be moved to the "Done" column.
    b. **Iterate design and functionality:** Once functionality has been implemented it can be tested to determine if it provides optimal functionality and usability. If it does not, then either the design should be iterated, or more user stories should be generated to make improvements.
4. **Use of finished product:** After a predetermined amount of work has been completed, such as the completion of a key epic, the software will be ready to use for its intended purpose. While agile relies on continuously improving functionality, meaning that the software could be used at any point, using a milestone ensures that all capability for the intended use, such as human in the loop evaluation, is present.

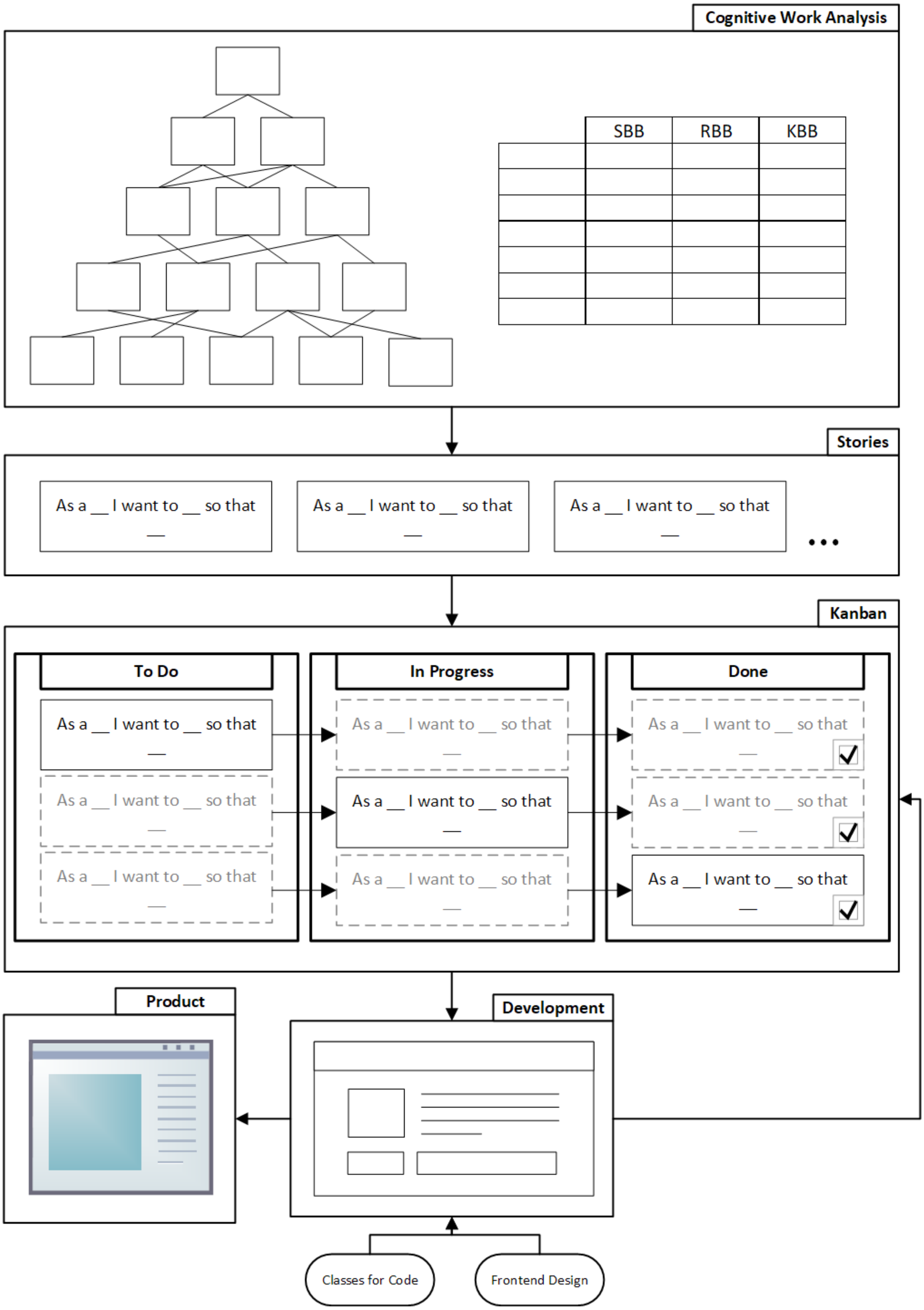


Figure 52 – Visual representation of using the outputs of Work Domain Analysis and Worker Competencies Analysis to use as inputs for Kanban-based agile software engineering.

8.4 Software Development

GIST was created using Simulation Engine II, a simulation engine created by Sonalysts designed to support multi-domain operations at multiple levels (Sonalysts, 2022) using the provided Software Development Kit (SDK). The frontend interface was designed using Qt in Qt Designer and the backend was implemented using C++ in Microsoft Visual Studio (various versions as updates became available). There were two software development streams. The first was undertaken as part of this thesis, based on processes from the previous sections. A wireframe of the initial design from Figure 50 and Figure 51 was created in Qt. After this, user stories were worked through to add functionality and fill in visual placeholders (e.g., replacing a blank box with an actual waterfall implementation) in an iterative manner. The abstraction hierarchies were systematically worked through, identifying user stories that would be pertinent to implement. For example, "as a *'TMA operator'*, I want to *'Gauge Solution Accuracy'*, so I can *'Assess Solutions'*". Given the dynamic nature of the interface, it was felt that it would be more appropriate to use a live prototype to understand what worked well and what could be improved, instead of creating all user stories from the start using a storyboard approach. The second stream was completed by Sonalysts, who provided support on the SDK, ensured all required capability was present, and assisted with developing parts of GIST.

Initial development of the designs was completed during a three-month visit to Sonalysts. They provided support on using SEII and the SDK. After the three-month development period, the foundations were present and adequate training had been provided to enable continuing development efforts. Using this training, GIST was progressed to a proof-of-concept with support from Sonalysts, who provided additional builds of the simulation engine to enhance functionality where required.

As with the frontend design, an object-oriented design and object-oriented programming approach were mainly used for the underlying code. While most code was within a class, exceptions included helper functions that were not placed in a class as a design choice, instead being available within the code's namespace, compile-time constants, or compiler instructions. These choices were either required by the compiler or were the result of design decisions made during development. Encompassing the design and creation of object-oriented software respectively, these approaches view software as being formed of multiple interacting classes representing entities, which encapsulate properties and methods of that entity. An instantiated class, which can be likened to a blueprint, is called an object. The methods were chosen as standard methods of software engineering (Bourque and Fairley, 2014), and by adopting an object-oriented approach to the design, they could be initially informed by that process. For example, it was known that other vessels were identified as an object to visually prototype and a series of properties were used to add visual elements; this same list could be used as the basis for the underlying class. The classes were refined as required during development, aided by the tools provided by Visual Studio and ReSharper (various versions; a suite of productivity tools).

8.5 Finished HMI and Key Features

A screenshot of the finished HMI is shown in Figure 54, showing the general layout. The similarities to the initial designs from Figure 50 and Figure 51 are visible, although it is clear that changes have been made. This section will discuss pertinent aspects of operating GIST and relate these back to the design directions identified in Chapter 5 and Chapter 6. Pertinent throughout is ensuring that the interface is as configurable as possible, arising from the design direction identified in Section 5.5.3. Where possible and appropriate, GIST was designed to give operators control over how they complete tasks and how data is represented to them.

The general layout is the same as the initial designs. The Ownship Information Panel is at the bottom, see Figure 53. From left to right, it contains the following capabilities:

- **Settings:** Toggles Sonar and TMA capabilities. Also allows for toggling sound.
- **Messages:** Shows messages to the operator that they need to acknowledge. These messages can either be from socio (operators) or technical (GIST automation) agents.
- **Cursor [Hull] and Cursor [Bow]:** Shows the true and relative bearings, and the range, of the cursor on the map in relation to the sensor the readout is for. The readout changes colour between green and red to match the associated relative bearing colour (starboard and port respectively), a mechanism which is applied to all bearing readouts throughout GIST.
- **Sensors:** A visual control for selecting which sonar sensor to display on the map view.
- **Stopwatches:** Two digital stopwatches that also include pause capability.
- **Ownship Parameters:** Readouts that show the depth of the submarine and associated keel depth, the course, and speed.

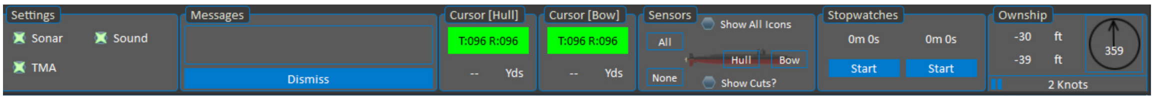


Figure 53 – GIST Ownship Information Panel screenshot

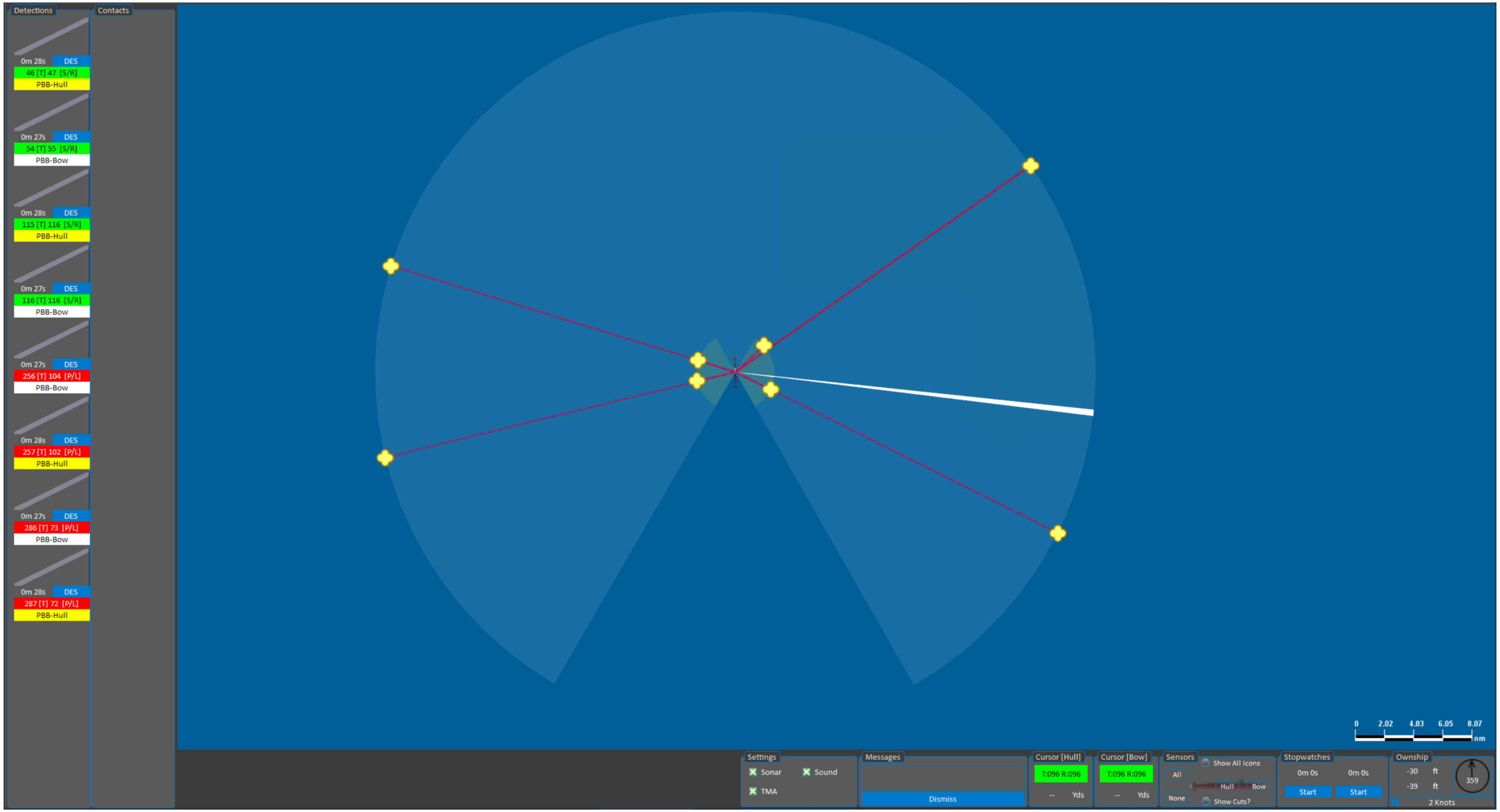


Figure 54 – GIST Screenshot

The main part of the interface is a top-down map view, occupying most of Figure 54. Ownship is shown as a representative image, see Figure 55, whose size is contingent on the zoom level. There are minimum and maximum sizes set on the image so that the submarine is always locatable. As this might be different to the actual dimensions, these are shown as a faint grey rectangle, allowing the operator to understand the submarine's actual dimensions at the zoom level. Characters are shown on the ownship image to represent information, such as 'S' for surfaced/periscope depth, and '!' for possible counter-detection (based on fictional, arbitrarily chosen figures in GIST). Other entities were shown as icons. Any contact without a solution would also have a warning label underneath its icon, see Figure 57.

Sonar sensor data is represented ecologically on the map, implementing the first recommendation from Section 5.5.2. Active sensors are represented by a circular polygon extending from ownship, showing their typical range. Sensor baffles are represented by removing a corresponding sector from the circle. Each sensor is assigned a colour, which was used on the map view and throughout the interface. The opacity was proportional to the background noise of the sensor, where more noise causes increased opacity. This allows operators to visualise the noise, and to intuitively understand that the sensor might not be detecting all traces as it will be occluding the environment underneath the opaque polygon. Detections and contacts were represented by icons and would initially appear at the edge of the sensor's range along the bearing line, but this is configurable, such as the use of pre-defined initial ranges. Live sonar sensor data was represented by a red line showing the most recent bearing data. Each sensor had a corresponding white sweep polygon, representing a 1° sector. It followed the mouse around for its angle and allowed the operator to play aural noise from that bearing.

Detections and contact cut information were represented as overlays on the map as well, see Figure 55. Current sonar detections were marked as red lines. As with the sensor range polygons, each detection line's opacity was a function of signal strength, which also affected the line width. They represented all signals being received. Once a detection had been made into a contact, cuts would start to be added to the map. This occurred at regular intervals, and as with the Dangerous Waters (DW) LOP were marked as cut lines. Their thickness and opacity were linked to the signal strength at the time of creation. There were multiple cut types and each on received a different colour: the leading (most recent) cut was coloured magenta, all cuts before this were coloured blue, and cuts that were redrawn were orange. Redrawn cuts were implemented to solve a graphics issue, where some cuts would not render when required; the operator could request the cuts be redrawn at any time, which would re-render all cuts using an orange colour.

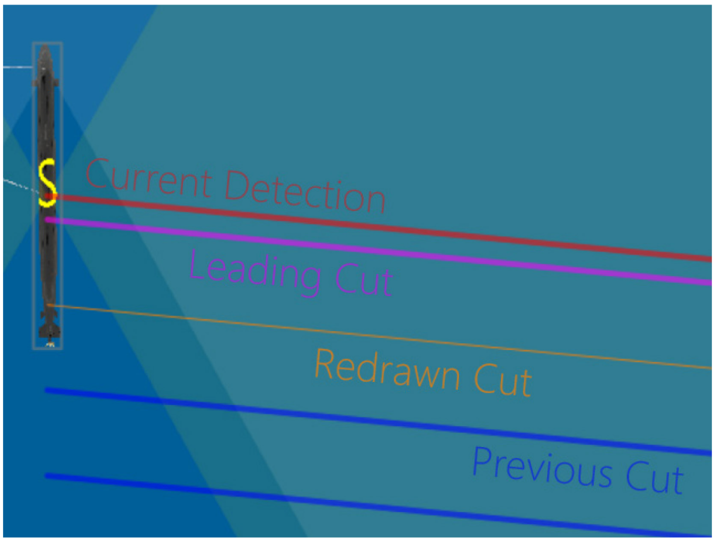


Figure 55 – Cuts for a contact and leading sonar data

Figure 56 – Example icon with a warning that a solution is required

On the left of the interface is Contact Management, see Figure 57. This was formed of two lists, one for detections, and one for contacts. These lists were populated by widgets (a generic term for an interface item) for each entity that displayed an image, the time it had been held, the bearing data, and the originating sensor. The widgets were sorted at regular intervals by their bearing, allowing operators to easily identify which contacts are merge candidates based on having similar bearings with different sensors. The blue “DES” (short for designate) button on the detection widgets was used to assign a tracker, causing it to be moved to the contacts list. The widget would remain the same apart from two small changes. Firstly, the button text would change to the tracker ID, which would open the Contact Information Panel when clicked. Secondly, the image area was populated with the default classification image, which is from the first possible entity in the database.

The Contact Information Panel had three tabs. The first was the General tab, see Figure 58, which contained general functionality for managing the contact. The operator could set the contact’s classification details using either the mouse or keyboard. If a contact was merged, will assign a colour to the merge, and highlight all trackers in the merge by colouring the left edge of the contact card. This data is also reflected underneath the “Split Into” button, which details what the merge is formed of. This was done to maintain operator SA, allowing them to understand which trackers form the contact they are working with.

The second tab was Sonar, see Figure 59, which is where Sonar analysis tasks were completed. The narrowband analysis was compacted into one readout, which used colours to represent the different datasets. Red lines for each frequency in the database entry would be shown as the operator cycled through the available classifications. Any frequencies from the contact that did not match were rendered in blue, and those that did were rendered in green. Operators could use this information to see if there was a match, and how complete partial matches were. Two waterfalls were implemented, which only showed contact data for the currently selected contact. This was to make the sonar data clear to operators. The bottom waterfall was a standard historic waterfall, while the top waterfall was a predictive waterfall that would use the most recent data to show what the waterfall would look like in the future using linear extrapolation. Both waterfalls had corresponding readouts underneath that would display information about the point underneath the cursor.

The third tab was TMA, see Figure 60, which is where TMA tasks were completed. There were three solution types with corresponding colours, shared (purple), working (blue), and private (yellow). The colour was used for all widgets pertaining to the solution, including items in the information panel and on the map. For example, the solution being drawn in Figure 61 is a working solution as it is blue. Solution parameters were shown in readouts within each solution area. The active solution, which is the one that would be drawn when interacting with the map, could be selected by clicking on the corresponding button in the “Solutions” group. Solution parameters could be copied between each solution from here by clicking the white arrows.

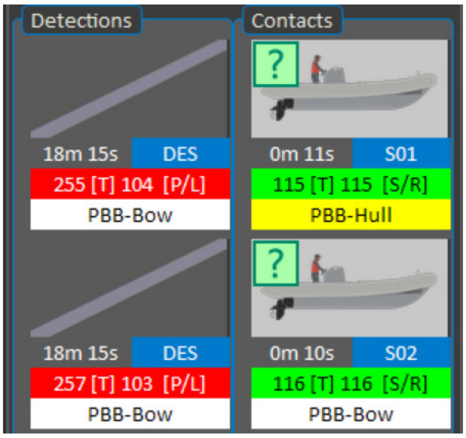


Figure 57 – GIST Contact Management screenshot

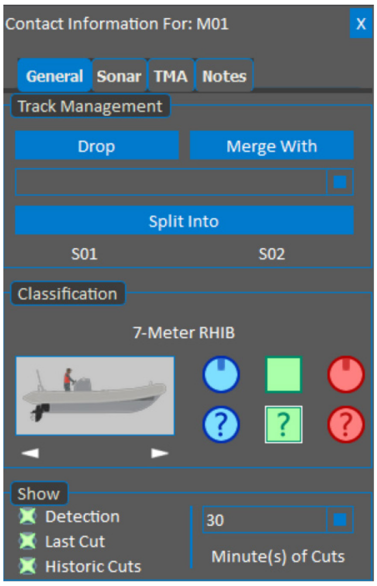


Figure 58 – GIST Contact Panel – General Tab screenshot

The mechanism for creating solutions was vastly different from DW and implemented the design direction to add more automation and enforce solution parameters, arising from the design direction identified in Section 6.4.2. Solutions were drawn directly onto the map as a series of straight legs, bookended by icons, by clicking on points, see Figure 61. This allowed operators to enter a historic path for a contact as well as its most recent leg, as opposed to just the single leg option that DW provided. Operators could directly manipulate the speedstrip using its icons, which could be added, moved, and deleted as required. They could also draw an entirely new solution.

The key difference from DW, however, was that the speedstrip would calculate the intersections to use for the solution automatically. This was achieved by taking the two most recent cuts that were intersected and marking the intersections, also shown in Figure 61. These two points were used to compute the solution components. The interface validated the intersections to make sure that the final cut had been intersected, and that it was the last intersection (by time), warning operators if these conditions had not been met. Once a solution had been entered, it had a predictive component that would show where the contact would be a short distance into the future. This predictive line turned red if the contact would violate ownship safety constraints. Restricting solution creation to use known data ensured that operators could not enter solutions that did not reflect cut data, and using the latest cuts ensured that out-of-date data was not used accidentally. Operators could filter the amount of data shown by using the 'Show' options in the General Tab.

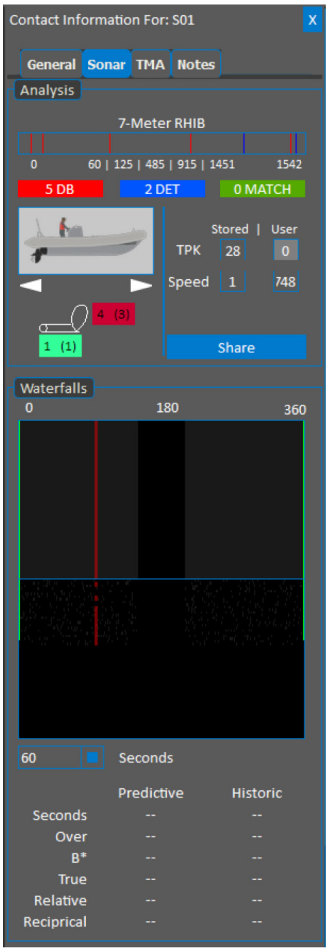


Figure 59 – GIST Contact Panel – Sonar Tab screenshot

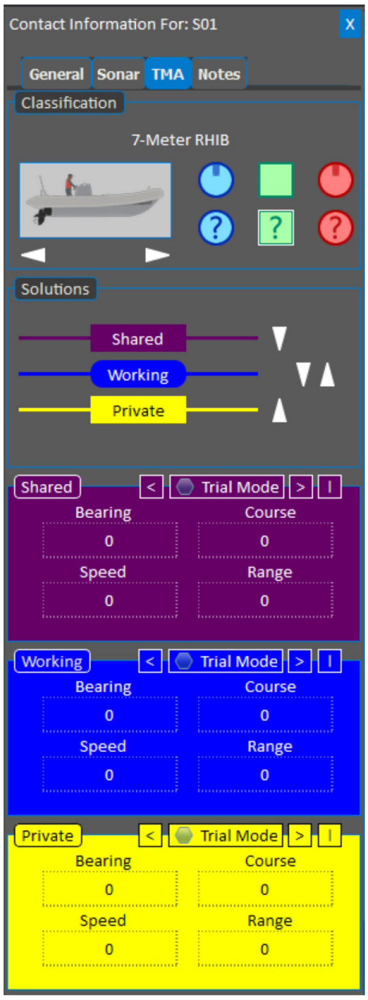


Figure 60 – GIST Contact Panel – TMA Tab screenshot

Operators were also provided with trialling tools to adjust their solutions. These tools were designed to assist operators in narrowing down solution possibilities, in addition to the other techniques, such as pattern recognition and range gating. Operators could use the optional tools to refine their solutions, or understand how a solution was arrived at, helping to develop their skills and understanding of the process, as per the design direction identified for Sonar and TMA. The first tool created a circular polygon around the mouse that showed the distance it would be possible for a contact to cover at the entered speed, see Figure 62. Operators could move their mouse along a cut and see whether it was possible to reach the next cut and by which course by looking for intersects, which would indicate the course(s) a contact would have to be travelling at the proposed speed to move between the two cuts being inspected.

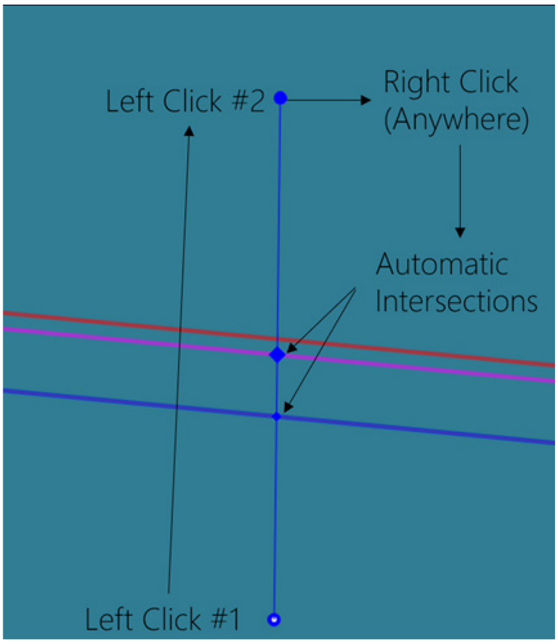


Figure 61 – Entering solutions

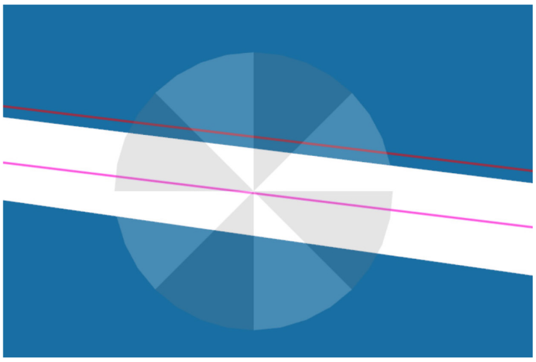


Figure 62 – Speed and course circle

The second tool placed the solution into a trial mode, see the controls atop each solution box in Figure 60, and allowed solution parameters to be entered numerically into the TMA readouts. These values could be previewed by pressing the “>” button, which would represent the solution on the map, see Figure 63 and Figure 64. Pressing the “<” would set the readout values to that of the most recent solution leg if required. The solution was represented by representing what the cuts would look like if the solution being trialled was accurate. The solid line is the leading cut, and the dashed line is the penultimate cut. The trial cuts are terminated by a green dot to show their location, both of which are linked by a line to show the contact’s path. If the trial cuts are not aligned with actual cuts, as in Figure 63, this meant that the solution was incorrect. The operator would then be expected to adjust their values until the cuts aligned, such as in Figure 64. This did not mean that the solution was correct, only plausible. Once the operator was finished trialling, they could input the solution using the “|” button, which would commit a single-leg solution using the entered parameters.

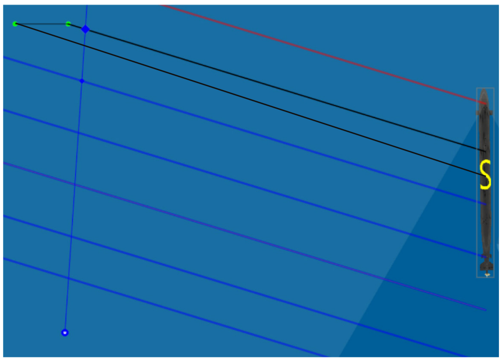


Figure 63 – Incorrect trial mode

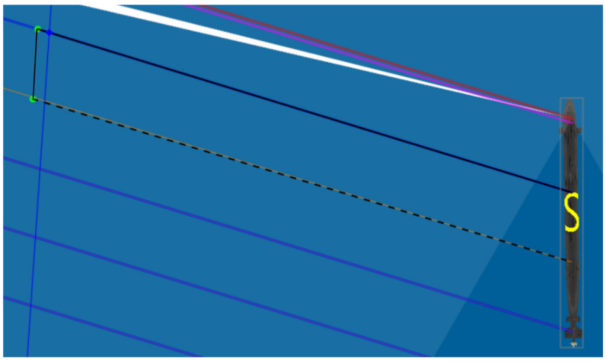


Figure 64 – Correct trial mode

The final tool was automatic solutions, which requested that the system automatically enter a classification for a solution for a contact based on known information. The classification was determined from the best-matched narrowband profile. The solution entered depended on whether the contact was merged or not. For unmerged contacts, the solution entered was a “good starting point” estimate, placing the contact at a reasonable location, ready for further analysis. This was not very accurate but could be used for operators to quickly assign contacts without reasonable solution positions to assess the tactical picture and assign priorities. Solutions computed for merged contacts were vastly more accurate and used the intersection points between each sensor cut to triangulate position. For example, if a contact was being detected on hull and bow sonar, then two cuts would be placed on the map at each cut interval, each originating from the source sensor; these cuts should intersect somewhere, see Figure 65, indicating that the contact is close to the intersect. While position (bearing and range) could be determined with one intersection, two were used to also determine the contact’s course and speed. With the positions and solution components known, a highly accurate solution could be added for the contact, see Figure 66.

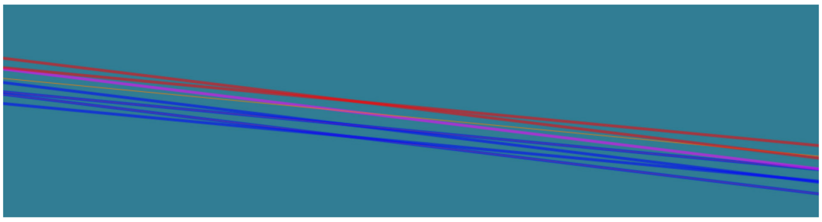


Figure 65 – Example of intersecting cuts from a merged contact

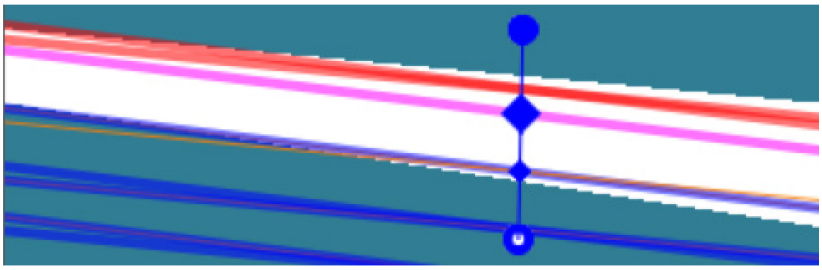


Figure 66 – Example of automatic solution applied to the intersecting cuts

The final tab in the Contact Information Panel was Notes, see Figure 67. This enabled operators to make notes, avoiding the need to store all information in their mind, or keep it written down. The latter is also a form of notes, but relies on the operator organising their notes manually, whereas GIST organises notes automatically. Each contact has its own notes, and they are chronologically organised, allowing the operator to assess whether the information is still accurate. Notes generated by the system are also shown here, such as those generated from using the automatic tool for unmerged and merged contacts, see Figure 68. Notes can also be promoted, which sends them to other operators, and shows it in the messages tab at the bottom of the interface. This allows operators to communicate data quickly without being misunderstood and could alleviate bottlenecks in communication identified in ComTET (Stanton, Roberts and Fay, 2017; Roberts, Stanton and Fay, 2018).

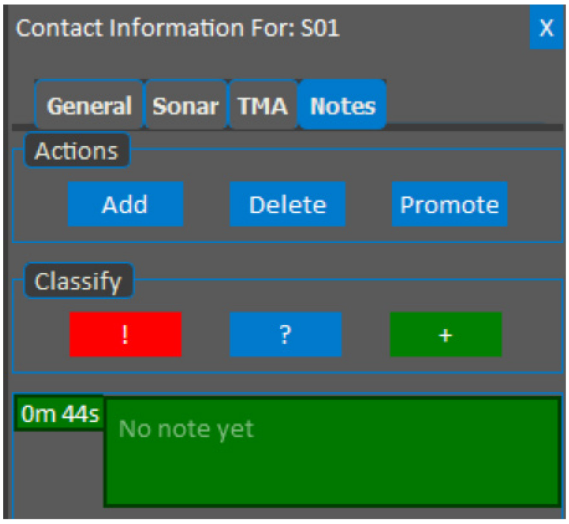


Figure 67 – GIST Contact Panel – Notes Tab screenshot

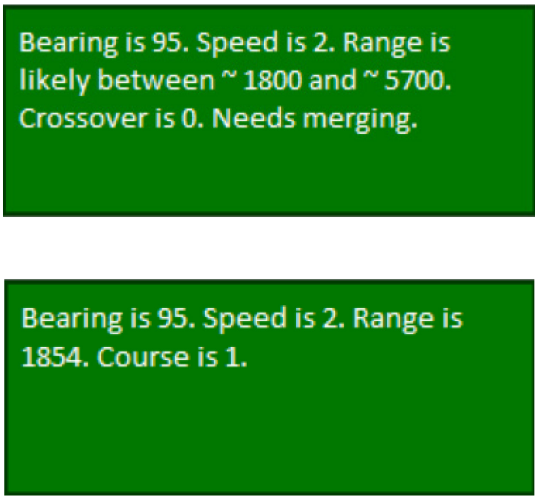


Figure 68 – Notes added for automatic solutions when the contact is not merged (top) and merged (bottom)

8.5.1 How were the design ideas met?

Table 22 presents an overview of how the design directions for Sonar and TMA were met, summarising the main points from the detail provided in Section 8.5.

Table 22 – A table summarising how the design directions from Sonar (Table 17) and TMA (Table 20) were met

Role	Design Idea	How was it met?
Sonar	The interface should be configurable and support effective data management and provide tools to process data into information for operators of all skill levels.	- Multiple configuration options are provided. - Tools for all skill levels are provided.
	Use information representations that reduce cognitive workload.	- Information is represented on the map view where possible, allowing for direct perception of the submarine's operating environment.
	Orient the interface around contacts and their information, as opposed to specific information gathering processes (DCLT) and provide configurability.	- Use of the contact widgets for operators to easily see contact information. - Use of the contact information panel to perform analysis tasks on a contact, instead of requiring multiple screens.
	Adopt a map-based display to show bearings and other related information with greater ecological validity.	- A map-based display was used.
	Represent time effectively on the map-based display.	- Cuts have different colours to represent whether they are the current detection, the latest cut, or a historic cut. - Cuts are placed on the map, allowing operators to intuitively detect when a cut was made, as it will start from a point along the historic path of ownship.
	Maintain affordances provided by traces.	- Detection markers and cuts used a combination of colour, opacity, and thickness to represent information to operators.
	Maintain utility of traces on waterfall for interpreting contact behaviour.	- A historic waterfall was retained for contacts. - A predictive waterfall was added to show functional predictive information, and to aid situational awareness.
	Maintain bimodal processing of sonar data.	- Audio was added to the interface, tied to the cursor sweep.

Role	Design Idea	How was it met?
	Move as much information management as possible into the interface so that separate sources do not need to be maintained.	- All required information can be logged in the interface. - Operators can use the notes functionality to log additional information.
	Provide operators with information from other stations if useful.	- The roles were merged, providing more information. - Notes can be used to provide information.
	Textual chat might be useful for operators.	- Notes could be used to chat between operators.
TMA	Add as much information and functionality as possible to support and/or enhance tactical picture compilation.	- A complete redesign of speedstrip entry. - Solution trials. - Speed and course circle.
	Information be better integrated to make it more accessible for analysis and decision making.	- All information is visually displayed on the map. - Information can be filtered to a specific contact, sensor, and timeframe. - All contacts and their pertinent information are shown using the contact widgets.
	Implement digital information data-sharing mechanisms for operators where possible.	- Notes functionality. - Data can be shared using the simulation engine.
	Add functional information to solutions to show their future state.	- Speedstrips had a predictor component. - There was a predictive waterfall.
	Warn operators if an entered merge appears incorrect.	- Operators were warned if merges were incorrect, based on contact bearings and source sensors.
	Always make contacts and their associated information visible.	- All contacts are shown on the left.
	Allow the operator to see an overview of all contacts at once, instead of one contact at a time.	- Pertinent information is shown using the contact widgets.

Role	Design Idea	How was it met?
	Consider merging the Sonar and TMA displays, as they will both use a map as a core component.	- The displays were merged.
	Visually represent sensor geometry in conjunction with ownship.	- Sensor geometry is visually displayed.
	Arrange contact information so that merge candidates can be identified easier.	- Contact widgets are ordered by bearing, so merge candidates can be easily identified, reviewed, and enacted.
	Add direct manipulation to speedstrips.	- Speedstrips can be directly interacted with.
	Allow multiple legs for speedstrips.	- Speedstrips can have as many legs as required.
	Add the capability to 'draw' solutions on the map and automatically enter intersections.	- Speedstrips can be drawn directly onto the map. - Intersections will automatically be found using the two most recent cuts.
	Automate tasks where there is little, or no, benefit to the operator completing them.	- Automatic cut intersection finding. - Automatic solution generation where possible.
	Warn operators if the speedstrip direction is not congruent with the cuts.	- Operators are warned when an entered solution does not intersect the last cut, or it is not the last intersection found.
	Validate entered solutions using logical constraints where possible.	- Various warnings were added to the interface, shown using notes.
	Add the ability to trial possible solutions to see if they match the constraints created by the available cuts.	- Trialling capability added, allowing operators to visually determine if a proposed solution is correct.
	Add more direct manipulation to the map to remove control buttons, freeing up more space for the map.	- Space for the map was maximised through utilising direct interaction.

8.6 Conclusion

While there are prevailing processes for arriving at an EID from a CWA, such as the work by Burns and Hajdukiewicz (2004) and Read *et al.* (2018), it remains commonly accepted that the process will involve creativity and fitting for specific purpose(s). This chapter presented design processes for GIST, accounting for both the frontend and backend design using modern software engineering practices. While they were created for GIST, they were designed to be generalisable and applied to other applications of EID where the expected output is software-based.

For the frontend, the method proposed by Burns and Hajdukiewicz (2004) was modified to be object-oriented and to directly address uncertainty, which is a key EID challenge (Vicente, 2002). Both changes were made with the underwater operating environment in mind, specifically the lack of fixed information that must be displayed, and the uncertainty involved in creating the tactical picture. This level of uncertainty exacerbates the design challenge of ensuring that operators do not perceive information being represented to them as a concrete truth; such a misinterpretation could degrade performance, and lead to incorrect actions being taken.

The backend was created using an agile Kanban approach. The backlog was populated using outputs from the Sonar and TMA Work Domain Analysis and Worker Competencies Analysis, along with the design directions from Chapter 5 and Chapter 6. The tasks were worked through to implement the underlying code for GIST. An object-oriented approach was used for the code, drawing general organisation from the objects identified when designing the frontend.

As the role of human factors continues to grow in organisations, it is important to continue to look for exploitation opportunities. By using the outputs from CWA as much as possible in a software engineering context, and reusing work between both parts of the software, it is hoped that it provides traction to the notion that concepts from both fields can be further integrated. This is especially important as software is a pervasive part of society and creating a clear path for the exploitation of CWA could see tangible benefits to software operability. The processes outlined in this chapter are complete procedures for using CWA for software engineering. However, as with other proposed EID design methods, variability in applications is assumed, and it is hoped that they serve as a useful starting point for others to adopt as required, furthering the link between CWA and software engineering.

The finished GIST was also presented in this chapter, providing an overview of its functionality and detailing key functionality aimed at supporting operators, which was cross-referenced against the design directions from previous chapters in Table 22. GIST utilises a top-down map view to present most of its data in an ecological manner, showing the submarine's operational environment as it is currently perceived. Entities on the map have counterpart information in the interface, such as contact cards, that gives the operator access to detailed information entities present, allowing them to carry out the full range of tasks required by their role. These tasks are also supported by tools that have been developed to simplify tasks that required more cognitive effort to be expended. An example of this is the course circle, which allows operators to visually identify which course(s) a contact is on, as opposed to calculating it from the waterfall. With a working version of GIST created in this chapter, the next chapter looks at evaluating it in a human in the loop experiment to determine the effect, if any, on key areas of performance.

Chapter 9 **Evaluation of Graphically Integrated Sonar and Target Motion Analysis**

9.1 Introduction

A key impetus for this thesis was to explore the suitability of Ecological Interface Design (EID; Vicente and Rasmussen, 1992) for use in the submarine control room to meet the challenges posed by future requirements. This chapter will explore whether GIST, created in Chapter 8, provides the expected improvements that EID purports to provide, and whether this research is in line with other similar experiments. Chapter 7 detailed the creation of the Graphically Integrated Sonar and Target Motion Analysis (GIST) interface, which was designed based on the analysis of Sonar (Chapter 5) and TMA (Chapter 6). While GIST is an EID like previous research in the domain, it operates using different concepts to prior work. The key differences are the merging of the Sonar and TMA roles, combined with the utilisation of a map view to display most information required. While there is substantive existing evidence that the underlying approach of EID could lead to improvements, GIST itself must be assessed to determine whether it will provide the same benefits to workload and performance as have been observed in other studies. Another key consideration was to assess whether GIST was as usable as contemporary interfaces, especially as it merged two roles together. The three factors of usability, performance, and workload must be balanced as much as possible to ensure maximal operability. For example, task performance might improve and workload lower, but if operators do not find the system usable, they might start implementing work arounds to ease their work, which could degrade overall system efficacy.

This chapter presents a Human in the Loop (HITL) experiment conducted in the Command Team Experimental Testbed (ComTET) simulator to evaluate GIST against contemporary baseline interfaces, Dangerous Waters (DW), for these three factors. Specifically, it was hypothesised that:

Hypothesis 1. **Subjective usability would be affected by the interface:** There would be a difference in perceived usability, measured by the SUS between the interfaces.

Hypothesis 2. **Objective task performance would be affected by the interface and scenario difficulty:** Performance in the three tasks identified above would be affected by the interface and scenario difficulty.

Hypothesis 3. **Workload would be affected by the interface and scenario difficulty:** Workload, as measured by the NASA-TLX and Bedford Workload Scale would be affected by the interface and scenario difficulty.

9.2 Method

Research into improving submarine control room HMIs has been ongoing for several decades, as evidenced in the work of Clarke (1999), Burns, Bryant and Chalmers (2000), Masakowski and Hardinge (2000), Dry *et al.* (2005), Ly, Huf and Henley (2007), and (Michailovs *et al.*, 2021; Michailovs *et al.*, 2022). Recent programs have sought to test interfaces and other changes using Human in the Loop (HITL) experimentation cycles, such as work in the Victoria Class Experimentation Laboratory (Bowden and Grosse, 2011; Hunter, Hazen and Randall, 2014) and ComTET (Roberts, Stanton and Fay, 2015; Roberts, Stanton and Fay, 2018).

Operators and other key stakeholders can be involved in designing an artifact to address a specific issue that is being faced, designs can be mocked up to an appropriate fidelity, and then they can be tested using robust experimental and statistical methods in simulator facilities such as ComTET. This approach is not without required time or cost, although it can significantly reduce resource requirements and create clear 'go/no go' points for designs. Consequently, multiple different ideas can be tested, and iterated if desired. In this case, it is an examination of GIST to determine if EID is a suitable design approach for Sonar and TMA. Experimental rigour also has the benefit of providing an understanding of whether the idea will work, backed by robust statistics, and where any improvements would need to be made for any adoption. As such, testing of a design in this manner is necessary for it to be considered for potential further investigation and inclusion in future submarine control rooms.

The creation of these facilities has created opportunity to test ideas created from design workshops, such as the work of Hall (2012), Fay, Roberts and Stanton (2020), and Salmon *et al.* (2016), at a reasonable level of fidelity. This was explored in Chapter 7, which aimed to validate the fidelity of the ComTET simulator so that it could reasonably be perceived as a submarine control room for the purposes of HITL experimentation. Consequently, a HITL scenario-based experiment in ComTET was chosen, as this would permit an assessment of how the participant would interact with the sensors and other underlying data when completing scenarios in an environment of reasonable fidelity. This approach was also chosen as completing a tactical picture takes time (Michailovs *et al.*, 2021) and uses complex data, which might not have been captured using alternative approaches that were did not use 'live' versions of the interfaces.

9.2.1 Participants

While similar to previous HITL experiments conducted in ComTET (Roberts, Stanton and Fay, 2017b; Stanton, Roberts and Fay, 2017; Roberts, Stanton and Fay, 2018; Stanton and Roberts, 2018; Roberts *et al.*, 2019; Stanton and Roberts, 2019; Stanton *et al.*, 2020a; Stanton *et al.*, 2020b), this experiment used an individual participant approach, as opposed to a team study. An individual participant study was chosen to avoid a confounding effect arising from changing both the interface, and merging the Sonar and TMA roles.

There was a target of recruiting 60 participants, consisting of a mix of novices and experts. Participants were mainly novices, as the relative differences were being assessed, in line with previous ComTET experiments (Stanton and Roberts, 2019), citing Walker *et al.* (2010c). They were primarily drawn from a student cohort as this overcame potential issues with submariner availability, an approach that has been adopted before in ComTET (Stanton *et al.*, 2020b). Novices were people with no experience of submarine control rooms, such as students, and experts were people familiar with the domain. Participation was entirely voluntary and had informed consent from each participant. The study received ethics approval from the University of Southampton ethics process with protocol number 10099, and MoDREC with protocol number 551/MODREC/14.

All participants were recruited through placing advertisements on message boards (physical and virtual) at the University of Southampton, sending requests for participation to industry partners (Dstl, Thales UK, BAE Systems, etc.) for dissemination, and through personal relationships (e.g., social media posts, colleagues). Requests were accompanied by recruitment posters that provided all required information. There were two versions available, each with different tagline to better appeal to the target audience; the novice version used 'experience being a submariner', and the expert version that used 'contribute to future control rooms'. This was to increase appeal to prospective participants; novices might not have felt they could contribute to future ways of working, but could be interested in the experience, whereas experts have the experience and might be more motivated to complete tasks like their job if it had an impact on future ways of working. Posters were placed in common areas around the University of Southampton, and a post was added to the internal communications platform with no specific target cohort. Companies were approached via professional connections who, at their sole discretion and in adherence to internal policies, advertised for voluntary participants from their organisations. The author's research team was highly collegiate, and an email was circulated to inform them of the study. It was made clear that participation was voluntary, and professional and personal relations would not be affected if they declined to participate or did not respond.

There were minimal restrictions on eligibility, with participants being required to be a citizen of the United Kingdom or the Commonwealth (a national security requirement of the funding) and have reasonable vision (i.e., enough to operate the interfaces without adjustments to avoid confounds). Additionally, participants had to be over 18 years of age.

Participants were offered £35 per day for expenses, pro-rated to the closest half-day increment if they exercised their right to withdraw. They were also offered lunch and there were refreshments available (tea, coffee, and various biscuits). All benefits were optional. Furthermore, some industry participants could not avail of the £35 for expenses as they were being funded by their company. This was a matter internal to each company, although no participant was excluded from being offered all benefits.

9.2.2 Measures

This thesis has concentrated on Sonar and TMA operators and their interfaces, which readily defined the roles to be included in the experiment. Sonar operators are key to maintaining the submarine's safety as this is often the primary sensor used. Combat system (a superset of TMA) operators are recognised as a key driver for combat system operation (Hautamaki, Bagnall and Small, 2005), enabling safety to be maintained. This defined the tasks for the experiment, with participants being asked to construct a reasonable tactical picture. However, not all tasks were chosen to be assessed, as this would be impractical and offer diminishing returns in assessing their objective performance. Thus, it was decided to assess the following tasks, which were core to the construction of a tactical picture (see Section 4.1.1):

- **Tracker Assignment:** Trackers assign an alias to a specific track that is used for all future analyses. Little, if anything, can be achieved without trackers. Therefore, participants should assign every trace a tracker, to ensure that they can assess it for inclusion in the tactical picture.
- **Contact Merging:** Multiple trackers can refer to one entity and should be merged to reduce participant workload. The merge process also increases the amount of information available for TMA and can improve solutions. Consequently, participants should merge trackers wherever possible to ensure they have the best possible data with which to generate solutions.
- **Solution Performance:** Solutions form the tactical picture, which is a representation of the submarine's operating environment. It is vital that this is correct, as decisions are made based on it. This affects all three of a submarine's operating tenets (remain safe, remain undetected, and complete the mission). Thus, participants should be entering the most accurate solutions possible and assessing them for ongoing accuracy.

The above data was collected by the software, with each interface recording objective data during scenarios. This included data about the submarine's environment (entity position and velocity), perceived tactical picture, and actions that operators took. As both interfaces recorded data using different formats, software was written that transformed the outputs into a common format for subsequent analysis. This analysis was performed using another program that computed desired metrics for each scenario run and compiled results into a single source.

All operators manage a large variety of information (Burns, Bryant and Chalmers, 2005), and the HMIs are a vital aspect of facilitating this. This is exacerbated by the fact that operators must complete various mission types, with decision-support systems being vital support (Burns, Bryant and Chalmers, 2000). However, it would not be enough to simply provide the information. The human dimension must be considered in the context of work to be performed (Chalmers, Easter and Potter, 2000). To consider this, it was decided that subjective measures of usability and workload would be collected; while objective performance improvements would be a valid result, this should not come at the expense of operators, which could negatively affect performance (Yan *et al.*, 2022), citing Nachreiner (1995). A reduced workload could also be indicative of GIST meeting the aim of moving as much processing as possible to SBB, which could offer increased capacity to evaluate their action space for unfamiliar and/or unexpected situations. Low- and high- difficulty versions of a scenario were created to generate a difference in workload to understand how this would affect performance, if at all.

The System Usability Scale (SUS; Brooke, 1996) was selected to assess perceived usability. It was chosen as a standard measure of usability, and one that has been assessed as reliable (Sauro and Lewis, 2012). The National Aeronautics and Space Administration Task Load Index (NASA-TLX; Hart and Staveland, 1988). Again, it was chosen as a standard measure of task load (Hart, 2006). To compliment the NASA-TLX, the Bedford Workload scale (Roscoe and Ellis, 1990) was also selected. This would allow participants to provide a single number for their workload, following the flowchart.

9.2.3 Equipment

The study was run in the ComTET simulator, which consisted of ten participant computers (two spares), one experimental computer, and a seat for experiments to monitor the experiments, see Figure 69. Participant computers were mounted in a representative Multi-Function Console (cabinets that house technical systems, see Section 4.1.1.3), with one ~22” 1920 × 1080 resolution monitor, a mouse, a keyboard, and a whiteboard. They were also provided with whiteboard pens, wipes, ballpoint pens, and paper.

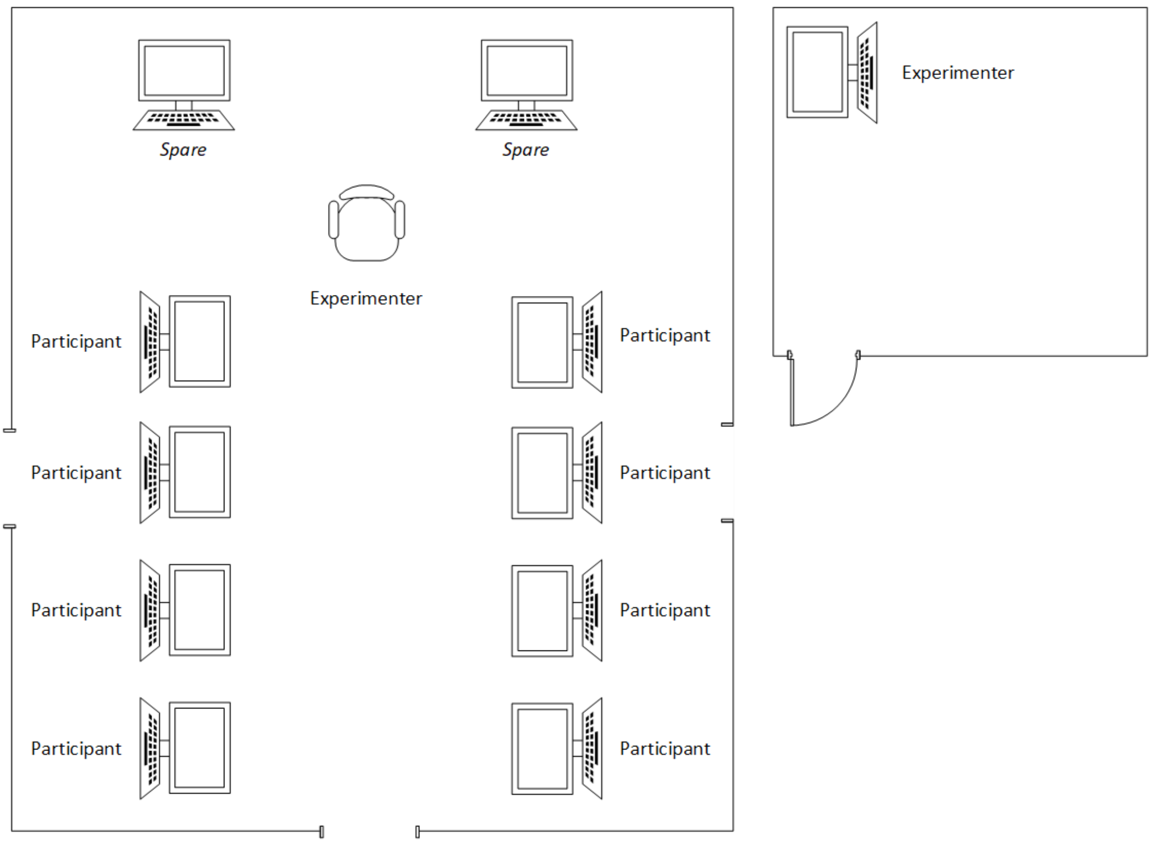


Figure 69 – Schematic layout of participants for the experiment. Not to scale.

The experiment was run digitally using Clamshell, a bespoke laboratory management software developed by the author (Fay, Stanton and Roberts, 2018), which includes digital versions of scales and measures. Clamshell was used to launch software, scales, and forms where possible, and tagged data with a unique participant identifier.

The first interface (DW) was formed of the Sonar and TMA screens in DW (Version 106, Build 651). These were representative of contemporary ways of working and served as the baseline. Three Sonar screens were used, Broadband, Narrowband, and the DEMON. The second interface (GIST) was formed of a single top-down map view that combined sonar and TMA functionality. It was built using the Simulation Engine II (SEII) Software Development Kit (SDK; Version 1.3.248, 2.8.264.0, Runtime 2.6; see Section 8.4 for details). SEII and GIST were configured to be the same as DW as much as possible, including the underlying sonar model where possible.

Finally, a paper-based demographics form was used to request optional data from each participant. The data from all forms was digitized, and, as with other data, collated into the compiled data.

9.2.4 Design

The study employed a 2 × 2 within-subjects repeated-measures design. The independent variables were the interface used (DW, GIST) and scenario difficulty (Low, High; described in Table 23). They will be referred to as “Interface” and “Scenario” henceforth. The dependent variables are detailed in Table 24. The group column shows which collective metric they were assessed as, which is reflected in Section 9.3. Participants were asked to complete Sonar and TMA tasks for each scenario, with the objective of building and maintaining an accurate tactical picture. To achieve this objective, they were asked to follow the Detect, Classify, Localise, Track (DCLT) initialism for Sonar (see Section 4.1.1.4 for an overview), and to continuously ensure that all entities had appropriate solutions for TMA (see Section 4.1.1.5 for an overview). Acceptable performance was generally defined, such as ensuring that solutions use known bearing and speed. However, exact measures were not provided, as these are usually provided in context by senior control room personnel. Consequently, participants were asked to complete tasks as accurately as they could in each scenario. Aside from the general order of tasks, no requests were made for task ordering, time allocation, or frequency.

Table 23 – Description of scenarios used

Difficulty	Entities	Key Features
Low	1 × Submarine 4 × Fishing Boat	- Consistent courses - Contacts not in baffles
High	1 × Submarine 5 × Fishing Boat 3 × Surveillance Boat in DW / Corvette in GIST (Similar models were chosen as not all models were available)	- Ownship changes course at 10 minutes - Contacts can enter and emerge from baffles

Table 24 – Description of independent variables

Independent Variable		Description
Subjective Usability	SUS Score	The overall score from the SUS after all scenarios were completed.
Tracker Assignment	% Trackers Assigned	The percentage of possible trackers that were assigned. The maximum number of trackers is the number of entities multiplied by the number of available sensors.

Independent Variable		Description
Contact Merging	% Merges	The percentage of entities that formed part of a merge at any point during the scenario.
	% Correct Merges	The percentage of completed merges that were correctly merged.
Solution Performance	Best Solution Position Δ	The smallest difference between an entity's location and its solution at the time it is entered or affected by contact management operations (merges and splits). This differs from the solutions range component, which is measured from ownship to the solution's position. This is to account for bearing as well.
	Best Dead-Reckoned Position Δ	The difference between an entity's location and solution after the "Best Solution Position Δ" has been dead reckoned by 30 seconds. This accounts for the solution's course and speed.
Subjective Workload	NASA-TLX	The raw NASA-TLX score, with a range of 0 – 120 (0 – 20 for each of the six questions).
	Bedford	The Bedford score, with a range of 0 – 10.

9.2.5 Procedure

Participants attended the ComTET simulator for two days for approximately eight hours each day and had regular comfort breaks and a 30-minute lunch break each day. Upon arrival on the first day, participants were seated and provided with a verbal general introduction to the study. They were asked to read through a Participant Information Sheet (PIS) provided to them and sign a consent form. Experimenters were available to answer any questions and assuage any concerns.

Participants were trained on both simulation engines using a combination of video tutorials and experimenter-guided practice over three hours, see Table 25. Due to restrictions on lab access arising from the Coronavirus 19 pandemic, all planned interface experiments for this thesis were merged into one large study using a repeated measures design. The tutorial videos for an interface presented in the next chapter, marked by a '*', are included for completeness. Tutorial videos were either created for this experiment or sourced from previous ComTET experiments. All videos were designed to be reusable for further experiments. Participants were invited to ask questions whenever required, except during the videos, as their answers might be covered later in the video being watched. The first video played was a general introduction to submarine control room operations and common underlying theory, such as what bearing, course, range, and speed are. After this, theory videos for both roles were played. A Sonar video explained Sonar theory, procedures, and waterfalls. This was followed by a counterpart TMA video that detailed how to generate and maintain solutions. Next, a patterns video was played, detailing common patterns in Sonar and TMA data that would help when completing tasks. Finally, a specific video was played for each interface, detailing their layout and operation. Each interface-specific video was followed by approximately 15 – 20 minutes of practice on the interface, with experimenters available to answer any questions, consolidate any common questions or misconceptions, and provide constructive feedback. Experimenters also asked participants to demonstrate tasks (e.g., "can you designate that contact?") to verify that they had understood the training material.

Table 25 – Overview of training that was provided to participants

Duration (minutes)	**Training Component**	**Purpose**
30	General video	Introduce submarine control rooms and associated general concepts.
40	Sonar video	Cover sonar theory.
20	TMA video	Cover TMA theory.
30	Patterns video	How to interpret patterns in sonar and TMA.
15	DW sonar video	How to operate DW sonar.
15	DW TMA video	How to operate DW TMA.
15	DW Practice	Familiarisation with DW.
20	Mashup video*	How to operate the Mashup interface.
20	Mashup practice*	Familiarisation with the Mashup interface.
20	GIST video	How to operate GIST.
15	GIST practice	Familiarisation with GIST.
30	Overall practice	Further familiarisation.
Σ 3 h 30 m	** Training activities for an interface evaluated in the next chapter.*	

Participants then completed all scenarios for most of the remaining time, with each scenario lasting for 20 minutes. Scenario order was counterbalanced using a Latin square to prevent order or practice effects. The square consisted of an arbitrary initial order of scenarios for each scenario run (i.e., the 1st, 2nd, …, Nth scenario), which was then shifted by one for each planned participant in the experiment, see Table 26 as an example. Each participant was assigned a line of counterbalancing from the table, and their scenarios were run in the dictated order, save for where running the planned scenario was not possible due to issues (e.g., simulation engine not starting, or sound not working). In this instance, the planned scenario was swapped with the next scenario in the order that could be run.

Table 26 – Exemplar Latin square counterbalancing used for the experiment. Arbitrary background colours are used for clarity.

		Scenario			
		1st	2nd	3rd	4th
Participant	1	Low – DW	High – DW	Low – GIST	High – GIST
	2	High – GIST	Low – DW	High – DW	Low – GIST
	3	Low – GIST	High – GIST	Low – DW	High – DW
	4	High – DW	Low – GIST	High – GIST	Low – DW

Legend				
	Low – DW	High – DW	Low – GIST	High – GIST

For each scenario, participants were asked to follow the steps defined in Section 4.1, described in their training, to construct and maintain a reasonable tactical picture. An experimenter was available for participants to request technical assistance but would not assist with the actual tasks. Once the scenario time elapsed the interface was paused, minimised, or closed, and participants were presented with the NASA TLX and Bedford scales, one at a time. Once they were completed, participants could opt for a short comfort break, while the next scenario was set up.

Once all scenarios were completed, participants completed a SUS for each interface. Participants were then invited to discuss the interfaces, followed by a debrief of the entire experiment. The latter was started by standardised prose so that all participants received the same information. Finally, participants were offered £35 for expenses, which concluded their participation.

9.2.6 Analysis of Data

All statistical analyses were conducted using IBM SPSS Statistics (version 28.1.1.0 (142)) and Excel (Microsoft 365). Data met most assumptions for conducting parametric tests, although normality was routinely violated when assessed using a Shapiro-Wilk test. This test was chosen as it is considered the most powerful test among the commonly used normality tests (Mohd Razali and Yap, 2011) and provides higher power levels (Mendes and Pala, 2003). As the sample size is greater than the commonly accepted size of 30 for the Central Limit Theorem (CLT; Field, 2018), parametric tests were still selected. All data, except for solution position data, was treated as-is. Solution position data was transformed using a common logarithm to reduce the effect of skew.

A Paired-Samples T Test was conducted on the SUS score. Tracker assignment performance was analysed using a two-way Analysis of Variance (ANOVA). The remaining categories were analysed using a two-way Multiple Analysis of Variance (MANOVA). Bonferroni corrections were used to account for repeated testing.

9.3 Results

There were 45 participants in total. Their ages ranged from 20 to 65 (M = 26.51, SD = 9.20), with 30 males and 15 females. Results are not disaggregated by these categories to maintain appropriate statistical power, especially for tests where the population was decreased due to only a subset of participants having completed the tasks (e.g., only 16 participants completed solutions for all four conditions).

9.3.1 Subjective Usability – System Usability Scale

Participants reported higher usability when using GIST than when using DW. The means and associated standard deviation for answers to each question, and the total scores, are presented in Table 27. SUS scores were higher for GIST (M = 54.67, SD = 23.10) than for DW (M = 44.11, SD = 20.79). The difference was significant, $t(44) = -2.12$, $p = .04$ (2-sided).

Table 27 – Means for each SUS question and the total score by interface, with cell shading representing a favourable rating for each question, based on the mean.

	DW		GIST	
	M	SD	M	SD
Q1. I think that I would like to use this system frequently	2.64	1.23	2.98	1.26
Q2. I found the system unnecessarily complex	3.38	1.22	2.60	1.14
Q3. I thought the system was easy to use	2.67	1.10	3.27	1.29
Q4. I think that I would need the support of a technical person to be able to use this system	2.96	1.23	2.96	1.28
Q5. I found the various functions in this system were well integrated	2.78	1.11	3.47	1.20
Q6. I thought there was too much inconsistency in this system	2.60	1.16	3.00	1.15
Q7. I would imagine that most people would learn to use this system very quickly	2.87	1.28	3.42	1.24
Q8. I found the system very cumbersome to use	3.76	1.25	2.64	1.16
Q9. I felt very confident using the system	2.96	1.28	2.78	1.28

Q10. I needed to learn a lot of things before I could get going with this system	3.58	1.27	2.84	1.25
Overall SUS Score	44.11	20.56	54.67	22.85

9.3.2 Tracker Assignment – % Trackers Assigned

Participants assigned more trackers in GIST than they did in DW. The percentage of trackers assigned was significantly affected by the interface used, $F(1, 36) = 58.21$, $p = < .001$. The mean percent of trackers assigned was 90.88% (SE = .77, SD = 4.68) in GIST and 65.46% (SE = 3.29, SD = 19.98) in DW.

Participants assigned more trackers in low difficulty scenarios. The percentage of trackers assigned was significantly affected by the scenario, $F(1, 36) = 6.99$, $p = .012$. The mean percent of trackers assigned was 80.27% (SE = 2.02, SD = 12.28) for the low scenario and 76.1% (SE = 1.73, SD = 10.5) for the high scenario.

It was revealed that scenarios conducted in GIST had a higher tracker assignment percentage than in their respective DW counterparts, see Table 28 and Figure 70. The interaction effect between the interface and scenario was significant, $F(1, 36) = 38.075$, $p = <.001$.

Table 28 – Tracker percentage assignment between DW and GIST for both scenarios

					95% Confidence Interval	
Interface	**Scenario**	**Mean**	**Std. Error**	**Std. Dev.**	**Lower Bound**	**Upper Bound**
DW	Low	72.97	4.04	24.56	64.78	81.16
	High	57.94	3.11	18.91	51.64	64.24
GIST	Low	87.5	.000	0.00	87.5	87.5
	High	94.26	1.54	9.36	91.14	97.37

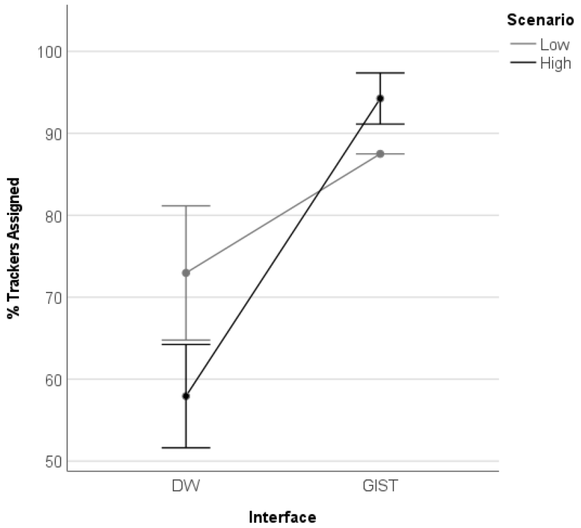


Figure 70 – Tracker percentage assignment between DW and GIST for both scenarios. 95% Confidence Intervals (CI) indicated.

9.3.3 Contact Merging – % Merges & % Correct Merges

Participants completed more merges, which were more accurate, in GIST when compared to DW. The MANVOA revealed that merge performance was significantly affected by the interface used, F (2, 35) = 49.37, V = .74, p = <.001. Subsequent univariate tests showed that there was a significant effect on the percentage of merges completed, F (1, 36) = 101.17, p = < .001, and there was a significant effect on the percentage of merges that were correct, F (1, 36) = 79.11, p = < .001. The percentage of merges completed was higher in GIST (M = 74.32%, SE = 3.78, SD = 23) than DW (M = 16.72%, SE = 3.95, SD = 24.03). The percent of correct merges was also higher in GIST (M = 88.92%, SE = 4.36, SD = 26.54) than DW (M = 24.78%, SE = 5.76, SD = 35.05).

Participants performed more merges in the low difficulty scenarios, and the accuracy of their merges was higher in low difficulty scenarios. Merge performance was significantly affected by the scenario, F (2, 35) = 9.76, p = <.001. Post hoc testing revealed that there was not a significant effect on the percentage of merges completed, F (1, 36) = .60, p = .442, although there was a significant effect on the percentage of merges that were correct, F (1, 36) = 11.788, p = .002. The percentage of merges that were completed was higher in the low scenarios (M = 46.62%, SE = 3.55, SD = 21.8) compared to the high scenarios (M = 44.43%, SE = 2.22, SD = 13.49). The percentage of correct merges was higher in the low scenarios (M = 63.74%, SE = 4.86, SD = 29.54) than the high scenarios (M = 49.95%, SE = 3.27, SD = 19.91).

Participants had better merge performance in GIST. There was a significant interaction effect between the interface and scenario, F (2, 35) = 15.44, p = < .001. Post hoc testing revealed that there was a significant effect on the percentage of merges completed, F (1, 36) = 29.47, p = < .001, and there was a significant effect on the percentage of merges that were correct, F (1, 36) = 8.87, p = .005. It was revealed that both measures of merge performance were higher in GIST for both scenarios, see Table 29, Figure 71, and Figure 72.

Table 29 – Merge performance for each interface and scenario combination

Measure	Interface	Scenario	Mean	Std. Error	Std. Dev.	95% Confidence Interval	
						Lower Bound	Upper Bound
Completed	DW	Low	26.35	6.05	36.78	14.09	38.61
		High	7.1	2.85	17.32	1.32	12.87
	GIST	Low	66.89	3.63	22.09	59.53	74.26
		High	81.76	4.46	27.10	72.72	90.79
Correct	DW	Low	37.84	8.08	49.17	21.45	54.23
		High	11.71	5.20	31.64	1.16	22.26
	GIST	Low	89.64	4.72	28.69	80.07	99.21
		High	88.19	4.81	29.26	78.44	97.95

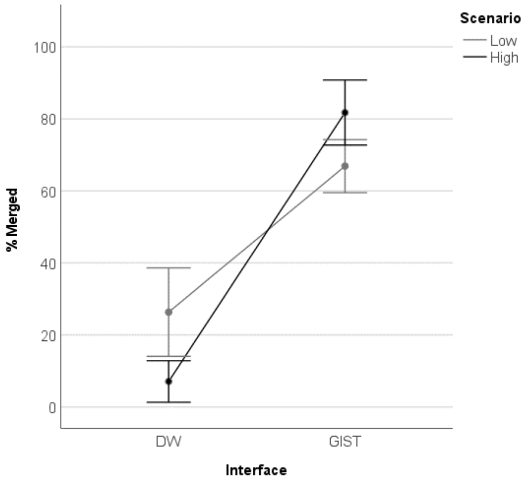


Figure 71 – % Merged for DW and GIST. 95% CI indicated.

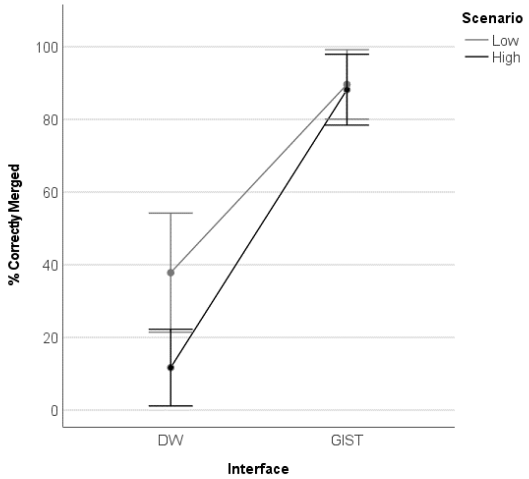


Figure 72 – % Correctly merged for DW and GIST. 95% CI indicated.

9.3.4 Solution Performance – Best Solution Position Δ & Best Dead-Reckoned Position Δ

Participants entered more accurate solutions using GIST when compared to DW. The MANOVA revealed that solution performance was significantly affected by the interface used, F (2, 14) = 8.87, V = .56, p = .003. The means for each metric are shown in Table 30, Figure 73 and Figure 74. Subsequent univariate tests revealed that there were significant effects for entered solution positions, F (1, 15) = 18.14, p = < .001, and dead-reckoned solution positions, F (1, 15) = 14.86, p = < .001. The most accurate solution position had a mean of 572.8 metres (SE = 1.27, SD = 5.08) in DW, whereas this was 70.47 metres (SE = 1.54, SD = 6.16) in GIST. Dead-reckoned solutions were also more accurate in GIST when compared to DW. The most accurate dead-reckoned solution position had a mean of 597.04 meters (SE = 1.25, SD = 5) in DW, whereas this was 64.86 meters (SE = 1.55, SD = 6.2) in GIST. Solution performance was not significantly affected by the scenario, F (2, 14) = 1.4, p = .279. There was not a significant interaction between the interface and scenario, F (2, 14) = .38, p = .694.

Table 30 – Solution Performance for DW and GIST

Measure	Interface	Scenario	Mean	Std. Error	Std. Dev.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Best Solution Position Δ	DW	Low	955.85	1.53	6.14	383.81	2,380.43
		High	343.08	1.38	5.52	172.71	681.50
	GIST	Low	82.74	1.76	7.03	24.88	275.16
		High	60.07	1.72	6.87	18.98	190.14
Best Dead-Reckoned Position Δ	DW	Low	958.61	1.54	6.17	380.93	2,412.35
		High	371.00	1.37	5.49	189.06	728.03
	GIST	Low	73.31	1.82	7.28	20.48	262.44
		High	57.30	1.73	6.92	17.83	184.16

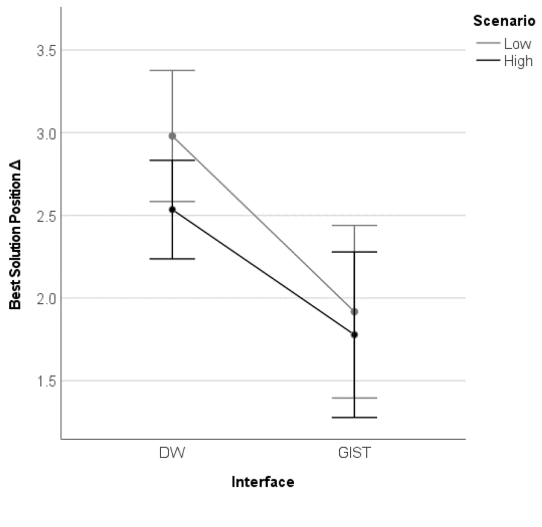


Figure 73 – Best Solution Position Δ for DW and GIST. 95% CI indicated.

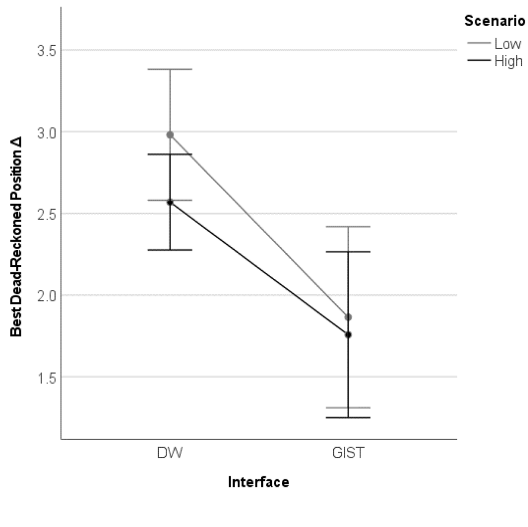


Figure 74 – Best Dead-Reckoned Position Δ for DW and GIST. 95% CI indicated.

As a limited number of participants had completed solutions across all 4 permutations (N = 16), subsequent one-way ANOVAs were completed to assert the effects of each independent variable. Testing with scenario as the only independent variable confirmed that the scenario did not affect performance within DW, $F(2, 31) = 1.21$, $p = .312$, and confirmed that the scenario did not affect performance within GIST, $F(2, 19) = .08$, $p = .928$. Testing with interface as the only independent variable revealed that there was a significant effect for low scenarios, $F(2, 26) = 12.56$, $p = < .001$, but there was not a significant effect for high scenarios, $F(2, 17) = 2.65$, $p = .99$. Subsequent testing for the low scenarios showed that there was a significant effect of both entered position, $F(1, 27) = 24.92$, $p < .001$, and dead-reckoned position, $F(1, 27) = 25.88$, $p < .001$.

9.3.5 Subjective Workload – NASA-TLX & Bedford

The MANOVA revealed that perceived subjective workload was not significantly affected by the interface used $F(2, 32) = 2.68$, $V = .14$, $p = .084$.

Participants rated the high difficulty scenarios more highly, signifying a higher perceived workload, when compared to the low difficulty scenarios, see Table 31 and Table 32. Note that the averages are different between Table 31 and Table 32 as the former excluded cases that were not complete (all Bedford and NASA-TLX scores present for a participant). Perceived subjective workload was significantly affected by the scenario, as measured by a MANOVA, $F(2, 32) = 25.59$, $V = .62$, $p = < .001$. Subsequent univariate tests revealed a significant effect for both the NASA-TLX, $F(1, 33) = 48.31$, $p = < .001$, and Bedford, $F(1, 33) = 25.36$, $p = < .001$. NASA-TLX scores increased with scenario difficulty, with low scenarios having a mean of 54.21 (SE = 1.99, SD = 11.63) and high scenarios having a mean of 66.04 (SE = 2.23, SD = 13.03). This was also true for Bedford scores, with low scenarios having a mean of 6.11 (SE = .38, SD = 2.19) and high scenarios having a mean of 7.91 (SE = .29, SD = 1.71).

The MANOVA revealed that there was no significant interaction effect between interface and scenario, $F(2, 32) = .752$, $p = .479$.

Table 31 – Workload scores for DW and GIST

Measure	Interface	Scenario	Mean	Std. Error	Std. Dev.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
TLX	DW	Low	58.27	2.61	15.22	52.95	63.58
		High	68.5	2.84	16.54	62.73	74.27
	GIST	Low	50.15	2.36	13.76	45.35	54.95
		High	63.59	3.01	17.57	57.46	69.72
Bedford	DW	Low	6.79	.44	2.58	5.89	7.69
		High	8.18	.35	2.06	7.46	8.89
	GIST	Low	5.43	.53	3.08	4.35	6.50
		High	7.65	.48	44.59	6.68	8.62

Table 32 – Means for each question of NASA-TLX

	DW				GIST			
	Low		High		Low		High	
	M	SD	M	SD	M	SD	M	SD
Mental Demand	11.2	4.48	13.57	4.86	8.97	5.02	12.54	5.23
Physical Demand	3.43	3.66	5.36	5.26	2.26	2.69	2.95	4.13
Temporal Demand	9.45	5.01	14.45	4.59	6.82	4.65	13.1	5.73
Performance	11.55	5.16	9	5.18	10.18	6.18	9.05	6.28
Effort	11.61	4.45	13.8	4.21	8.79	5.26	12.10	5.09
Frustration	9.8	5.87	11.43	6.02	10.82	5.74	11.97	5.29
Total	57.05	14.98	67.61	16.51	47.85	14.29	61.72	17.73

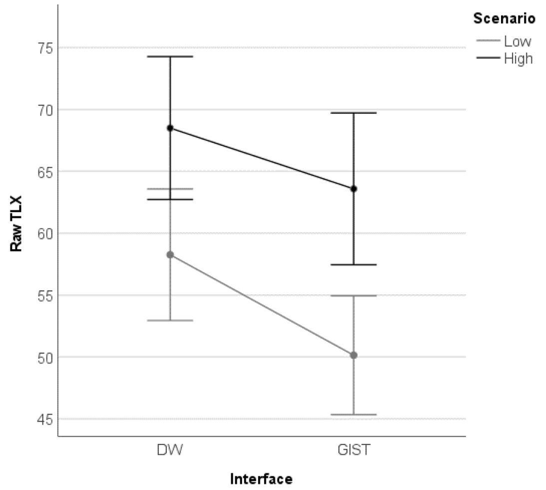


Figure 75 – TLX scores for DW and GIST. 95% CI indicated.

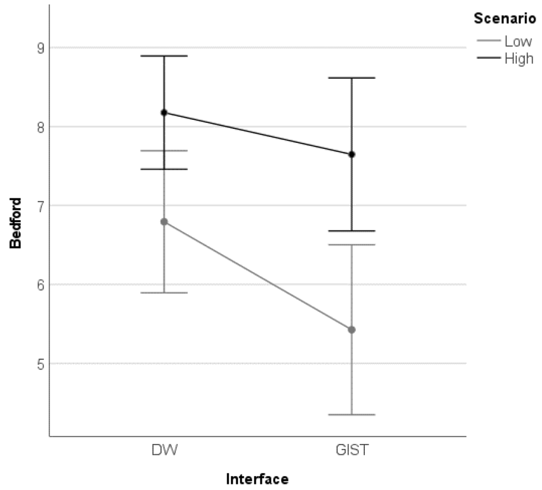


Figure 76 – Bedford scores for DW and GIST. 95% CI indicated.

9.4 Discussion

9.4.1 Hypothesis 1. Subjective usability would be affected by the interface

Brief Results Summary: Participants rated GIST higher than DW for subjective usability using the SUS scale. This supports the hypothesis.

The increased SUS score supports previously seen results that usability could be improved through a holistic consideration and integration of system functionality (Hall, 2012) and implementation of appropriate levels of automation (Calhoun *et al.*, 2013). It also supports previous experiments that found improvement in usability when using EID (Behymer, 2017; Tran, Hilliard and Jamieson, 2017; King, Read and Salmon, 2022). However, while GIST was rated higher overall, it was not rated most favourably in all component questions for the SUS, see Table 27. This provides avenues to explore for future usability improvements. This is in line with a study conducted by Burns, Kuo and Ng (2003), which yielded performance improvements, but participants gave feedback that future iterations could be tweaked to improve usability. GIST was rated less favourably for question 6 (inconsistency) and question 9 (confidence). Inconsistency in the interface could have increased task completion times, as has previously been observed with ecological displays (Mazaeva and Bisantz, 2014). This would suggest that an area of future development for GIST would be to increase interface consistency. However, it has been argued that seeking a completely consistent interface may not be appropriate when there is a focus on tasks that users complete (Grudin, 1989). Bennett *et al.* (2011) argue that consistency could allow operators to bypass information processing limits, utilising Skill-based Behaviour as much as possible, arising from homogenous expected responses in relation to associated triggers (Rasmussen, 1983; Rasmussen and Vicente, 1989). They further argue that valuable consistency arises from creating *"consistent mappings between interface representations and functional properties of the task situation"*. In the context of GIST, this suggests that future iterations should seek to ensure consistent tooling for similar tasks. An example of this would be the utilisation of speedstrips, trials, and the course and speed circle tools to create solutions. They all have the same goal but have different interaction methods. While they were designed to address different objectives to meet the overall goal of solution entry, it could be beneficial to remove interaction discrepancies in joint objectives.

The need for improving consistency might have also affected participant responses to the confidence question of the SUS. It is proposed that participant confidence was diminished for the use of GIST in relation to its consistency. This is supported by the experiments of Schneider and Shiffrin (1977), who demonstrated that consistent stimuli increased task performance, suggesting that the inconsistency might have degraded participants' confidence in their ability to appropriately use the provided functionality. This degradation might have been related to participant training, with them not being sure if they were using tools correctly due to inconsistent utilisation from their point of view. The effect of inconsistent mapping has been shown to degrade training outcomes (Goettl, 2006), citing (Schneider and Shiffrin, 1977). Therefore, while participants were provided with full training and had the opportunity to ask questions, this suggests that interface training for participants could be improved. Conversely, it might be possible that the interface training was suitable, and that participants were not sufficiently trained in the underlying concepts of Sonar and TMA operation, providing them with the enough knowledge to know why the tools were useful.

Regardless of GIST showing a relative improvement, neither interface scored above the commonly accepted baseline of 70 (Bangor, Kortum and Miller, 2008). Though, usability is not a dichotomous distinction, rather a scale. Bangor, Kortum and Miller (2008) proposed an adjective style grading system to categorise SUS scores. GIST met the threshold for being described as "Okay" (≥ 52.01 and ≤ 72.75), whereas DW would be in the "Poor" category (≥ 39.17 and ≤ 52.01). Conversely, Sauro and Lewis (2012) suggested that it might be fairer to grade on a curve, owing to the difficulty in achieving a top grade (GIST was graded 'D' for being ≥ 51.7 and ≤ 62.6, and DW was graded 'F' for being ≤ 51.7). While GIST needs further usability improvement, the literature suggests that the current usability difference was enough to distinguish it sufficiently as an improvement, both in relative terms (a simple comparison) and by categorising using aggregated data as Bangor, Kortum and Miller (2008) and Sauro and Lewis (2012) suggested.

9.4.2 Hypothesis 2. Objective task performance would be affected by the interface and scenario difficulty

Brief Results Summary – Tracker Assignment: Participants assigned more trackers in GIST than they did in DW, with the number of trackers being significantly affected by the interface. The number of trackers assigned was also significantly affected by the scenario difficulty, with slightly more trackers being assigned in the low difficulty scenarios, see Table 28. Scenarios completed in GIST had a higher tracker assignment percentage, which was statistically significant, see Table 28.

Brief Results Summary – Contact Merging: Participants completed more merges, which were more accurate, in GIST when compared to DW, see Table 29. Merge performance was statistically significantly affected by scenario difficulty; more merges were completed in low difficulty scenarios, and the percentage of correct merges was higher in the low difficulty scenarios. A combination of interface and scenario had a statistically significant effect on merge performance, with both measures being higher in GIST for both scenarios.

Brief Results Summary – Solution Performance: Solutions were more accurate in GIST than they were in DW, see Table 30, a difference that was statistically significant. There was not a statistically significant difference between the scenarios, nor the interaction between scenarios and the interface.

The results mostly support the hypothesis, save for solution performance, which was only affected by the interface used. Improvements to task performance from using GIST affirm that an EID can improve performance (Torenvliet, Jamieson and Vicente, 2000; Vicente, 2002; Bennett and Flach, 2019). The experiment also confirms the results of an experiment conducted by Michailovs *et al.* (2021), who found that teams using integrated information displays based on EID principles constructed a more accurate tactical picture. It also shows that EID is a viable option to ensure that the increasing amount of data in a control room does not exceed operator capability to process it, a problem identified by Woods, Patterson and Roth (2002). The results are also in agreement with studies in various other domains that have shown an improvement when using EID, such as manufacturing systems (Cravens, 2021), healthcare (Shier *et al.*, 2018; Zestic *et al.*, 2019), driving (Jamson, Hibberd and Merat, 2015; Schewe and Vollrath, 2020), and air traffic control (Borst *et al.*, 2017). It has been stated that interactions between perceptual and cognitive processes can affect operator performance (Hanisch, Kramer and Hulin, 1991; Masakowski and Hardinge, 2000), as can overly complex screens (Coll and Wingertsman, 1990). The reduction of these factors through an EID approach to reduce required cognitive processing, and simplification of the screens through merging them into one, may have also contributed to better task performance.

Task performance was improved when using GIST, despite not reaching the recommended 70 threshold for a SUS score (Bangor, Kortum and Miller, 2008). This is congruent with the result of an optronics study conducted by Michailovs *et al.* (2022), which compared a contemporary system, showing an image from a specific bearing, to a new design, showing all bearings at once by stacking five segments of the 360°; while there was not a statistically significant difference between the mean SUS scores (contemporary = 60.10 and new design = 65.47), the new design showed improvements in some aspects of performance. This also demonstrates that a new interface need not score ≥ 70 in the SUS to create credible improvement and supports the notion that the SUS is a relative scale (Bangor, Kortum and Miller, 2008; Sauro and Lewis, 2012). However, it would be pertinent to consider whether the different aesthetics contributed towards the usability difference, which Tractinsky, Katz and Ikar (2000) identified a connection between. The sole contribution of aesthetics to the SUS would be interesting to assess, although as GIST aimed to make both aesthetic and functional improvements, which are often implemented together, this could be difficult to substantiate with appropriate confidence to draw conclusions from within the context of the current study.

The improvement in tracker assignment percentages between the two interfaces is interesting as both tasks are simple Rule-Based Behaviour tasks from the training (Vicente, 1999b), with participants being instructed to assign trackers to all sonar traces that did not already have one. Thus, there should have been no differences, if at all, between the number of trackers assigned between interfaces using the same scenario. However, the difference can likely be attributed to integrated sonar design and detection automation that GIST uses. Participants were required to switch between arrays, with only one being shown at a time, and manually find traces in the provided waterfalls. These traces might not have been immediately clear. Conversely, GIST had the option of displaying all sonar data at once, and started in this mode, with an automation feature that automatically located traces that should be made into contacts. If participants did not actively switch arrays in DW, then they would have missed up to 50% of the available trackers for a scenario (the counterpart trackers to the array that they were viewing, sans those in the baffles). This suggests that Michailovs *et al.* (2022), citing Posner (1980), were correct when they proposed that a singular screen could cause a natural attentional spotlight.

Training is another likely causal factor, as participants did assign trackers, although some might not have switched arrays to check for more. While this process was covered in the training, it is a learning point for future studies, suggesting that reworked training might be appropriate. Taking the results as-is, however, supports the growing evidence that making information accessible to operators in an integrated or consolidated fashion can (Dominguez *et al.*, 2006; Michailovs *et al.*, 2021; Michailovs *et al.*, 2022) and does improve performance. This was demonstrated by experiments trialling a Victoria class Integrated Information Display (Hunter, Hazen and Randall, 2014), representing all 360° of optronics in one screen (Michailovs *et al.*, 2022), tactical picture compilation using distributed data Michailovs *et al.* (2021), and colocation of Sonar and TMA operators (Stanton and Roberts, 2020).

Neither interface prompted to assign trackers to traces that did not have trackers, despite them being core to subsequent tasks. Future iterations to GIST, or other interfaces, should seek to explore this, ensuring that operators are appropriately prompted or guided to perform key tasks that have been neglected. This would be in keeping with the EID notion of supporting behaviour at all levels of the SRK Taxonomy, with less experienced operators provided with more assistance. Such functionality would have to be implemented with considerations of an objective of EID in allowing operators to retain autonomy in their actions (Vicente and Rasmussen, 1992; Borst, Flach and Ellerbroek, 2015), ensuring that operators are not harangued into completing an action that was not part of their plan. Having to pay attention to an alert might also force processing into a Knowledge-Based Behaviour, which could increase cognitive workload, potentially degrading performance. Conversely, it might be appropriate to draw operators attention to certain neglected tasks that they could have a detrimental effect on task performance, and this could affect the ownship safety (Meshkati, 1991; National Transportation Safety Board, 2001; Marine Accident Investigation Branch, 2016). This could be achieved using alarms in increasing levels of intrusiveness, selecting a method suitable for the situation (Hautamaki, Bagnall and Small, 2005). An alternative approach could be for the automation to complete certain tasks when the operator is experiencing a high workload (Hou *et al.*, 2015). Another consideration arises from how submarine control rooms operate, with lots of information stored in the command teams' minds, especially at the command level (Dominguez *et al.*, 2006); operators may have already completed a task but the interface would be unaware of this, unless operators were required to enter all decisions into the interface, which could increase cognitive workload and time taken to process information.

As with tracker assignment performance, merge performance was better in GIST. The lower scores for the percentage of entities entered into a merge (% Merged) in DW likely have the same causal factors as the tracker assignment. Participants could have created merges from all trackers they assigned, but without a complete collection of trackers, these merges would only be a subset of entities. This is supported by the low percentage of correct merges data in DW, see Table 29, which could have been caused by participants incorrectly merging trackers on the same array, a fundamentally incorrect possibility in the scenarios. This suggests that the work domain was not able to be directly perceived, a goal of EID (Gibson, 1979; Rasmussen and Vicente, 1989; McIlroy, 2016), correctly. Consequently, operators failed to understand that they were violating a constraint and that their action was incorrect.

Another possible explanation for the lower number of trackers in DW could be that participants switched between arrays to assign trackers but did not assign all trackers before the scenario ended as they were completing subsequent analysis tasks. However, this might not be the case based on the low percentage of merges correct for DW, as higher scores would be expected. A further possible contributory factor is that GIST produced explicit warning messages when merges were performed incorrectly (not on different arrays and/or component contacts had a sufficiently large bearing disparity), whereas DW did not. This provided feedback to participants, which could have led them to revise their merge. It is known that feedback can improve performance in submarine control rooms, and has been demonstrated for Sonar operators following a training program (Winchell, Panell and Pickering, 1976), as well as periscope operators when using training software (Landsberg *et al.*, 2012; Van Buskirk *et al.*, 2019; Schroeder *et al.*, 2020). Furthermore, DW only showed physical information regarding merge correctness by showing all cuts at once. An incorrect merge could be determined by understanding that the cuts from the merge candidates were not closely aligned with each other. GIST added functional information to this process using the warning notes, confirming the assertion that displaying both types of information can lead to better performance (Torenvliet, Jamieson and Vicente, 2000; Vicente, 2002). It would be interesting for a future study to explore why operators appeared to not understand that the merge was uncertain using the cuts as a visual hint alone though, as visually showing uncertainty can improve performance (Kirschenbaum *et al.*, 2014). In the context of EID theory, the cuts should have offered a directly perceivable hint that they did not refer to the same object. A vital consideration is that this study was conducted on individuals, as opposed to teams, with the latter being how work is completed in submarine control rooms. It is plausible that the quality control performed by the Sonar Controller and Operations Officer would catch these errors, see Table 13, acting as another layer of feedback.

Following the trend for track management metrics (assignment and merging), solution performance was better in GIST. This could be explained by the difference in solution generation processes. DW opts for a Local Operations Plot that allows operators to place a solution anywhere without validation. This results in a practically infinite number of possible solutions (Cunningham and Thomas, 2005), which must be narrowed down by constraining known parameters, such as bearing and speed, to reduce available parameter permutations. This is important as cuts can match, but the solution can still be incorrect (DeAngelis and Green, 1992). GIST instead provides a drawable speedstrip that automatically calculates intersections, removing the need to perform manual alignment. The speedstrips location can be constrained by fixing parameters, and additional tools are provided to trial solutions, see Section 8.5. It is posited both factors contributed to more accurate solutions in GIST, and at the very least reduced egregious solution entry by ensuring that solutions used actual cut data, instead of being based on where the speedstrip was placed in DW. Solutions being more accurate supports the positive results of previous experiments examining fault detection (Jamieson, 2002; Reising and Sanderson, 2004; Rechard *et al.*, 2015), as participants would have been able to explore how their solutions were constrained and identify faults with how their solution had been entered based on the generated parameters.

While these factors were likely contributory to the better solution accuracy in general, GIST's automated solution entry for merged contacts is a stronger possibility. This functionality took advantage of the fact that the sensors were sufficiently offset to provide a position from triangulated data in merges. This was a simple implementation using two sets of intersections across time, and although quite accurate there is research on generating a solution using more advanced algorithms (Aidala, 1979; DeAngelis and Green, 1992; Lee *et al.*, 2008; Ince *et al.*, 2009; Genç, 2010; Annabattula *et al.*, 2015; Punchihewa *et al.*, 2022), which could improve positional performance further. Regardless of the accuracy of automated solution generation, it is vital to ensure that operators still quality-check the work to avoid automation bias, as it has been shown that humans monitor automated systems poorly, relying on warnings instead of manual checks (Mosier and Skitka, 1996; Lützhöft and Dekker, 2002). Automation checking could also allow operators to understand the doubt, if any, of automation output, potentially mitigating this automation bias (Wickens *et al.*, 2015; Man *et al.*, 2018). While improved accuracy is generally better, more accurate algorithms might yield diminishing returns past a certain point (e.g., 10- vs 1- meter fidelity) and have no perceptible effect on the decision to stay clear of other vessels by sufficient distances for safety and covertness. Instead, it might be preferable to represent areas of uncertainty, showing where vessels could, or will, possibly be, which has shown to be effective in multiple studies (Dry *et al.*, 2005). This would be aligned with the concept of EID, allowing the submarine command team to understand constraints on their positioning and navigation so that they can plan to achieve their goals.

9.4.3 Hypothesis 3. Workload would be affected by the interface and scenario difficulty

Brief Results Summary: There was not a statistically significant effect of interface, nor an interaction effect between interface and scenario, on workload. Participants rated the high difficulty scenarios more highly, signifying a higher perceived workload, when compared to the low difficulty scenarios, see Table 31 and Table 32. Thus, the hypothesis is partially supported, with scenario difficulty affecting workload.

The experiment demonstrated improved performance, although this should not be to the detriment of operator workload. Apart from consideration of an operator's working conditions, performance is affected by workload (Yan *et al.*, 2022), citing (Nachreiner, 1995). The subjective usability change is positive, with reduced workload means for GIST. The NASA-TLX results, see Table 31 and Table 32, suggest specific workload types contributed more than others. Comparing the means of scenarios between the two interfaces suggests that all workload types except for frustration, and performance for low scenarios, were improved for GIST. This suggests that any future improvements to GIST should work on removing causes of frustration. This might also be related to the inconsistency identified using the SUS, which could have been a source of frustration for participants.

A reduction in workload supports SRK literature (Rasmussen, 1983; Rasmussen and Vicente, 1989), EID literature (Vicente and Rasmussen, 1992; Nielsen, Goodrich and Ricks, 2007; Bennett and Flach, 2019), and previously conducted experiments on EID interfaces (Hall, Shattuck and Bennett, 2012; Selkowitz *et al.*, 2017; Calhoun *et al.*, 2018; Schewe and Vollrath, 2020; Michailovs *et al.*, 2021). It is thought that the workload benefits arose from adhering to guidelines in the EID literature. This includes representing the information ecologically, allowing for direct perception (Gibson, 1979; McIlroy, 2016), capitalising on human perception and psychomotor abilities (Dinadis and Vicente, 1996), transitioning of cognitive tasks to perceptual tasks so that control is at the lowest possible level (Van Dam, 2014; Cravens, 2021), and displaying both physical and functional information (Pawlak and Vicente, 1996). It is proposed that these benefits have arisen from a thorough analysis of Sonar (Chapter 5) and TMA (Chapter 6) to understand how they could be redesigned using an EID approach (Chapter 8), with changes being focused on representing as much information as possible using the map view.

The significant effect of scenario on subjective workload partially validates the scenario design, confirming that the low- and high- difficulty scenarios were perceived as such. This also confirms previous research where workload was an indicator of task performance for a track management task (Loft *et al.*, 2018). Subjective workload only being significantly affected by the scenario suggests that the number of contacts is a more important contributory factor to workload than the interface for this experiment. This is congruent with some accident reports, where workload was a casual factor (National Transportation Safety Board, 2001; Marine Accident Investigation Branch, 2016). Solutions in low scenarios being more accurate could also support this, as if participants had a lower workload, then they could have spent more time creating, evaluation, and refining each solution. As such, future work should look to further ensuring that operators are further supported in high workload situations. This would build on the features currently built into GIST, which were designed to let the operator work freely within constraints, but also to provide guidance where a mistake was known to be committed. For example, participants could merge any contacts together, but would be warned if the merge was not correct. Potentially incorrect actions were not disallowed, as the system might be incorrect. This might lead to accidents if operators cannot override the software if required (Favarò *et al.*, 2013). For example, the bearings on two trackers might be different if only one was in the baffles and its information was not being updated, but the merge would be correct. Submarine command teams are well versed in handling ambiguity, and care must be taken not to incorrectly exclude a valid conclusion that they, either as individuals or a team, might arrive at.

9.5 Conclusions

This chapter has presented a HITL study conducted to examine the effects of GIST as an EID on individual performance when conducting Sonar and TMA tasks with the aim of creating and refining a tactical picture. Three main areas of assessment were subjective usability, objective performance, and subjective workload, each of which had the associated hypothesis that they would be affected by the interface and scenario used.

EID was selected for the design of GIST because synergies were observed between the work of the command team in constructing the tactical picture, and the goal of EID in making the constraints of a system and its environment apparent to operators (Van Dam, Mulder and van Paassen, 2008; Fay, Roberts and Stanton, 2019). The study was conducted to determine whether the use of EID was an appropriate design choice to address future challenges posed. These challenges stemmed from requirements for future submarine control rooms to handle more data from more sensors (Carrigan, 2009; Henley, Schmitt and Huf, 2013; Stillion and Clark, 2015), and the potential for more screens to be introduced to do so (Chalmers, Easter and Potter, 2000). While sufficient operational capability is currently possible, these future requirements are necessitating a step change. It was posited in Section 4.2 that this step change should be to EID. EID was proposed that the introduction of EID would improve performance compared to traditional (non-EID) systems (Vicente, 2002; Bennett and Flach, 2019), increase work domain transparency (Van Dam, Mulder and van Paassen, 2008), reduce workload (Nielsen, Goodrich and Ricks, 2007), and reduce memory requirements (Lau and Jamieson, 2006). Therefore, this thesis has explored how EID could be implemented for Sonar and TMA, detailing the Cognitive Work Analyses conducted (Chapter 5 for Sonar, and Chapter 6 for TMA) and the construction of GIST (Chapter 7).

While previous research has shown that EID provides a reduction in workload and improves task completion (Borst *et al.*, 2017; Shier *et al.*, 2018; Zestic *et al.*, 2019; Cravens, 2021; Michailovs *et al.*, 2021; Michailovs *et al.*, 2022), there was a need to verify this for GIST specifically. A key aspect of designing for future ways of working in submarine control rooms is taking advantage of the ability to test ideas in facilities under robust experimental conditions. This provides early evidence for any benefits and provides an understanding of where improvements might be made should the design be taken forward. The conducted experiment has demonstrated that the benefits of EID are achieved when using GIST, confirming similar studies in the domain (Michailovs *et al.*, 2021; Michailovs *et al.*, 2022), and has demonstrated that a merged interface between Sonar and TMA is a viable option as well.

Comparing the SUS results between the two interfaces demonstrated that there were statistically significant differences in favour of GIST, and that participants rated it as generally being more usable. However, it was not rated preferably in all SUS questions, receiving a less favourable response for question 6 (consistency) and question 9 (confidence). Inconsistency can increase task completion times (Mazaeva and Bisantz, 2014), but conversely can be appropriate when there is a focus on tasks (Grudin, 1989; Bennett and Flach, 2011). Consequently, tools for specific tasks, such as solution entry, in future versions of GIST would be reviewed to ensure that are consistent while still maintaining effectiveness in their use-case(s). Inconsistency in GIST might have been related to participants not feeling as confident using it as with DW. It has been shown that consistency can aid learning (Schneider and Shiffrin, 1977; Goettl, 2006), suggesting that confidence could be improved if consistency was improved. This would be supplemented by enhanced training, giving operators a more comprehensive understanding of GIST, which might also help their confidence. A final point is that there is still room for improvement to achieve the commonly accepted score of 70 for being reasonably usable.

Each objective measure of performance also showed statistically significant improvement in GIST, supporting literature that proposes that EID will improve task performance (Torenvliet, Jamieson and Vicente, 2000; Vicente, 2002; Bennett and Flach, 2019). It is congruent with other EID research that assessed integration (Michailovs *et al.*, 2021) or consolidated display (Michailovs *et al.*, 2022). It is also congruent with studies in other domains that demonstrated EID can improve performance (Borst *et al.*, 2017; Shier *et al.*, 2018; Zestic *et al.*, 2019; Cravens, 2021). Moreover, the improved performance would have direct implications for the safety of the boat, which is a key consideration of submarine command teams (Mack, 2003). The improvement in all tasks would mean that operators had a more complete and accurate tactical picture in GIST. This is most stark when looking at the solution positioning, where entered solutions were just over 500 meters more accurate. Such large differences would undoubtably degrade ownship safety, as contacts would be perceived to be in a completely different location, which could detrimentally affect ownship safety if maneuverers were attempted using the tactical picture, creating the possibility of collisions.

Finally, subjective workload was shown to decrease when using GIST, supporting literature that states this as a benefit of EID (Dinadis and Vicente, 1996; Nielsen, Goodrich and Ricks, 2007). While there might be a contributing factor of the different aesthetics affecting the perception of workload (Tractinsky, Katz and Ikar, 2000), a decrease in workload is a positive result and could have improved performance (Yan *et al.*, 2022), citing (Nachreiner, 1995). While aesthetics are likely to have contributed to a lower subjective workload, it is a stronger possibility that the benefits arose from adhering to guidelines in the EID literature, which would have yielded more substantive workload benefits than looks alone. This appears to be the case based on the improved task performance observed. Finally, a lower routine workload level suggests it would be possible that operators would have more capacity to handle unfamiliar and/or unexpected events that might negatively affect safety.

Overall, the results of this HITL experiment have been positive and have demonstrated the benefits of GIST, and its underlying EID principles, using robust experimental and statistical rigor. This is the first step to introducing what GIST has to offer into future ways of working. The next steps would be to make changes based on lessons learned from this experiment, and then run a team study to identify how GIST affects the wider tactical picture compilation sociotechnical system.

Chapter 10 **Comparing Graphically Integrated Sonar and Target Motion Analysis to a User-Centred Design**

10.1 Introduction

A method of designing new capabilities, including interfaces, that is gaining ground for military purposes is the use of operator-centred approaches (Hamburger, Miskimens and Truver, 2011), including workshops such as Tactical Advancements for the Next Generation (Hall, 2012; Turner, 2017), or GV design workshops (Fay, Roberts and Stanton, 2020). These workshops involve operators as the final end-users to be involved in 'blue-sky' design exercises that yield design ideas that can be taken forward for prototypical development, testing, and possibly implementation. The inclusion of operators is a positive step, as User-Centred Design (UCD; Norman and Draper, 1986) methods have been shown to yield better interfaces. The inclusion of the operators also marks a shift from UCD alone to participatory design (Read *et al.*, 2015b), where all parties (domain experts, designers, researchers) are actively engaged in the design process together (Sanders and Stappers, 2008). However, Beevis, Vicente and Dinadis (1998) argued that it is not enough to adopt UCD alone. These workshops do not typically include Ecological Interface Design (EID) as a consideration, meaning that the operational environment and its constraints might not be fully considered, leaving the associated benefits of EID unexplored. While EID incorporates the user into the design process, it differs because it has a focus on the environment and its constraints (Ho, Dal Vernon and Jamieson, 2003; Kwok, 2007; Ellejmi *et al.*, 2018). It is not contradictory to a UCD approach (Ho, Dal Vernon and Jamieson, 2003), rather complimentary (Howie and Vicente, 1998; Kwok, 2007), aiming to provide support for unanticipated events, which can be omitted from an operator only perspective (Lau and Jamieson, 2006). This could result in unexploited capability gains unless EID, or at least CWA as a precursor stage, is not also considered. Thus, it may be pertinent to include EID principles in these design workshops. However, previous workshops have been successful, meaning that any proposed inclusions should show additional value to the resultant designs to warrant changes to a working process.

There is literature to support consideration of both EID and UCD, both in terms of process (Henley, Schmitt and Huf, 2013; Read *et al.*, 2015b; Read *et al.*, 2018; Revell *et al.*, 2018; Revell *et al.*, 2019), and experimental results (Ho, Dal Vernon and Jamieson, 2003; Mendoza, Angelelli and Lindgren, 2011; Varga, Winkelholz and Traeber-Burdin, 2017). Revell *et al.* (2018) proposed a method called User Centred Ecological Interface Design (UCEID) that would incorporate both approaches, being best suited to application in complex sociotechnical systems where the user plays a critical role in the interaction. The process was represented as a flowchart, building on successive stages of output to produce candidate designs for development. The general process was to conduct user research and apply thematic analysis, before conducting all CWA stages, and then use a design workshop for concept generation and development.

The Cognitive Work Analysis Design Toolkit (CWA-DT; Read *et al.*, 2015a; Read *et al.*, 2015b; Read *et al.*, 2018) is another option, with the design process being designed around the use of CWA to guide the design activities. There are multiple stages provided, although, as the name suggests, its utilization can be flexible and context-driven (Read *et al.*, 2018). The stages are groups of individual activities or concepts that can be used to generate outputs, which can then be fed forward into connected process stages. Although primarily focused on CWA, Read *et al.* (2018) noted that it had been used with other methods, and could be expanded to incorporate others in the future. This was primarily concerned with systems-based analysis methods, although UCD could be applicable to the stages provided. There is growing evidence of applications of the CWA-DT (Read *et al.*, 2016; Salmon *et al.*, 2017; Beanland *et al.*, 2018), demonstrating both its utility, but also elucidating how it could be used to run a design workshop. These options could both be utilised to run updated workshops for the design of interfaces. Although as stated previously, current user-centred workshop approaches have been shown to work, meaning that sound reasoning must be presented for change to occur.

This brief chapter explores the differences between Graphically Integrated Sonar and Target Motion Analysis (GIST) and a Mashup display, created as part of a design workshop. The aim is to compare the interfaces to identify whether an EID could offer benefits that a UCD does not, providing evidence to demonstrate that military design workshops should also incorporate EID. This comparison was performed using data from the Human in the Loop (HITL) experiment from the previous Chapter, instead of its own experiment. Due to restrictions on lab access arising from the Coronavirus 19 pandemic, all planned interface experiments were merged into one large study using a repeated measures design, which was described in Section 9.2. The hypotheses for the comparison were as follows:

Hypothesis 1. **Subjective usability would not be affected by the interface:** There would not be a difference in perceived usability, measured by the SUS between the interfaces, as both incorporated input from the user community.

Hypothesis 2. **Objective task performance would be affected by the interface and scenario difficulty:** Task performance would be affected by the interface and scenario difficulty.

Hypothesis 3. **Workload would be affected by the interface and scenario difficulty:** Workload, as measured by the NASA-TLX and Bedford Workload Scale would be affected by the interface and scenario difficulty.

10.2 Method

10.2.1 Design Workshop

The mashup interface was designed over the course of a three-day design workshop. Twenty-six participants, three facilitators, one technical partner, and one member of the author took part in the design sprint process held across three days at BAE Systems Farnborough. The participants were subject matter experts, being either operators or supervisors for Sonar and TMA. Participants were asked to provide their expertise and ideas to generate the designs. Facilitators would support syndicates of participants in completing this process providing relevant materials and facilities to do so. The technical partner answered questions about the work, including the impetus and expected outcomes, although it is important to note this was to ensure the interfaces afforded requirements rather than constrain the designs. They would also have the final say should a large decision arise or if syndicates were unable to make decisions. The author was to document the process and answer any questions about the ComTET project and the empirical data published to date. Again, this information was used to ground the concepts designed, ensuring they had all required affordances, but not to contain the designs by legacy considerations (e.g., hardware and/or software).

The GV design sprint is a process for answering questions through design, prototyping, and customer testing (Google Ventures, 2016). It aims to remove the requirement to build and deploy a new product to receive user feedback, instead relying on quick prototypes from a streamlined process to do so. This feedback can then be used to inform future design directions. Table 33 summarizes the process and the expected outputs at each stage.

Table 33 – Overview of the modified GV design sprint process

Day	Task	Time (Hours)	Output
1	Introduction	1.5	--
	Lightning Talks	1.75	Affinity maps
	How Might We		
	Affinity Mapping		
	Goal Setting	1.25	Goal and information journey
	Information Journey		

Day	Task	Time (Hours)	Output
	Boot Up Review		
	Idea Generation	1.25	Material for solution sketch
	Crazy Eights		
	Solution Sketching	1.25	Solution sketches
2	Voting	1.5	Winners and Maybe-laters
	Rumble or all-in-one	1.75	Storyboard
	Storyboarding		
	Prototyping	2.5	Draft Prototype, Prototype
3	Interview-based Testing	1.5	Interview notes
	Interviewing and Analysis	1.75	Guidance for improvement
	Presentations	2	Reviewed prototypes
	Debrief		

10.2.1.1 **Day One**

Introduction (1.5 Hours). Participants were provided with an introduction to the GV design sprint process and introduced to individuals running the event. A brief overview of the schedule was provided, and all participants were introduced. Participants were requested not to let rank seniority impact the design process, and to leave rank at the door. To facilitate this the event prescribed a casual dress code. Participants were instructed that there could be no stupid ideas and that the primary aim of the event was to promote 'outside-the-box' concept development.

Lightning Talks, How Might We, Affinity Mapping (1.75 Hours). A series of short talks were delivered to provide participants with a general overview of the purpose of the event and to clearly identify the problem space being addressed. The first talk was from a technical partner who described the future challenges that submarine control rooms will face (e.g., next-generation sensor integration) and why these must be addressed to maintain effective performance. The three design concept seeds to be created were also outlined: A mash-up display, an overlay display, and an automated sonar track management assistant. It was indicated that the integration of Artificial Intelligence (AI) was to be considered wherever suitable, regardless of whether it was believed that such capacity currently exists. Prior to the second talk, a short video was shown to participants that detailed the design sprint process and its overarching philosophy. The second talk was delivered by a submariner Officer of the Watch (OOW), who described the submarine control room from a senior officer's command and control perspective. This included information concerning how the submarine would be commanded, decisions that must be made, and the focus on safety as part of the three submarine tenets (remain safe, remain undetected, complete the mission; Mack, 2003).

The third talk provided insights into the information flow bottlenecks identified in contemporary submarine control rooms as part of the ComTET project.

The next phase saw participants split into their design syndicates (three in total) and paired with a facilitator who would guide them through the design process and help resolve any differences of opinions within syndicates. During the lightning talks, participants were asked to capture ideas on post-it notes, using a note-taking method called “How Might We’s”, which consisted of three components:

- **How:** Suppose that opportunities exist
- **Might:** The process might not find something
- **We:** All participants should be involved

Using this method participants generated ideas, one per note, to take forward. The ideas were requested to be succinct, and not too broad or narrow. Each idea was a participant’s thoughts on what could be explored and achieved, based on the lightning talks. These ideas formed the initial basis for design discussions within the syndicates.

Ideas generated as part of the “How Might We” process were stuck on a wall, with participants grouping them into categories as they went. Participants were instructed not to worry if categories were not immediately apparent, and that overlaps, or duplication were good places to look for groupings. Groupings could be changed as required to create the most useful mapping, but there was a time limit of ten minutes. Once this time had elapsed, participants were given three tokens each to vote on their preferred ideas. They could vote for their own idea and could place multiple votes on a single idea.

Goal Setting, Information Journey (1.25 Hours). Participants were guided by the facilitator to create a goal for their interface. They were instructed to keep the goal simple and could move relevant sticky notes around the whiteboard paper the goals were being written on. Each goal was written by the facilitator and discussed by the syndicate to shape its exact nature. Once overall goals were set, participants were asked to set success metrics that would facilitate an understanding of whether the interface was successful. With goal setting complete, participants designed an information journey. This described the process through which end-users would interact with the product, starting from initial engagement and carrying through to them using the product and any steps afterwards.

Boot up Review, Idea Generation, Crazy Eights, Solution Sketching (2.5 Hours). At this stage of the process, participants started sketching out their ideas. However, before doing so they reviewed their progress as a group to produce a list of ideas they would like to develop further as drawings.

The first step was to sketch as many initial ideas as necessary and to circle their best ideas. After this had been completed, each participant was provided with a sheet of paper folded into eight sections. They were asked to create variations of their best ideas in each of the frames. They did not have to generate eight ideas; however, they were asked to develop at least two. Once this stage had been completed, participants selected their favourite idea and drew it on paper to illustrate it in detail. Participants were instructed to make the sketches self-explanatory but were told that ‘ugly’ was okay, as the process should not be hindered by a lack of artistic talent. Each solution was to be given an encompassing, easy to remember title.

10.2.1.2 **Day Two**

Voting (1.5 Hours). Participant sketches from the previous day were stuck to a wall. When participants arrived, they viewed the sketches and were given three stickers each to stick to aspects that they liked, to form a heat map. Once participants had had a chance to view the sketches and vote on the aspects that they liked they entered the speed critique phase. As a group, they discussed each sketch and captured standout ideas and important objections on post-it notes, that were subsequently stuck to the design. After the critique, the sketcher was requested to present their design and could highlight anything the group had missed. Requiring sketchers to present at the end of the critique was important as it prevented stifled feedback stemming from feelings towards an individual. This enabled honest feedback and a reduction of ownership biases. Finally, each participant was provided with a big dot sticker to vote for their favourite idea.

Rumble or all-in-one, Storyboarding, Prototype (1.75 Hours). Should multiple sketches have been voted to a similar degree, the syndicate members were asked to either pick one or create a composite design. Each syndicate opted to include elements of highly voted interfaces into an idea that would be taken forward. Other interface aspects were also chosen as ‘maybe-laters’ that could be incorporated into future designs should they be required. Once a design idea was chosen, participants storyboarded their idea using whiteboard sheets mounted to a wall. Through this process, participants were requested to consider the information journey they had designed on day 1. The storyboard did not have to be perfect, but complete enough to illustrate the intention and state of the interface at particular points in time. This would facilitate participants taking ideas forward and prototyping them. At the prototyping stage, participants were informed that the development should be realistic enough to test, but that it did not need to be in a high state of maturity (i.e., for operational deployment). They were informed that the prototype could be created using tools of their choice, such as interactive presentations, drawings, prototyping software, or developing an example application. To facilitate effective construction of each prototype, roles were assigned to each syndicate member, such as assembly, finding resources, and ensuring a consistent user experience. Facilitators helped push the prototypes forward and ensured participants had all required resources.

10.2.1.3 **Day Three**

Interview-based Testing, Interviewing and Analysis (3.25 Hours). Once the prototypes were complete, an interviewer was nominated from each syndicate to lead interviews of participants from other syndicates. To facilitate the interview process, a short presentation was created to highlight the key features of the prototype. Interviews followed a schedule created by facilitators, which allowed members of each syndicate to visit all other syndicates as interviewees. While each interview was with one individual, multiple syndicate members could join. An interview was also held with the technical partner to keep them abreast of the designs. In each interview, a demonstration of the syndicate's prototype was provided by the interviewer. Afterwards, the interviewee could ask questions and provide feedback. The interviewer would also elicit feedback throughout. Other members of the syndicate would record any pertinent points using sticky notes, organising them using a matrix on a wall or whiteboard (areas of interest as rows, interviewees as columns). Once all interviews were completed, syndicates consolidated feedback into consistent themes and groupings.

Presentations, Debrief (2 Hours). With the process completed and prototype feedback collected, each interface was presented to project stakeholders by the designing syndicate. Each syndicate prepared a presentation, with time allocated for feedback and questions. There was a general discussion about each interface and the practicalities of implementing them operationally. The interfaces typically included artificial intelligence and futuristic technology, therefore it was important to understand what capabilities would be required to actualise the designs and what was 'blue-sky thinking'. This facilitated an understanding of whether the proof-of-concept designs could be developed as proof-of-concept software in a suitable timeframe for testing in ComTET. Finally, the lead facilitator provided a debrief of the process, and thanked everyone for their time, concluding the design sprint event.

10.2.1.4 **Resulting Prototype**

The design sprint process was successful and yielded three prototypes that could be taken forward to create proof-of-concept designs to undertake formal testing of their effectiveness compared to a contemporary baseline. Of these, a display known as a 'Mashup' was selected to take forward for testing, and will be described below.

The mashup display was an interface that combined information from a variety of sensors and sources, utilizing artificial intelligence to assist with classification and detection. The overview screen in Figure 77 would allow operators to see a combined view of sensor (Sonar, Electronic Warfare, Periscope, etc) data against time on the left, and tactical information on the right. A playback of received sonar signals was also available. The artificial intelligence could learn and flag previously seen signals to operators, along with a classification. Learning would be facilitated by information from different sensors being corroborated to provide a comprehensive signal profile. By including extensive artificial intelligence capabilities, it was posited that a reduction in operator numbers would be possible and any remaining operators would be placed into a common warfare branch. These operators would be supported by the AI performing tracking and classification in busy environments, but they retained the ability to veto any decisions or actions it made.

Operators could see contact-specific data by selecting it from the overview, which would open the contact mashup page, see Figure 78. The left-hand side shows sonar and underwater data. The right-hand side shows a tactical picture overview and Automatic Identification System (AIS) data. Customisation was possible to facilitate operator work as much as possible, and predefined customisation options could be accessed using the buttons in the bottom left, which would configure the screen for common scenarios.

The idea for the prototype was communicated to an industry company, who implemented the interface ready for experimentation using the ComTET facility, see Figure 79. As with GIST in Chapter 8, the finished product differed from the initial designs due to various factors, such as time and feasibility. The finished Mashup display consisted of several screens separated into tabs. Each displayed a selection of data that was combined from all Sonar sensors, meaning that participants did not have to perform any merges as these would be completed automatically. Solutions were entered using a calculator that accepted known solution parameters and used them to calculate unknown parameters. These parameters could then be shared as a solution that appeared on the waterfall, showing what the track would be doing if the solution was accurate, meaning that it had to align with the underlying data.

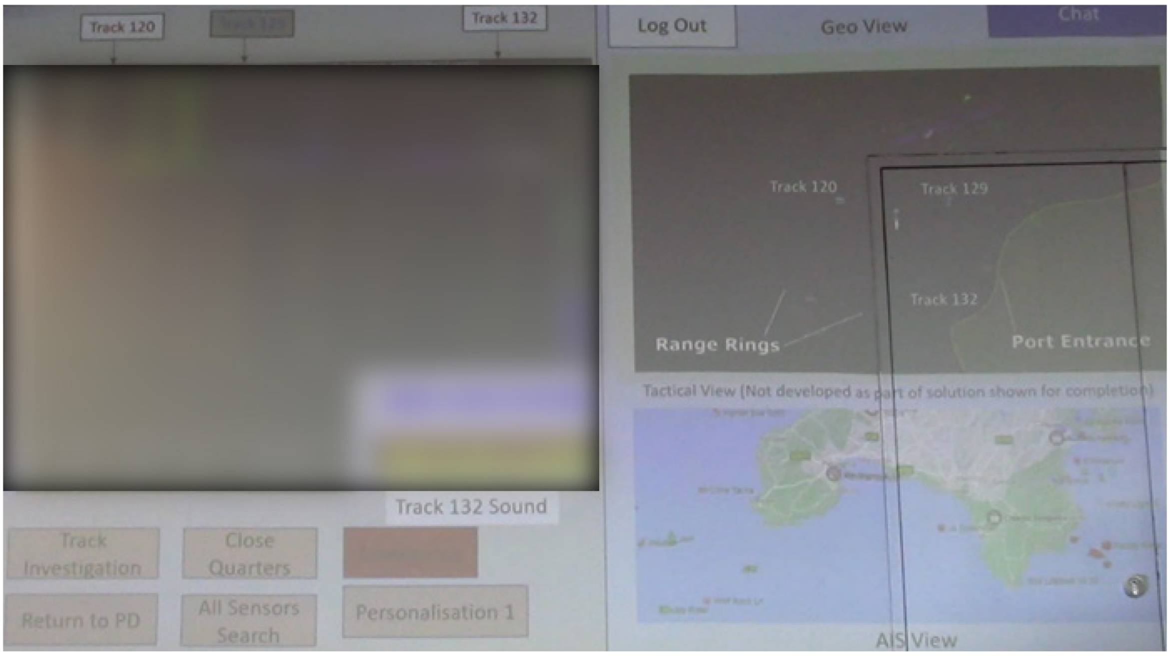


Figure 77 – A screenshot of the mashup display overview screen

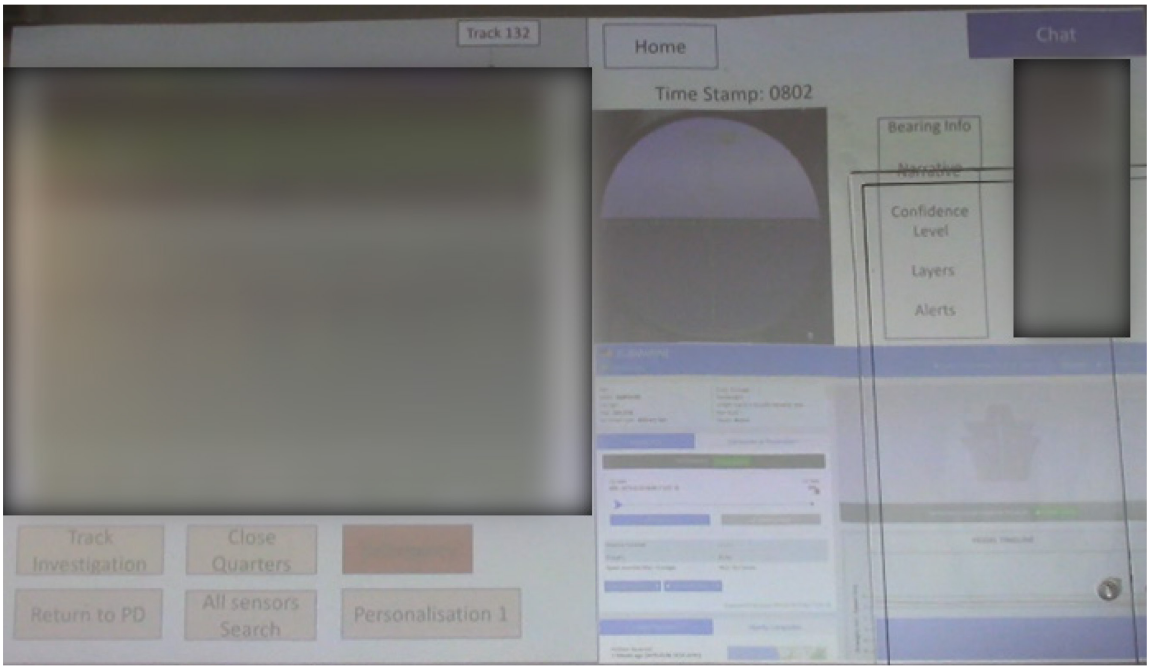


Figure 78 – A screenshot of the mashup display mashup screen

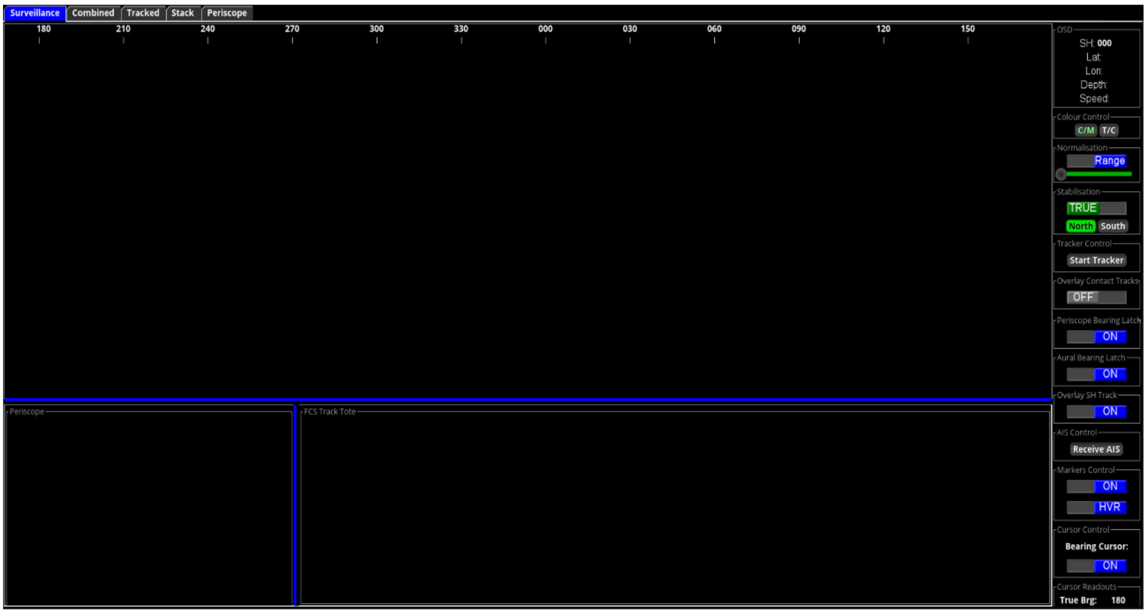


Figure 79 – A screenshot of the implemented mashup display

10.2.2 Experiment

As mentioned in the introduction, the data for this thesis was gathered in a combined repeated measures study owing to the Coronavirus 19 pandemic. The method for this Chapter remains mostly the same as described in Section 9.2, although merge performance was not evaluated as the Mashup automatically completed this, meaning that there were no merges to compare against.

10.3 Results

10.3.1 Subjective Usability – System Usability Scale

Participants rated the Mashup (M = 56.83, SD = 21.66) slightly higher than GIST (M = 54.67, SD = 23.11) for usability. The difference was not statistically significant, t (44) = -2.17, p = .66 (2-sided). The means and standard deviations for each question, and the total scores, are presented in Table 34.

Table 34 – Averages for each SUS question and the total score by interface, with cell shading representing a favourable rating for each question, based on the mean.

	GIST		Mashup	
	M	SD	M	SD
Q1. I think that I would like to use this system frequently	2.98	1.26	3.02	1.18
Q2. I found the system unnecessarily complex	2.60	1.14	2.69	1.23
Q3. I thought the system was easy to use	3.27	1.29	3.36	1.18

Q4. I think that I would need the support of a technical person to be able to use this system	2.96	1.28	2.78	1.19
Q5. I found the various functions in this system were well integrated	3.47	1.20	3.60	1.00
Q6. I thought there was too much inconsistency in this system	3.00	1.15	2.64	1.16
Q7. I would imagine that most people would learn to use this system very quickly	3.42	1.24	3.42	1.20
Q8. I found the system very cumbersome to use	2.64	1.16	2.78	1.13
Q9. I felt very confident using the system	2.78	1.28	3.27	1.18
Q10. I needed to learn a lot of things before I could get going with this system	2.84	1.25	3.04	1.09
Overall SUS Score	54.67	22.85	56.83	21.42

10.3.2 Tracker Assignment – % Trackers Assigned

Participants assigned more trackers when using the Mashup. The percentage of trackers assigned was significantly affected by the interface used, $F(1, 35) = 33.38$, $p < .001$. It was revealed that participants assigned more trackers in the Mashup than they did using GIST, as shown in Table 35 and Figure 80. The mean percent of trackers assigned was 90.8% (SE = .79, SD = 4.72) in GIST and 97.67% (SE = .93, SD = 5.57) in the Mashup. There was not a significant effect of scenario $F(1, 35) = .604$, $p = .44$. Tracker assignment was higher when using the Mashup, across both scenarios completed. There was a significant interaction effect between the interface and scenario, $F(1, 35) = 22.680$, $p < .001$.

Table 35 – Tracker percentage assignment between GIST and the Mashup for both scenarios

					95% Confidence Interval	
Interface	**Scenario**	**Mean**	**Std. Error**	**Std. Dev.**	**Lower Bound**	**Upper Bound**
GIST	Low	87.5	.0	.0	87.5	87.5
	High	94.1	1.57	9.55	90.9	97.37
Mashup	Low	100	.0	.0	100	100
	High	95.35	1.86	11.62	91.58	99.12

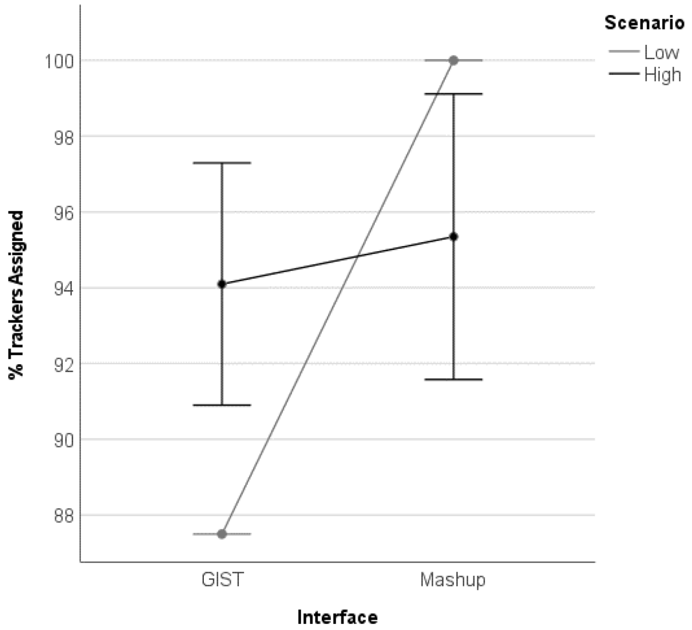


Figure 80 – Tracker percentage assignment between GIST and the Mashup for both scenarios. 95% Confidence Intervals (CI) indicated.

10.3.3 Solution Performance – Best Solution Position Δ & Best Dead-Reckoned Position Δ

Participants entered more accurate solutions when using GIST. The MANOVA revealed that solution performance was significantly affected by the interface, F (2, 15) = 18.69, V = .714, p < .001. The means for each metric are shown in Table 36, Figure 81, and Figure 82. Subsequent univariate tests revealed that there were statistically significant effects of the interface used for both the entered solution positions, F (1, 16) = 27.614, p < .001, and for the dead-reckoned solution positions, F (1, 16) = 24.19, p < .001. The most accurate solution position had a mean of 36.48 meters (SE = 1.55, SD = 6.39) in GIST and 537.03 meters (SE = 1.33, SE = 5.48) in the Mashup. The dead-reckoned solutions were also more accurate in GIST (M = 38.55, SE = 1.48, SD = 6.1) than in the Mashup (M = 601.17, SE = 1.3, SE = 5.36). Solution performance was not affected by the scenario used, F (2, 15)

= .72, V = .09, p = .502, nor was it affected by an interaction between the interface and scenario used, F (2, 15) = .21, V = .03, p = .809.

Table 36 – Solution Performance for GIST and the Mashup

						95% Confidence Interval	
Measure	**Interface**	**Scenario**	**Mean**	**Std. Error**	**Std. Dev.**	**Lower Bound**	**Upper Bound**
Best Solution Position Δ	GIST	Low	40.49	1.79	7.39	11.74	139.6
		High	32.79	1.67	6.88	11.06	97.22
	Mashup	Low	796.03	1.39	5.73	395.79	1,601.01
		High	362.49	1.54	6.35	145.05	905.9
Best Dead-Reckoned Position Δ	GIST	Low	41.77	1.77	7.29	12.48	139.81
		High	35.63	1.62	6.67	12.87	98.67
	Mashup	Low	872.68	1.35	5.55	464.13	1,640.83
		High	414.43	1.47	6.05	184.03	933.32

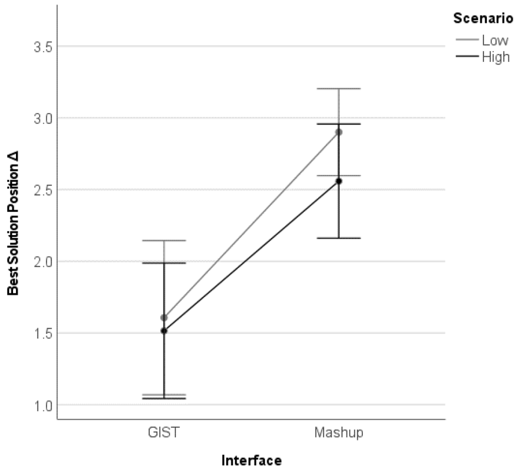


Figure 81 – Best Solution Position Δ for GIST and the Mashup. 95% CI indicated.

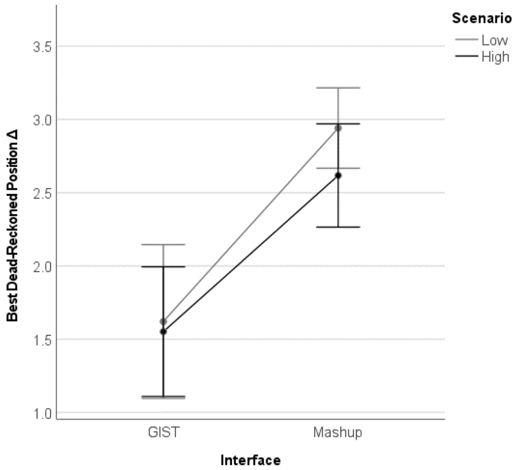


Figure 82 – Best Dead-Reckoned Position Δ for GIST and the Mashup. 95% CI indicated.

The results of the above were confirmed using further ANOVA tests as only 17 participants had completed all permutations. Testing with scenario as the only independent variable confirmed that GIST was not affected by scenario, F (2, 19) = .01, p = .928. This was also confirmed for the Mashup, F (2, 28) = 2.75, p = .081.

Testing with interface as the only independent variable revealed that there was a significant effect of interface for the low scenarios, F (2, 24) = 16.01, p < .001. Subsequent testing showed that there was a significant effect of both entered position, F (1, 25) = 27.92, p < .001, and dead-reckoned position, F (1, 25) = 33.03, p < .001, with solutions being more accurate in GIST. High scenarios were not significant, F (1, 18) = 5.09, p = .018.

10.3.4 Subjective Workload – NASA-TLX & Bedford

The Mashup and low difficulty scenarios had lower subjective workload. The MANOVA revealed that subjective workload was significantly affected by the interface used, F (2, 30) = 6.54, V = .30, p = .004. Subsequent univariate tests showed that there was not a significant effect of the NASA-TLX, F (1, 31) = 1.10, p = .30, but there was a significant effect of the Bedford, F (1, 31) = 10.72, p = .003. The MANOVA also revealed a significant effect from the scenario used, F (1, 30), V = .729, p < .001. Subsequent univariate tests showed that there were significant effects for both the NASA-TLX, F (1, 31) = 62.14, p < .001, and the Bedford, F (1, 31) = 49.26, p < .001. There was not a significant interaction between the interface and scenario used, F (2, 30) = .02, V = .001, p = .98.

Table 37 – Workload scores for GIST and the Mashup

						95% Confidence Interval	
Measure	**Interface**	**Scenario**	**Mean**	**Std. Error**	**Std. Dev.**	**Lower Bound**	**Upper Bound**
TLX	GIST	Low	48.75	2.44	13.80	43.77	53.73
		High	63.81	3.20	18.10	57.28	70.35
	Mashup	Low	45.5	2.38	13.46	40.65	50.35
		High	61.16	3.26	18.44	54.52	67.8
Bedford	GIST	Low	5.23	.53	3.00	4.16	6.31
		High	7.66	.48	2.72	6.71	8.64
	Mashup	Low	3.69	.33	1.87	3.01	4.36
		High	6.27	.46	2.60	5.33	7.20

Table 38 – Means for each question of NASA-TLX

	GIST				Mashup			
	Low		High		Low		High	
	M	SD	M	SD	M	SD	M	SD
Mental Demand	8.97	5.02	12.54	5.23	7.55	4.09	11.66	5.07
Physical Demand	2.26	2.69	2.95	4.13	2.75	3.60	3.93	5.36
Temporal Demand	6.82	4.65	13.10	5.73	6.16	4.01	11.46	5.57
Performance	10.18	6.18	9.05	6.28	15.59	2.49	10.54	4.61
Effort	8.79	5.26	12.10	5.09	8.95	4.26	11.93	4.99
Frustration	10.82	5.74	11.97	5.29	5.57	4.04	10.63	5.25
Total	47.85	14.29	61.72	17.73	46.57	13.42	60.15	18.78

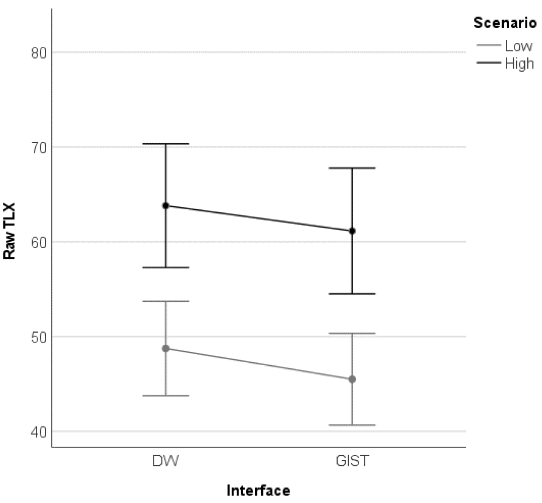


Figure 83 – TLX scores for GIST and the Mashup. 95% CI indicated.

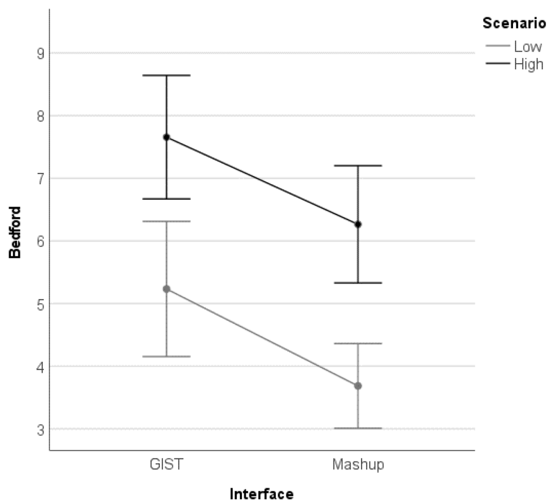


Figure 84 – Bedford scores for GIST and the Mashup. 95% CI indicated.

10.4 Discussion

10.4.1 Hypothesis 1. Subjective usability would not be affected by the interface

Brief Results Summary: Participants rated the Mashup as being more usable than GIST on the SUS, although the difference was not statistically significant, see Table 34. This confirms the hypothesis.

Both interfaces receiving similar usability scores is reasonable, given that they were both designed with user input, and therefore would reflect what they would consider to be usable in an interface.

While the Mashup display was rated higher, the difference was not statistically significant, confirming previous experiments that have assessed the difference in usability between an EID and a display designed around the existing human-centred guidance (Young and Birrell, 2012). This suggests that EID could be used in design workshops without sacrificing usability (i.e., not focusing solely on the user as with UCD can still produce usable designs). However, this may also be evidence that a common design choice between the two interfaces, such as holistic integration of system functionality (Hall, 2012), is the actual causal factor for the interface usability scores over the method used to design them. Young and Birrell (2012) also anecdotally noted that participants could take some time to get used to the EID, and that usability would surpass the traditional display once the learning curve was completed. However, this could also be true for the Mashup display, as the novice participants had typically not encountered or used Sonar and TMA displays before. This is supported by Anokhin, Ivkin and Dorokhovich (2018) testing an EID interface for a nuclear power plant control room, where they confirmed that the interface and its components were intuitively obvious to operators without additional explanation to operators. This suggests that if participants in the current study were more familiar with the underlying concepts, then they might have found the interface more understandable, and therefore more usable. However, this does not suggest that GIST would be ‘natural’ or ‘intuitive’ in all situations, requiring minimal training, which is a common misconception regarding EID (Borst, Flach and Ellerbroek, 2015). Rather the suggestion is that familiarity with a domain could greatly aid in the interpretation of a domain-specific interface (Borst, Flach and Ellerbroek, 2015). A final possible factor is that because the experiment had two interfaces sharing similar interaction mechanisms (DW and the Mashup), this might have expediated user familiarisation, and therefore increased their perception of usability.

The results show that the Mashup was preferably rated in most (N = 6; plus one equally rated) questions and overall, suggesting that concentrated and more explicit involvement of the users can lead to a higher usability score. This study was conducted with mostly novice users, which might have affected usability. Grier (2013) noted in their discussion of the SUS that military users have specific training to be able to use interfaces for their job, and that operators tend to have an “I can make it work” mentality, compared to civilian technology, which is designed to be operable with as little training as possible. Consequently, it would be of interest for future work to evaluate using an expert-only cohort to determine if their training has any effect on the usability scores of both interfaces. For example, the novice participants in this experiment would have had to learn about waterfalls, the prevailing mechanism for representing Sonar data in the Mashup display, whereas expert users would be used to working with waterfalls as part of their job and might have found them more usable than the map view of GIST, a completely different representation. The potential for differences between subjective usability ratings between novices and experts is supported by experiments that have shown the latter can pick up EID displays more easily (Borst, Flach and Ellerbroek, 2015), citing Jamieson (2007) and Burns *et al.* (2008). Another dimension to explore could be between different skill levels of experts, as those with more experience might rate waterfall-based displays as easier to use as they can apply their knowledge easier than another display that requires them to expend more cognitive effort translating the representation shown to something which they can act on. Their interpretations of the current state of the environment might also be hampered as they cannot use a 'library' of saved Sonar waterfall patterns in their long-term memory to facilitate quick memory recall (Vicente and Wang, 1998; Vicente, 2001).

10.4.2 Hypothesis 2. Objective task performance would be affected by the interface and scenario difficulty

Brief Results Summary – Tracker Assignment: Participants assigned more trackers when using the Mashup, with there being a significant effect of interface, and an interaction effect between interface and scenario, see Table 35 and Figure 80. This confirms the hypothesis.

Brief Results Summary – Solution Accuracy: Solutions were more accurate in GIST, as shown in Table 36, Figure 81, and Figure 82. Performance was significantly affected by the interface used, which also had significant effects on both components of accuracy (solution as entered, and dead-reckoned). Solutions entered in GIST were also more accurate for both low and high difficulty scenarios. This confirms the hypothesis.

Work in a submarine control room is explicitly tied to the environment in which it is operating, and the command team's accuracy is fundamentally tied to this. The level of uncertainty arising from observing the environment predominantly via Sonar (Kirschenbaum, 2001; Roberts, Stanton and Fay, 2018) means that some tasks, such as range or solutions, can be completed within specified constraints and still be incorrect. This contrasts with other domains, such as nuclear power plant control rooms, where maintaining known, measurable, parameters within specified bounds can be sufficient to maintain correct performance across a vast majority of tasks. Instead, the command team relies on a process of iterative refinement, working to constraint the solution space bounded by constraints (Michailovs *et al.*, 2021), to understand when their tasks have been completed adequately. Consequently, while a UCD approach to submarine HMI design may improve the user experience and completion rate of tasks such as contact designation, it does not reflect the awareness of constraint-driven refinement that is inherent to submarine control room work. It is proposed that this is a core explanation of why improved solutions were entered with GIST, as it provided participants with explicit representations of evolving constraints (e.g., the map view for spatial plotting, or contact widgets arranging themselves via bearing) to situate their tasks within.

This supposition appears to have been reflected in the results, with the interfaces showing stronger performance in the task related to their design underpinning.

Tracker assignment is a task-based process, with operators being required to follow steps to request that the system follow a signal whenever something of interest is detected, and to ensure that trackers are added for all sensors it is available on. The contacts from these trackers are then required to be merged to unify the information. The Mashup removed the need operators for operators to achieve these tasks, automatically utilising data from all sensors for tracking. This is likely the cause for the higher tracker assignment percentage in the Mashup could be that participants only had to designate one tracker per contact, instead of assigning two trackers per contact (one per sensor) in GIST. This meant that better performance was easier in the Mashup as fewer actions were required to achieve the same result, a user-centred design decision to make.

Future versions of GIST could incorporate user-centred ideas to improve task performance, adding functionality that aids operators to complete their tasks, and facilitating general workflows. While there was an element of this in GIST, the primary focus of EID is on the environment (Ho, Dal Vernon and Jamieson, 2003; Kwok, 2007; Ellejmi *et al.*, 2018), instead of concentrating on specific tasks that would need to be optimised with the user in mind. This could be addressed in future workshops aimed at exploring how expert operators use GIST, and where improvements could be made to their processes, addressing an identified need to improve usability from Section 9.4.1. For example, whenever a tracker is assigned on one sensor, corresponding trackers are automatically added on all other sensors if the contact is also present. This type of thinking could be applied to other workshops, creating a completely new interface, and then investigating how it could be used, and how usage could be optimised.

In contrast with tracker assignment, solution generation is a process heavily driven by environmental constraints, requiring operators to consider these to construct an accurate tactical picture. There is a task-based component in that operators should be periodically reviewing all solutions, although this only addresses scheduling, not accuracy. As operators discover or generate more information about contacts, this is used to constrain their solutions from near-infinite possibilities to those within an acceptable area of uncertainty; they are literally constraining their solution space on the map. This was supported in GIST by displaying all information on the map, supporting visuospatial resolution of bounding constraints, such as the trial mode overlay (Figure 63 and Figure 64). This contrasts with the contemporary TMA interface, where the cuts are represented using cartesian coordinate space, but are displayed without a map background, other contacts, and other functional Information. Consequently, operators can reason visually about whether a speedstrip matches the provided cuts, but are still required to change screens, and/or perform tasks in their head (i.e., transitioning speed from the Sonar screen to enter into TMA), to reason fully about solutions.

This transmutation from a predominantly cognitive task to a perceptual task, allowing operators to work at the lowest possible level of cognitive control, is a key goal of EID (Van Dam, 2014; Cravens, 2021). Solution performance being better in GIST supports the core EID premise that representing the constraints of an environment can lead to better performance (Torenvliet, Jamieson and Vicente, 2000; Vicente, 2002; Bennett and Flach, 2019). While there is limited literature in the maritime domain, EID exhibiting better performance than a UCD is supported by a driving experiment that resulted in improved performance braking performance in a braking task (Harre and Lüdtke, 2018; Harre, 2019). It is proposed the visualisation of Ownship's environment as a map view in GIST was a key factor for the improved solution performance, as this was the biggest differentiator in implementation from the Mashup display. Additionally, GIST implemented solution trialling and entry tools (see Section 8.5) that were designed around the map view. This was complimented by automation solution algorithms that could enter reasonable solutions for discrete contacts, and highly accurate solutions for merged contacts. The Mashup display also had trialling tools, calculating unknown parameters from known parameters, and showed operator's the resultant solution overlaid onto the waterfall. However, the waterfall does not represent range and there was not a predictive waterfall to show future solution state. Additionally, predictive information about a solution's location until the next cut was expected was shown on the map as a line extending from the solution icon. Adding this information could have been beneficial, as including predictive information to an interface has been shown to improve performance over an observation or checklist approach (Agnisarman, Madathil and Bertrand, 2019). For GIST, participants would have been able to evaluate solution accuracy by evaluating the line, and whether it was possible (i.e., did not show the contact as traversing the globe in 30 seconds) and congruent with what they wanted to enter (i.e., *"This should move more over the time period, so it should be edited"*).

Additionally, a lack of predictive information would have limited perception of functional information about the solution, which may have reduced performance. This is supported by previous studies that have found showing both physical and functional information yields better performance than either alone (Christoffersen *et al.*, 1994; Bennett and Flach, 2019). GIST's provision of functional information, supporting purposeful (functional) action (Lintern, 2006), would have given participants more information to understand the accuracy of their solutions, as they could have perceived the accuracy in the environment using the map display. Finally, the note-based messages provided by the automation in GIST could have assisted the operator in determining when to make changes to the solution. Whenever the automation was asked to enter a solution, it output a note that contained the results of its evaluation, see Figure 68 for examples. This provided additional functional information regarding how certain the automation was, and whether there were any components that would need manually investigating, such as confirming a range for contacts that were not part of a merge. While the design workshop included creating tools that would help create solutions in the Mashup display, better performance for solution entry in GIST implies that environmental representations must be considered to ensure that any designs optimally factor in both the user and their work environment.

In summary, mixed performance between the interfaces suggests that each had its own strengths and weaknesses for facilitating tasks, correlating with the focus that either had during their design. The Mashup's improved performance in the tracker assignment task appears to be reflective of its focus on supporting operators in completing their expected workflows and making this as efficient as possible. While it also introduced new capability for solution creation over the contemporary DW TMA interface, participants performed better in GIST, where they were provided with an ecological representation of their environment in the form of a map. Both interfaces improving a task each is a positive result, although this also means that they could have performed better at the other task. As the tasks are iterative and inter-related, future interfaces will be required to complete all tasks using the best available principles for capability design. Therefore, future versions of Sonar and TMA interfaces, designed in workshops or otherwise, should seek to capitalise on the benefits each method brings to tasks in their forte by merging them; the user-centred approach would ensure that operator requirements are accounted for, and an ecological approach would ensure that the environment can be readily perceived.

10.4.3 Hypothesis 3. Workload would be affected by the interface and scenario difficulty

Brief Results Summary: There was a statistically significant effect of interface, with the Mashup being rated better across all scenarios, see Table 37, Figure 83, and Figure 84. Subsequent univariate testing showed that there was only a significant effect for the Bedford. Participants rated the high difficulty scenarios more highly, signifying a higher perceived workload, when compared to the low difficulty scenarios. This confirms the hypothesis.

In the context of EID, the Mashup display could be considered a traditional interface as it did not explicitly also include functional information (Bennett and Flach, 2019). The Mashup display having lower workload scores contradicts literature that suggests workload would decrease (Rasmussen, 1983; Rasmussen and Vicente, 1989; Vicente and Rasmussen, 1992; Nielsen, Goodrich and Ricks, 2007; Bennett and Flach, 2019). This also contradicts prior experimental results that have shown EID has a lower workload than UCD (Wu *et al.*, 2016), or contemporary non-EID designs (Hall, Shattuck and Bennett, 2012; Selkowitz *et al.*, 2017; Calhoun *et al.*, 2018; Michailovs *et al.*, 2021; King, Read and Salmon, 2022). One possible explanation for this could be that participants required more familiarisation time (Young and Birrell, 2012), although this does not hold much weight by itself as an anecdotal observation. This could be substantiated if operators were mainly operating using Knowledge-based Behaviour on the Skills, Rules, and Knowledge Taxonomy, as this would require more cognitive workload (Ellejmi *et al.*, 2018). This might have been the case, as solution entry in the Mashup display required operators to determine known solution factors using provided tools and use these to calculate the full solution, which is a series of activities that could be completed at lower levels of the taxonomy.

In contrast, while GIST solution entry allowed for direct perception of the solution entry and automated the entry process in parts, a requirement to critically assess and transform visual data could have made the task cognitive instead of perceptual, raising workload (Van Dam, 2014; Cravens, 2021). This suggests that while the ecological map interaction in GIST improved task performance, providing operators with other non-EID methods to work with information could be beneficial. For example, the GIST solution trialling tool, see Figure 63 and Figure 64 in Section 8.5, made the trialling a perceptual task as users could determine if the solution was correct by comparing lines. However, less of a focus was given to how the user might achieve the task of identifying the component (bearing, course, range, speed) numbers to change and how much by to move the trial lines around. While design workshops could solely focus on the user-centred aspect to reduce workload for operators, it would be disadvantageous to not factor in how task performance could be improved using EID as well, given other literature that supports and evidences the workload benefits.

10.4.4 Recommendations for Approach Combination

The work of Read *et al.* (2018) and Revell *et al.* (2018) has demonstrated that multi-method design approaches can be defined and utilised. The results of this experiment offered an opportunity to derive recommendations for doing so, that will contribute to the discourse surrounding the integration of EID and UCD.

10.4.4.1 **Signposting**

A misconception regarding EID is that the resultant displays are inherently natural and/or initiative, requiring minimal training (Borst, Flach and Ellerbroek, 2015). They argue that Ashby's Law of Requisite Variety (Ashby, 1956) is a fundamental principle to EID's consideration of ecology (Vicente, 1991a), which necessitates that effective interfaces be as complex as the domain under consideration. It is proposed that this can be a detrimental factor to usability for complex domains, and this is corroborated by the GIST being rated lower in the SUS than the Mashup interface, designed around simplifying operator tasks. Expert operator participation might have increased perceived usability as they are very familiar with the domain (Borst, Flach and Ellerbroek, 2015). However, as even experienced users might not be familiar with the entire work domain and/or be abstracted from its underpinning constraints (Burns and Hajdukiewicz, 2004), representing the full work domain might have also detrimentally affected usability. The focus of UCD on effective usability for users (Abras, Maloney-Krichmar and Preece, 2004) could prove effective to address this, implementing mechanisms that provide requisite simplicity.

Considering the work domain as an area that can be operated within, UCD could be used to provide "signposts" that guide users between stable conditions (i.e., between goals, or states where no constraints are broken) in EID interfaces. If common tasks are known, then they can be added as guided processes, similar to "wizards" in some complex software. Signpost identification should be an explicit consideration of any consideration of any EID design process. Design directions could be based on whether users understand the system's state and know their action plan, prompting designers to focus on a particular area, see Table 39. Ensuring that they can navigate the interface in action terms might also help their response to events requiring time-constrained responses, as the system would guide them. For example, if a contact is identified as being too close to Ownship, a popup titled "Trigger safety procedures" could appear and prompt the user to complete steps associated with this.

Table 39 – EID and UCD Focus Matrix

	System Status Understood	**System Status Misunderstood**
Action Plan Known	*Trust the user(s), but verify through testing to unsure understanding and actions are both appropriate.*	Refine EID aspects so status is understood, as misunderstandings could lead to incorrect action plans.
Action Plan Unknown	Refine UCD aspects to facilitate user discovering or being signposted to what actions they could take based on the status.	Refine EID aspects so status is understood. Refine UCD aspects to facilitate user discovering or being signposted to what actions they could take based on the status.

10.4.4.2 **Software Support**

A pervasive criticism of the methods that could be used to inform EID and UCD is that they are time-consuming (Stanton and Harvey, 2017; Stanton and Roberts, 2018; Revell *et al.*, 2019; van Velsen, Ludden and Grünloh, 2022). Stanton and Harvey (2017) further note that the data collection methods can be potentially disruptive to the system under investigation. This has been mitigated by the creation of software tools to support analysis, such as CWA tools (Jenkins *et al.*, 2007; Hingu *et al.*, 2017), and running human in the loop studies, such as Clamshell (Fay, Stanton and Roberts, 2018). However, there is still a large amount of data collection and processing required, which might be a prohibitive factor in applying both methods as a cohesive whole. No matter how well-crafted proposed approaches are, if the requisites are excessively onerous on a project's resource constraints, they will not be applied. This is particularly important in this context, where a joint application of two methods is being advocated for, as either could be selected, and applied individually.

Therefore, rather than proposing another method, it is recommended that the data collection aspects should be focused on for future work in combining the approaches. Similar work has already been explored by work in the ComTET project, where tools were delivered to automate analyses being conducted where possible. Future work could explore generalising these tools and making them available to others to accelerate their analysis activities. This could significantly reduce the resources required to complete analyses, potentially increasing the range of design methods that could be employed. Additionally, this type of software could facilitate asynchronous data collection from subject matter experts, increasing their participation by allowing them to provide input when convenient to them, as opposed to at arranged timeslots. It is acknowledged that this would likely incur extra effort on an analyst's part for refining and integrating responses, although it is argued that this would be acceptable, within limits, to achieve higher response rates.

Another related avenue of exploration is the extension and instrumentation of interfaces, both for testing and production (deployed and in-use) capability, to automate performance and feedback data. This could be implemented easier than in the past as there has been a shift in maritime platform development to move to modular, extensible, and open capability (Sea, 2005; Scott, 2006; Hobson, 2008; Scott *et al.*, 2011; Smith *et al.*, 2013; BAE Systems, 2015). Such changes are likely to have been implemented in other domains as well, owing to the field of software engineering driving these changes. By integrating human factors data collection directly into systems, it would become possible to collect vast amounts of data to understand them for insight generation, and done correctly, it would not interfere with routine operation, an issue identified by Stanton and Harvey (2017). Should direct integration not be possible, another approach could be the use of transformative tools that take existing log outputs and extract relevant data.

10.5 Conclusion

This Chapter has presented the results of a HITL experiment with the aim of examining how interfaces created using EID and UCD approaches differed to determine if there were any potential benefits for inclusion of ecological design principles to the user-centred workshops that are being utilised for the design of new military interfaces. While UCD has benefits for interface design (Norman and Draper, 1986), Beevis, Vicente and Dinadis (1998) argued that it is not enough to adopt UCD in isolation. This is pertinent to the workshops being conducted, as UCD does not explicitly consider the work environment and its constraints as EID does (Ho, Dal Vernon and Jamieson, 2003; Kwok, 2007; Ellejmi *et al.*, 2018). Workshops could be changed or adapted to incorporate joint processes (e.g., UCEID (Revell *et al.*, 2018; Revell *et al.*, 2019) and the CWA-DT (Read *et al.*, 2015b; Read *et al.*, 2018)).

Considering both design approaches was supported by the results of the experiment in multiple ways. Firstly, subjective usability was rated quite closely between the two interfaces. This indicates that there would be no usability losses by not focussing completely on the operators. Perceived usability could change over time for GIST as an EID interface (Young and Birrell, 2012), although this was anecdotal, and could also be true for the Mashup, as novice participants had not encountered Sonar and TMA displays in any format before.

Secondly, each interface had better performance in a specific task related to their design, and the task could have been improved utilising a combined approach. The Mashup performed better for the tracker assignment task, indicating that while GIST focused on a representation of the work environment to improve performance, further consideration of functionality with a user-oriented focus could improve performance in some tasks. Conversely, GIST had better solution performance that the Mashup display, which is thought to be because of an ecological approach, which could be incorporated into design workshops to ensure that a myopic focus on the user does not preclude consideration of their work environment. A combined approach between the two for the design of future Sonar and TMA interfaces could lead to better task performance than ether in isolation.

The third and final reason for support of a combined approach came from the workload scores, where the Mashup was rated better. This contradicted previous literature where EID has been shown to improve workload. It was though that the increase in workload stemmed from capability being designed around interpreting the environment, with some tasks in GIST forcing higher cognitive workload as a result. With UCD and EID both having been shown to lower workload compared to each other, in addition to interfaces designed without them, it would be pertinent to consider both approaches in tandem when conducting future workshops.

The design, implementation, and evaluation process of GIST has been described throughout the thesis, although the focus on EID could leave potential adopters querying the benefits when compared to UCD workshop approaches that have become commonplace in the design of military interfaces. This chapter addressed this, comparing and contrasting the difference between two interfaces designed using differing approaches. It demonstrated that either approach alone did not yield better performance in all areas, and that a mixed approach could be beneficial for future designs to merge the strengths of both approaches.

Chapter 11 **Conclusions**

The aim of this thesis was to explore whether Ecological Interface Design (EID) would be a suitable design paradigm to employ in future submarine control rooms, and how the associated design process from Cognitive Work Analysis (CWA) could be framed within the context of eventual software development efforts. There was a focus on Sonar and Target Motion Analysis (TMA) as key components of the tactical picture process (Stanton and Roberts, 2018), a key aspect of maintaining ownship safety (Mack, 2003). This chapter will summarise the work in this thesis, explain its significance, and finally propose future work that could be of interest for future research endeavours.

11.1 Summary of Work

This research had three key objectives, which are presented below, along with how they have been met.

11.1.1 Objective 1: Creating a detailed understanding of Sonar and TMA operation

There was a paucity of consolidated and comprehensive operation regarding how submarine control rooms operate that could be used to make informed decision decisions as part of this thesis. Previous literature had addressed specific aspects of operation, such as the application of Cognitive Work Analysis (Stanton and Bessell, 2014) or Event Analysis of Systemic Teamwork (Stanton, 2014) to the process of a submarine returning to periscope depth. Other work provided overviews of how submarine control rooms operated, contextualised by the work that was being presented (National Transportation Safety Board, 2001; Bisantz *et al.*, 2003; Mansell, Tynan and Kershaw, 2004; Arrabito, Cooke and McFadden, 2005; Ince *et al.*, 2009; Kirschenbaum *et al.*, 2014; Marine Accident Investigation Branch, 2016; Michailovs *et al.*, 2021; Michailovs *et al.*, 2022). However, there remained a need to synthesise this information into a holistic overview of how tactical picture compilation worked on submarines.

Chapter 4 addressed this, highlighting the complexity of the work that submarine command teams conduct, but also the amount of uncertainty they must account for. The submarine control room is a complex sociotechnical system, owing to the combination of highly trained operators and advanced technology working together in a goal oriented manner (Ly, Huf and Henley, 2007; Walker *et al.*, 2008; Stanton, 2014; Stanton and Bessell, 2014). This was accounted for, detailing the role of the social and technological agents in creating a tactical picture. It was noted that some agents, such as the Officer of the Watch, face challenges as part of this in managing the amount of data that must be aggregated (Dominguez *et al.*, 2006). Breakdowns in the flow of information, degrading tactical picture accuracy, have been identified as causal factors in accidents (National Transportation Safety Board, 2001; Roberts and Tadmor, 2002; Marine Accident Investigation Branch, 2016), confirming that a poorly performing subsystem can negatively affect an entire system (Meshkati, 1991). This review of control room operation and associated problem contextualised the subsequent case for application of EID to Sonar and TMA. This necessitated a CWA to be undertaken as a formative assessment (how work could be conducted; Naikar, 2013; Stanton *et al.*, 2013; Stanton *et al.*, 2017a), building on the understanding of how work should be conducted from Chapter 4. Individual stages of CWA can be chosen depending on analysis requirements (McIlroy and Stanton, 2015b). Therefore, two stages of CWA were chosen as they are directly associated with EID (Vicente and Rasmussen, 1988;1990; Burns and Hajdukiewicz, 2004; Jenkins *et al.*, 2009).

Chapter 5 explored the Sonar CWA in detail, presenting how the system is currently used to contextualise formative recommendations for redesign, which were summarised in Table 17. There was a focus on the waterfall, which is an enduring and prevailing component of most, if not all, modern sonar systems (Chen and Burns, 2007), citing from (Waite, 2002). This is thought to be because of the information density it provides, along the interpretability of this information. Consequently, the waterfall was identified as a leverage point, a concept proposed in the CWA Design Toolkit (Read *et al.*, 2018). Read *et al.* (2018) describe leverage points as system aspects that could yield large system changes through small changes to aspect being considered. While common, the waterfall is not without issue, and it was posited that the difference between its perceptual form and what it represented may cause an increase in cognitive workload (Hanisch, Kramer and Hulin, 1991; Masakowski and Hardinge, 2000). This is important because more sonar data is being provided to operators (Dominguez *et al.*, 2006; Gosling, 2008; Defence Equipment and Support, 2010; Jacobus, Yan and Barrett, 2012; Smith *et al.*, 2013), and the increased amount could exceed an operators ability to process it or exceed their working memory capacity (Mason *et al.*, 1989; Woods, Patterson and Roth, 2002). Consequently, recommendations were focused on improving the ecological validity of the sonar data, transitioning processing of the waterfall data to a perceptual task where possible so that operators could operate at the lowest possible level of cognitive control (Van Dam, 2014; Cravens, 2021). Additionally, while multiple screens could create a natural “attentional spotlight”, multiple screens require operators to store more information in working memory (Michailovs *et al.*, 2022). Consequently, another theme for insights was to remove the working memory load required to work between screens and information representations, creating a single-screen representation of the data required.

Chapter 6 explored the TMA CWA in detail, presenting how the system is currently used to contextualise formative recommendations for redesign, which were summarised in Table 20. The Local Operations Plot was chosen as the leverage point to investigate. It was also considered as a pain point from the CWA Design Toolkit, defined as problems or issues that represent user frustration, conflicting goals, or information bottlenecks (Read *et al.*, 2018). This was because interaction with the speedstrips can be cumbersome for operators and hamper solution entry efforts. The Local Operations Plot is a prevailing method of entering solutions (Clarke, 1999), and a component of combat systems (a superset of TMA functionality, and actual systems used onboard). With newer combat systems, such Submarine Combat System – Next Generation (BAE Systems, 2015) or Aegis Weapons System (Threston, 2009), becoming more capable, there is a drive to introduce more automation, which could reduce error by reducing cognitive workload (Breton and Bossé, 2003). This is especially pertinent as the role combines information from around the control room, which may overload operators, especially if solution possibilities remain large. However, any automation added should ensure that operators remain ready to step in when the automation fails or is incapable of undertaking the task at hand (Bainbridge, 1983). These findings led to proposing that more automation should be introduced into the TMA system as it is predominantly a computational process (Punchihewa *et al.*, 2022). This has led to multiple algorithms being created, although it would be inappropriate to remove operators completely from the process as they are required to process 'soft' data (Punchihewa *et al.*, 2022). Furthermore, an operator's capability to improvise might outperform automation solutions (Chalmers, Easter and Potter, 2000; Van Dam, Mulder and van Paassen, 2008), meaning that they should be included in the solution generation process. The automation could offer varying levels of support for operators working at each level of the Skills, Rules, and Knowledge Taxonomy. Introduction of automation to tactical picture compilation has been observed to be beneficial to task performance and workload, although impaired non-automated task performance (Tatasciore *et al.*, 2020). However, subsequent increases in the degree of automation did degrade performance any further (Tatasciore *et al.*, 2018; Tatasciore *et al.*, 2020). This would suggest that as much automation should be introduced as possible, although Tatasciore *et al.* (2022) observed that automation failure detection was degraded with higher levels of automation. Therefore, it was proposed that operators be provided with automation tools that complete specific tasks, which would not require monitoring over a long period of time. This would build on top of the introduction of automating the solution entry to remove cumbersome interaction that had been identified as part of the analysis, freeing cognitive workload capacity for generating solutions, instead of expending unnecessary cognitive effort entering them.

11.1.2 Objective 2: Creating a documented analysis and design process, considering software engineering

The process for designing GIST has been documented throughout this thesis, with each part building on the last to present a detailed methodological account of how EID was applied to Sonar and Target Motion Analysis. Section 11.1.1 provided a detailed account of the steps taken in Chapter 4 (research the domain), Chapter 5/Chapter 6 (generating design recommendations from the Sonar and TMA Cognitive Work Analyses), which relate to other parts of the process that have been documented in this thesis. Consequently, only these parts will be presented in this section to avoid repetition.

Method selection is an important part of any design process, and the selection of EID was robustly justified. Chapter 2 presented the theoretical underpinnings of EID, which included two stages of Cognitive Work Analysis, Work Domain Analysis and Worker Competencies Analysis, and the Skills, Rules, and Knowledge Taxonomy. This was used as a base for this thesis, identifying and exploring the necessary theory to support the arguments made and conclusions drawn. Next, the argument against using User-Centred Design (Norman and Draper, 1986; Norman, 1988) in this instance was presented. While User-Centred Design is a widespread design approach (Vredenburg *et al.*, 2002), a lack of explicit focus on the environment and its constraints might lead to them being omitted from created designs. This omission was shown to be unsuitable in Section 4.1, which explored how work is conducted in a submarine control room, demonstrating that a core aspect of work is understanding their working environment and its constraints through a tactical picture. This contextualised Section 4.2, where this synergy was highlighted as an argument for using EID for Sonar and TMA. Finally, User-Centred Design was revisited in Chapter 10, where GIST was compared to a Mashup display. The results of Chapter 9 and Chapter 10, which will be discussed in more detail in Section 11.1.3, demonstrated that while EID does yield benefits over contemporary designs for both Sonar and TMA, there is still room for improvement, which could be enacted using a joint User-Centred Design approach. This has implications for current defence design workshops, which use a User-Centred approach (Hamburger, Miskimens and Truver, 2011; Hall, 2012; Turner, 2017; Fay, Roberts and Stanton, 2020). Based on the results of the conducted human in the loop experiment, it was proposed that future design endeavours should be taking a joint methodological approach, including EID in future workshops, as opposed to the User-Centred only approach that is currently utilised.

Chapter 3 documented the creation of a taxonomy of constraints that could be used for Work Domain Analyses, with possible use cases presented in Section 3.5. It was created to formalise the learning that was provided by being apprenticed to an expert, a common method of learning

Cognitive Work Analysis (Naikar, Hopcroft and Moylan, 2005), when creating the abstraction hierarchies for Chapter 5, Chapter 6, and Chapter 7. It was recognised that not all companies seeking to use the methods applied within this thesis will have access to a Cognitive Work Analysis expert, and therefore a literature-driven taxonomy of constraints that should be explored and accounted for would be aspect of the process to document. The checklist of constraint categories is useful in itself, and can also be combined with the works that describe prompts to elicit information (e.g., Naikar, Hopcroft and Moylan, 2005; Jenkins *et al.*, 2009; Read *et al.*, 2016) to systematically approach an analysis. However, the taxonomy was also presented in a linked form, showing how the different categories interconnect. By documenting this, future practitioners could direct their attention to categories that are linked to those already found, but have limited, or no, constraints within them.

Linking EID to software engineering was also explicitly explored and addressed at different points throughout the thesis. Chapter 8 explored the design (Section 8.2) and implementation (Section 8.3) stages, considering contemporary software practices. This is important as while there is substantial literature on creating the frontend design (e.g., Burns and Hajdukiewicz, 2004; Hajdukiewicz and Burns, 2004; Upton and Doherty, 2008; Read *et al.*, 2018) there is very limited literature on considering the backend (e.g., Wells *et al.*, 2011; e.g., Oosthuizen and Pretorius, 2018), and a dearth of literature that considers both (e.g., Dhukaram and Baber, 2016). If EID is to move from simulator studies to production instances, it is vital that research in the domain further explores how this will be achieved. While the frontend and backend of software are separate, they are intertwined, and given that parallels have been drawn between software engineering systems analysis artefacts and Cognitive Work Analysis, it is pertinent that EID approaches should start considering both aspects, especially if there is an eventual end-goal of real-world implementation; duplication of work for the front- and back- end, especially where convergence has been clearly identified is intolerable for time and cost reasons for companies, and can contrast with established development principles such as agile.

The work in Chapter 7 sought to address this for the frontend by adapting the approach of seminal approach of Burns and Hajdukiewicz (2004) to be object-oriented, a key design principle of modern software development. This change was deliberately aligned with backend implementation principles, allowing work to be shared between the two, and to streamline development. Another key modification was to explicitly account for the uncertainty faced by command teams (Hunter, Hazen and Randall, 2014), relating to the challenge of sensor uncertainty identified by Vicente (2002). The challenge of uncertainty is especially pertinent in submarine control rooms (Dry *et al.*, 2005; Hunter, Hazen and Randall, 2014; Kirschenbaum *et al.*, 2014), making it a priority to address in any documented methodology. The backend method in Section 8.3, proposed that the objects and associated properties identified from the design could serve as a base for the underlying software design. However, modern software does not all rely on explicit designs, and some approaches, such as agile, might not use formal modelling methods. An example of this is when changes are made to an existing system, which would not require formal modelling to take place, rather the changes could be made directly in the code base by a software engineer familiar with the code. Therefore, it was proposed that the WDA be directly used to create or extend the agile backlog for software engineering teams, allowing them to structure their work as desired.

11.1.3 Objective 3: Assessment of a novel Sonar and Target Motion Analysis HMI

The final objective was to assess the created EID interfaces for Sonar and TMA to determine if they yielded the improvements that were presented in the EID literature, such as increased usability (King, Read and Salmon, 2022), improved task performance (Vicente, 2002), and reduced cognitive workload (Lau and Jamieson, 2006; Nielsen, Goodrich and Ricks, 2007). When commencing this thesis, the plan had been to create two new EID interfaces, one each for Sonar and TMA. However, the Cognitive Work Analyses conducted for Sonar (Chapter 5) and TMA (Chapter 6) suggested that a combined interface would be the best approach to take for their redesign. Only one interface, GIST, was designed as a result, allowing operators to conduct tasks from both roles. Consequently, experiments in this thesis assess participants completing the tasks of both roles. Regrettably, the Coronavirus-19 pandemic affected the planned human in the loop studies, and therefore they were merged into a single experiment using a repeated measures design. A study with individuals was chosen to avoid a confounding effect of changing a key aspect of the control room sociotechnical system.

Before completing any human in the loop study examining the interfaces in the ComTET simulator, it was vital to ensure that results could be reasonably be expected to be applicable to real-world environments. Chapter 7 presented a comparison of the ComTET abstraction hierarchies against those created during a study at HMS Drake’s Talisman trainer to compare the two work domains, following the approach of similar to the approach taken by Burns, Bisantz and Roth (2004) and St-Maurice and Burns (2018). The comparison was contextualised among other work looking to validate the simulator, conducted by other members of the ComTET project team (Roberts *et al.*, 2020), and beyond the scope of this thesis. The comparisons focused on establishing functional and physical fidelity for the versions of Sonar and TMA used in ComTET, which would serve as the baseline interfaces. It was revealed that the ComTET Sonar and TMA HMIs could be comparable to their real-world counterparts, although they only contained a subset of functionality and capability, in line with their status as an experimental research facility, compared to Talisman being a high-fidelity trainer.

Chapter 9 examined the differences between contemporary interfaces and GIST as an EID. It was hypothesised that usability, task performance, and workload would be affected. The results were promising for further exploration of EID in a submarine control room for Sonar and TMA, with all hypotheses being supported by the data, and in favour of GIST. Subjective usability was significantly higher in GIST, confirming previous studies where an improvement was found (Behymer, 2017; Tran, Hilliard and Jamieson, 2017; King, Read and Salmon, 2022). Potential avenues for further usability improvement were identified, focusing on addressing inconsistency and operator confidence, as these questions were rated in favour of the contemporary interface. This need for improvement was highlighted by the absolute SUS score. While a significant relative improvement was identified, GIST's score still fell below the commonly accepted baseline of 70 (Bangor, Kortum and Miller, 2008), demonstrating that there is room for continued improvement. However, it was noted that usability is a scale, not a dichotomy, meaning that meeting or exceeding the threshold would not automatically guarantee usability in all circumstances.

Objective task performance was shown to be better in GIST, with participants assigning more trackers, having improved merge performance, and assigning more accurate solutions. This confirmed the results of previous experiments which have shown EID can improve task performance (Jamson, Hibberd and Merat, 2015; Borst *et al.*, 2017; Shier *et al.*, 2018; Zestic *et al.*, 2019; Schewe and Vollrath, 2020; Cravens, 2021), and confirmed results from a domain-specific study by Michailovs *et al.* (2021). It was proposed that the principles of EID contributed to the improvements. This included a reduction of cognitive workload, transitioning to perceptual workload, by integrating information, a common approach for military EID designs (e.g., Hunter, Hazen and Randall, 2014; Michailovs *et al.*, 2021; Michailovs *et al.*, 2022). Another key factor was representing all data on a map-based display, which was posited to have made constraints affecting merges (i.e., sources should have the same bearing, but different sensors) directly perceivable, a key goal of EID (Gibson, 1979; Rasmussen and Vicente, 1989; McIlroy, 2016). Additionally, making broken constraints more visible to the operator when entering solutions, a factor which was has been shown to improve fault detection performance in previous experiments (Jamieson, 2002; Reising and Sanderson, 2004; Rechard *et al.*, 2015). Incorrect solutions could still be entered, there were several cues to inform operators that their solution was likely incorrect, prompting them to update it. While an updated solution that matches the data can still be incorrect (DeAngelis and Green, 1992), based on a practically infinite number of solutions without sufficient constraining data (Cunningham and Thomas, 2005), encouraging operators to enter plausible solutions reduced the possibility of misplaced contact detrimentally affecting ownship safety. The additional information for solutions was functional information, which has been shown to improve task performance (Torenvliet, Jamieson and Vicente, 2000; Vicente, 2002). Other functional information was also included in GIST, such as a predictive waterfall, allowing operators to directly perceive a contact's future state.

Finally, subjective workload was shown to have improved when using GIST, although this was non-significant. There was a significant effect of scenario observed, however, validating that the low- and high- difficulty scenario designs were perceived as such by the participants. A workload reduction in GIST supported SRK literature (Rasmussen, 1983; Rasmussen and Vicente, 1989), EID literature (Vicente and Rasmussen, 1992; Nielsen, Goodrich and Ricks, 2007; Bennett and Flach, 2019), and previously conducted experiments on EID interfaces (Hall, Shattuck and Bennett, 2012; Selkowitz *et al.*, 2017; Calhoun *et al.*, 2018; Schewe and Vollrath, 2020; Michailovs *et al.*, 2021). As with the improved task performance, it was posited that the reduction in workload arose from GIST focusing on a map-based view of the information to facilitate direct perception, and therefore reducing cognitive demand to the lowest level required on the Skills, Rules, and Knowledge Taxonomy (Gibson, 1979; Dinadis and Vicente, 1996; Van Dam, 2014; McIlroy, 2016; Cravens, 2021).

Chapter 10 examined the differences between GIST and a User-Centred Design, which was called a Mashup display. It was hypothesised that task performance and workload would be affected, but that usability would not be affected. This was because both had been designed with the user as a core consideration, so it was proposed that usability would be rated similarly, although task performance and workload would differ due to the methods used to create the interfaces (EID and User-Centred Design). The first hypothesis was supported by the data, with no significant difference between the SUS score for each interface, following the same pattern as previously observed when comparing a User-Centred Design to an EID (Young and Birrell, 2012). The Mashup display was rated slightly higher though, which could be explained by the focus on usability arising from its creation in a design workshop (see Section 10.2.1).

The Mashup display showed significantly improved tracker designation over GIST. It was proposed that this was because the merged display of sensor data only required participants to designate half of the trackers they would be required to in GIST. This is an example of usability improvements that could be made, as recommended in Section 9.4.1, as GIST was designed to make the environment and its constraints directly perceivable using an EID approach (Ho, Dal Vernon and Jamieson, 2003; Kwok, 2007; Ellejmi *et al.*, 2018). While user feedback was incorporated into GIST's design, this does not preclude further improvements being made. The merged sensor display removed the capability to adequately compare merges, as this was not a task that could be completed in the Mashup display. GIST had significantly better solution accuracy, supporting the argument that displaying the constraints of an environment can lead to better performance (Torenvliet, Jamieson and Vicente, 2000; Vicente, 2002; Bennett and Flach, 2019). It was posited that this was a key differentiating factor between the two interfaces, as they both contained trialling tools, calculating unknown parameters from known parameters, and showed operator's the resultant solution overlaid onto the waterfall. However, GIST represented additional functional information in the form of the map-based view, allowing purposeful (functional) action (Lintern, 2006), and the predictive waterfalls. Presentation of physical and functional information has been shown to improve performance (Christoffersen *et al.*, 1994; Torenvliet, Jamieson and Vicente, 2000; Vicente, 2002; Bennett and Flach, 2019), as has showing predictive information over an observational or checklist approach (Agnisarman, Madathil and Bertrand, 2019). Finally, subjective workload was significantly improved when using the Mashup display. This contradicted prior experimental results that have shown EID has a lower workload than User-Centred Design (Wu *et al.*, 2016), or contemporary non-EID designs (Hall, Shattuck and Bennett, 2012; Selkowitz *et al.*, 2017; Calhoun *et al.*, 2018; Michailovs *et al.*, 2021; King, Read and Salmon, 2022). One possible reason for this was an increase in familiarisation time that might be required for GIST as an EID (Young and Birrell, 2012). However, it was proposed that the increase in workload was because participants were required to work using Knowledge-Based Behaviour on the Skills, Rules, and Knowledge Taxonomy (Ellejmi *et al.*, 2018). This was because while GIST utilised a map-based representation of the environment, some of the tools provided might have required cognitive, instead of perceptual, work when using them, increasing workload (Van Dam, 2014; Cravens, 2021). By contrast, as a simplification of the actual process, the Mashup tools required text entry, which showed visual output on the waterfall, only requiring participants to use perceptual cognition to determine whether two lines were aligned.

11.2 Contributions

This thesis has made theoretical, methodological, and practical contributions. They are presented in this section, explaining their rationale and what they add to existing knowledge.

11.2.1 Theoretical

11.2.1.1 **Confirmation of Constraints Across the Literature**

Constraints are a core component of Cognitive Work Analysis and Ecological Interface Design, bounding work within a system within which activity can take place (Vicente and Rasmussen, 1988; Borst, Flach and Ellerbroek, 2015). As EID can be applied to a wide variety of systems, there are a multitude of constraints that might have to be considered to apply these methods. Prior literature has addressed this through prompts to elicit constraints (Naikar, Hopcroft and Moylan, 2005) and examples of constraints discovered in analysis walkthroughs (Jenkins *et al.*, 2009; Stanton *et al.*, 2017a). However, they did not present a holistic overview of what constraints have actually been found in the literature across all domains, and how these categories interacted. This presented two issues, which were addressed by the taxonomy of constraints in Chapter 3. Firstly, while the list of prompts by Naikar, Hopcroft and Moylan (2005) were comprehensive, without analysis of EID literature, it was not possible to confirm that they were exhaustive. Performing a systematic literature review to identify all constraint found from EID applications addressed this, creating twenty categories of constraint that have been observed in the literature. Secondly, previous treatments did not fully explore how categories of constraint might be related, despite these interactions being acknowledged in the literature (Burns, Bisantz and Roth, 2004; Bennett and Flach, 2011). The taxonomy addressed this by applying Social Network Analysis to reveal which categories of constraint were linked and how strong the links were. By addressing these issues, the taxonomy contributes to future theoretical discourse by providing a concrete literature-backed assertion of the types of constraints that sociotechnical systems have been assessed to exhibit, and how these constraints are connected.

11.2.1.2 How Submarine Control Rooms Operate and Why Ecological Interface Design is Suitable for Sonar and Target Motion Analysis

EID, or its principles, have been applied to the maritime domain before (Burns, Bryant and Chalmers, 2000; Chalmers, Easter and Potter, 2000; Burns, Bisantz and Roth, 2004; Burns, Bryant and Chalmers, 2005; Dry *et al.*, 2005; Chen, 2007; Ly, Huf and Henley, 2007; Hunter, Hazen and Randall, 2014; Michailovs *et al.*, 2021). There was also literature that provided an understanding of aspects of how submarine control rooms operate (Murphy, 2000c; Arrabito, Cooke and McFadden, 2005; Baggeroer, 2005; Dominguez *et al.*, 2006; Matthews *et al.*, 2006; Carrigan, 2009; Hamburger, Miskimens and Truver, 2011; Kirschenbaum *et al.*, 2014; Stanton, 2014; Stanton and Bessell, 2014; Philippe *et al.*, 2016; Michailovs *et al.*, 2021). However, there was not a consolidated, holistic, overview of how the submarine control room sociotechnical system operated with regards to tactical picture compilation, including how the interfaces of Sonar and TMA were used to facilitate this. Furthermore, there was little explicitly highlighting how this operation made them especially suitable for the application of EID, especially for Sonar and TMA. This was addressed by Chapter 4, published as Fay, Roberts and Stanton (2019), which presented the work of submarine control rooms in Section 4.1, and used this to contextualise the case for the application of EID in Section 4.2. This argument was based on a need to move from an evolutionary approach to design, which has unnecessarily retained legacy constraints, to enact a step change that ensures the challenges of future requirements can be addressed. These challenges include maintaining or reducing crew sizes (Masakowski, 2000; Ly, Huf and Henley, 2007; Stanton and Roberts, 2018), processing larger volumes of data from improved or new sensors (Chalmers, Easter and Potter, 2000; Duryea, Lindstrom and Sayegh, 2008; Roberts, Stanton and Fay, 2015), and utilisation of more displays (Chalmers, Easter and Potter, 2000). An evolutionary approach, driven by a need to maintain training readiness (Hall, 2012), reduce cost as well as risk (Gosling, 2008), and onboard factors (Defence Equipment and Support, 2010), has previously worked to address these challenges. However, enhanced capabilities of modern combat systems are necessitating consideration of new designs to ensure that submarine control rooms remain at the vanguard of capability. By proposing EID for this step-change and presenting an in-depth case of why, linked to the underlying theory, future work can use and modify this theoretical contribution as a robust justification for utilising an ecological approach to submarine control room system design.

11.2.2 Methodological

11.2.2.1 **Taxonomy of Constraints**

The taxonomy of constraints presented in Chapter 3 has enhanced the method for conducting analysis using Cognitive Work Analysis, with a focus on Work Domain Analysis. Previously, this was either achieved by practitioner experience, being apprenticed to a Cognitive Work Analysis expert, exploration with a domain expert, and/or using the prompts provided by Naikar, Hopcroft and Moylan (2005). However, there was no literature that provided a systematic overview of constraints that had been found across previous analyses of complex sociotechnical systems for EID, and there was no indication of how the constraints were interconnected. Consequently, categories of constraints might have been missed from a conducted analysis, and there was not a method to direct the order of exploration for constraints. The taxonomy addressed this by providing a list of twenty categories that were informed by the literature, and how they were linked. Section 3.5 explored how the taxonomy could be used. The first option was as a check list, see Table 12 and Figure 10, of categories and the top constraints found in them, which would be systematically worked through as part of the analysis. Another option was to use a guided checkbox approach, see Figure 11, where practitioners would be cognisant of other categories being mentioned during exploration of the category they were currently exploring, and jump to those categories to direct the process, instead of blindly working down the checklist. The final proposed approach was to use the visual representation of the taxonomy, see Figure 9, to identify linked categories to the one currently being explored, evaluating whether they should be explored as there is little to no constraints present from them. This has the advantage of directing the analysis process to categories which might not have been considered, hence the lack of constraints from the category in the analysis, to ensure that all known categories of constraint within complex sociotechnical systems have been considered.

11.2.2.2 **Utilising Implicit Representations in Interfaces for Work Domain Analysis**

Section 5.2.1 presented an alternative approach for Work Domain Analysis, largely following the prevailing methodology (Burns and Hajdukiewicz, 2004; Jenkins *et al.*, 2009; Stanton *et al.*, 2017a). However, the Physical Objects level of the Abstraction Hierarchy included components from the interfaces under consideration. This was because it is was observed that the existing Sonar and TMA designs already had already implicitly included representations of the work domain, thought to be caused by them being a mix of law- and intent- driven domains as they were in a military context (Bennett, 2014). Another driving factor for the prevalence of implicit metaphors was that the work domain is seldom directly observable to submariners, meaning that the control room has developed using representations of the work domain, such as the tactical picture or the Local Operations Plot. Consequently, categories of items in the interfaces would provide a representation of the work domain under consideration. Categories of items were used to maintain adherence to the methodological requirement of describing categories and not instances (Naikar, Hopcroft and Moylan, 2005). While abstracted away from the complexities of the complete work domain, the modified approach was desirable to set an appropriate scope for the analysis, as advocated for by Kortschot *et al.* (2017). Practically, it can also help to address potential issues with Cognitive Work Analysis (Vicente, 2002; Stanton *et al.*, 2013; Hou *et al.*, 2015), requiring less subject matter expert contact, saving time, and potentially cost for analyses. Another identified benefit was that the generated artifacts could be used for training and documentation purposes as well; providing operators with the information could assist them with training, facilitating an understanding of what items in the interfaces do or how they can achieve their goals, and system suppliers could use the information to understand how a system is expected to perform, and why functionality is required. These approaches could also be used to bolster the "Physical Object Cards" concept in the Cognitive Work Analysis Design Toolkit (Read *et al.*, 2015b), which explores new ways of utilising Physical Objects within a system. The method proposed in this thesis was cognisant of the iterative nature of Cognitive Work Analysis, so an initial iteration could be developed using the interface Physical Objects and subsequently modified after applying the method of Read *et al.* (2015b). A potential downside to the approach is that workers can become dissociated from the underlying domain (Rasmussen, Pejtersen and Goodstein, 1994; Naikar, Hopcroft and Moylan, 2005), degrading the completeness of the resultant abstraction hierarchy. This was addressed explicitly through comprehensive questioning of experts ("How …?", "Why …?") when constructing the abstraction hierarchies to ensure that the domain was accounted for. Extensive operator training addressed the issue implicitly, with all operators being fully cognisant of the complexities of the domain they worked in.

The completed Work Domain Analysis and Worker Competencies Analysis for Sonar (Chapter 5) and TMA (Chapter 6) are also contributions. Other work has applied Cognitive Work Analysis to maritime control rooms (Burns, Bryant and Chalmers, 2000; Burns, Bisantz and Roth, 2004; Burns, Bryant and Chalmers, 2005; Burns, 2012; Hunter, Hazen and Randall, 2014). However, this work distinguishes itself through the scope of analysis, concentrating solely on the interfaces, and presenting a comprehensive walkthrough of each station's operation to identify shortfalls. Another important distinction is that the roles in ComTET were designed to be as general as possible, maximising generalised applicability, and creating a resource that others could adapt for future analyses of Sonar and TMA.

11.2.2.3 **Enhancing the link between Ecological Interface Design and Software Engineering**

Chapter 8 presented how existing Cognitive Work Analysis to Ecological Interface Design methodology could be updated to better connect with contemporary software engineering practices, considering both the front- and back- end designs. Providing this link is a vital contribution as companies have often heavily invested in their software processes as a cornerstone of modern business, and concepts must be compatible with this to gain traction (Baxter and Sommerville, 2010). As the role of human factors continues to grow in organisations, it is important to consider how best to disseminate and promote action on the important findings that are generated. While there are multiple methods for the creation of the frontend design (Burns and Hajdukiewicz, 2004; Read *et al.*, 2018), there was little for the design of the backend (e.g., Wells, 2011; Wells *et al.*, 2011; Oosthuizen and Pretorius, 2018), and a dearth of literature that considered both (e.g., Dhukaram, 2016; Dhukaram and Baber, 2016). The methods proposed in Chapter 8 addressed this by updating the design method of Burns and Hajdukiewicz (2004) to be object-oriented in line with modern software practices (Section 8.2.1), and arguing against the translation of Cognitive Work Analysis to software engineering modelling outputs, as proposed by Dhukaram and Baber (2016), instead utilising an agile Kanban approach for project management (Section 8.3). It was proposed that the current outputs (*"Work Domain Analysis"*, *"Classes for Code"*, or an *"Interface Design"*) involved in the creation of an EID could be integrated into a single process that mapped how each interlinked, and showed how they could be converted, either following proposed processes, or using the various seminal works (Burns and Hajdukiewicz, 2004; Wells *et al.*, 2011; Dhukaram and Baber, 2016; Oosthuizen and Pretorius, 2018; Read *et al.*, 2018). It was argued that an object-oriented approach was more appropriate for modern software designs, especially for Sonar and TMA, where the screen content is not fixed and the number of objects present would change, such as the number of contacts represented on the tactical picture. The backend was based on a modern agile approach, as software engineering is moving away from utilisation of the waterfall model (working from designs created at the start of a project/product) to agile approaches that facilitate responsiveness to change and quicker delivery. By applying the concepts of contemporary software engineering to EID, it presents an opportunity to leverage these methods to deliver EID interfaces using software methods that have been proven to work for large-scale and complex software systems. The proposed methods can be used for the design and implementation of ecological interfaces by established software teams, bringing the potential for their widespread adoption closer. Furthermore, the nature of agile software methodologies aligns better with the iterative aspect of CWA, improving on previous works that have not explicitly accounted for this.

11.2.3 Practical

11.2.3.1 **Demonstrating Ecological Interface Design is Suitable for Sonar and Target Motion Analysis**

This thesis has made a practical contribution by demonstrating that Ecological Interface Design is a viable option for the Royal Navy. Documenting the design and development processes, as detailed in Section 11.1.2, provides ComTET stakeholders with a practical demonstration of how the results were achieved, and detail on how ecological design could be applied to other stations or domains. Chapter 4 identified that while defence companies are making great strides with the products that they create, there is a reluctance to move away from something that has been demonstrated to work (Gosling, 2008; Hall, 2012). This has led to an evolutionary approach, which is valid, although it may have inadvertently and unnecessarily retained obsolete constraints. This could be hindering operators from taking full advantage of the cutting-edge capabilities offered by these systems, especially with the complexity of the future maritime environment. There is risk associated with making changes to 'tried and tested', so a practical demonstration of improvements provides justification to make changes. By demonstrating that a step change to Ecological Interface Design can have clear benefits above and beyond evolutionary designs, see Chapter 9, it is envisaged that companies will use this as an impetus to explore how the principles in this thesis could be applied to their own products. Such a step change would perhaps be even more pronounced than presented in Chapter 9, given the resources and expertise available to these companies, which would no doubt be demonstrated in products arising from this approach.

11.2.3.2 **Graphically Integrated Sonar and Target Motion Analysis: Sonar and Target Motion Analysis can be Merged**

The creation of Graphically Integrated Sonar and Target Motion Analysis is a practical contribution for two reasons. The first is that it demonstrates that it is possible to merge the two roles, which is a novel contribution to the literature. Previous literature has investigated the co-location of the roles (Roberts *et al.*, 2019; Stanton and Roberts, 2020; Stanton *et al.*, 2020b), and the sharing of information between them (Michailovs *et al.*, 2021). However, to the author's knowledge, there has not been publicly available research into the merging of both roles into one interface. This could have practical implications for the design of submarine control rooms and could be used to inform crewing levels. The second reason is that it practically demonstrates how the number of screens in a control room could be reduced, addressing related concerns identified by Dominguez *et al.* (2006) and Hamburger, Miskimens and Truver (2011), who argued that the number of screens in the control room could cognitively overwhelm operators. Displaying the information from two roles on one interface would reduce the required cognitive effort to interpret all information in a control room. Furthermore, it creates a basis to add information from other roles and screens that is compatible with a map-based representation (i.e., tactical picture information). While there is obviously a gulf in fidelity and capability between software for testing in ComTET and software ready for deployment in submarines, Graphically Integrated Sonar and Target Motion Analysis is a practical demonstration of what is possible, going above and beyond rhetoric on the possibilities of reducing screens and combining roles.

11.3 Evaluation of Project

11.3.1 Subject Matter Experts

11.3.1.1 **Input – Project and Researcher expertise**

The work in this thesis would not have been possible without the contributions of the many subject matter experts who provided their input, assistance, and feedback. It was fortunate that this research was supported by the ComTET research project, which afforded access to these subject matter experts, and allowed the author to become one themself. It is felt that this had a tangible impact on the structure of this research and the ideas presented. In terms of structure, regular project meetings with subject matter experts, and availability of ad-hoc support meant that ideas could receive regular feedback, as opposed to having to arrange specific workshops. Utilising expertise provided during the project also ensured that their time was appropriately used, gaining feedback on multiple work streams at once, instead of requiring individual meetings for each, which would have been unreasonably demanding on the experts’ time, and could have added significant cost to the project. This is prevalent in this research, whereby changes were enacted from regular contact with experts and feedback on reports, identifying where small changes could be made, as opposed to a workshop where sweeping changes were identified. Such an approach is akin to the software engineering practices employed in Section 8.3, seeking to use a responsive, agile, approach to be responsive to emergent requirements and input. This was facilitated by the author’s expertise gained throughout the project, allowing them to make informed decisions about what would be suitable, which could be confirmed or altered by subject matter experts. However, it is recognised that future work in this area may not be based within a project, and therefore it is encouraged to adopt a more traditional approach to the analysis, seeking input from experts using workshops and other fixed points as recommended in the literature (Burns and Hajdukiewicz, 2004; Stanton *et al.*, 2017a; Read *et al.*, 2018).

11.3.1.2 **Demographics**

How subject matter experts input into the project also affected data collection about them, limiting demographic data that would usually accompany data sourced from a population sample. There were two limiting factors: the lack of ethical approval, and more importantly individual consent, to collect data from individuals generously providing their time and input; and limitations on demographic data that could be collected from them (Stanton and Roberts, 2018; Stanton and Roberts, 2020), such as the validation study.

It was not possible to collect demographic data from experts contributing to the research outside of the context of a study. As described above, this was a significant source of information. Their input was incorporated into the research at all possible points, although as no ethical approval was in place, demographic data could not be collected and reported. Even with ethical approval to do so (i.e., a long-term study to collect the demographic data of ComTET subject matter experts), they might not have felt comfortable providing their demographic data along with their input outside of a study, such as at project review meetings. This could have risked deterring potential contributors, detrimentally affecting the amount of expert input. Possible reasons include the perception of gatekeeping input (i.e., researchers preferring input from more senior personnel) and/or being required to provide information for their input to be considered (i.e., *"Why is information about my experience required? I am an expert, and my input is valid"*). Given the small population of experts, and the amount of input required for a program such as ComTET, it was instead deemed more appropriate to only collect data on study participants.

However, the collection of demographic data within studies was also limited, as with other ComTET studies, by security considerations and recommendations from the Ministry of Defense Research Ethics Committee (MoDREC; Stanton and Roberts, 2018; Stanton and Roberts, 2020). Combined, with sample sizes prohibitive of disaggregation, it is recognised that these are limiting factors for utilisation beyond reporting descriptive population statistics. This is especially pertinent in Chapter 7, where the demographics are mostly temporal, with the assumption that time $\propto$ experience. While rank and role were collected as more evidence-based indications of differing skill levels, the small population size would have prohibited attribution of results to a specific demographic cross-section without potentially uniquely identifying participants.

Therefore, future work within the area would be encouraged to utilise more robust mechanisms for collecting contributor data, and expanding on the data that is collected. If following the experiment only approach recommended in Section 11.3.1.2, this would be through the collection of ethics panel approved demographic data. Conversely, if future research was again project based, it would be prudent to gain ethical approval for collecting demographic data for contributing experts outside of specific studies in a structured manner, such as optional demographic questionnaires for attendees of feedback and review meetings. While this would not address the possibility that some subject matter experts might still decline, it could positively drive participation by becoming a pre-emptive part of the project, as opposed to a reactive request when input is provided.

A more robust approach to organising, processing, and analysis data would be used. This could be via an associative system, such as the NVivo coding approach used by Read *et al.* (2022). Consequently, it would be possible to provide more detailed demographic information (i.e., *"The ____ simulator was designed using input from n = _ subject matter experts. _ were currently operational submariners with _ years of experience. A matrix of contribution topics, organised by role, is detailed in table _"*). This approach could also be expanded to cover information gained by osmosis, see Section 5.2.2.1, improving data provenance for more traceable construction of analysis artefacts.

11.3.2 Taxonomy of Constraints

The taxonomy of constraints developed in Chapter 3 was created in response to identifying a gap in the methodological practice after conducting the CWAs for Chapter 5, Chapter 6, and Chapter 7. Consequently, it was fully developed after the analyses were approved by domain experts, after being created using guidance from the supervisory team and literature on constraints that could be expected to be found (e.g., Naikar, Hopcroft and Moylan, 2005; Jenkins *et al.*, 2009; Read *et al.*, 2016). This input led to complete and comprehensive abstraction hierarchies, which contained the expected categories of constraints from the taxonomy once it was finished. Section 5.2.2.1 stated that this was expected owing to the involvement of an expert CWA practitioner, confirming what Naikar, Hopcroft and Moylan (2005) commented on the process, in that being apprenticed to an expert makes the process more accessible. In the context of this thesis, it meant that applying the taxonomy to the finished analyses would have had limited, if any benefits, as the constraint categories found had already been suggested by the experts. Consequently, the taxonomy was used as a brief confirmatory checklist, instead of being fully integrated into the data collection method in this instance. Future applications of the CWA method would incorporate the completed taxonomy from the start, seeking to apply it fully to maximise its utilisation. While the inclusion of an expert CWA practitioner could minimise the need for the constraint taxonomy, it would still be worthwhile for experts to utilise it to formalise the application of knowledge from experience that is stored mentally and provide a literature-driven checklist of possible constraints.

Another limiting aspect of the taxonomy is the limited Cohen's κ for interrater reliability for the constraint categorisation, shown in Section 3.2.6. While a level of agreement was shown using the scale provided by Landis and Koch (1977), there was room for improvement in these scores to reach an optimal value ($\kappa \geq .8$). Discussions with colleagues during the process suggested that this was because constraints could be reasonably categorized into multiple categories, instead of the one category that NVivo allowed. This was partially accounted for by the provision of a secondary category that the additional rater could provide, although this could not be represented within the software and only affected the inter-rater reliability calculation through manual comparison. Future iterations of the taxonomy would address this limitation by moving to different software that allows for classifying constraints into multiple categories. Doing so may also improve the connections between each category, as the variable nature of some constraints could be accounted for. Another option would be to subset the taxonomy for different domains and types of domains, which would allow for more concrete implementations to be created, considering the exact nature of what each constraint means in the context of the work domain specified.

11.3.3 Testing after the Coronovirus-19 pandemic

The human in the loop experiments in this thesis were conducted in the ComTET simulator facility using novice participants. The use of novices to understand relative differences in performance has been demonstrated to be appropriate to increase statistical robustness (Walker *et al.*, 2010c; Stanton and Roberts, 2019), with the direction of change being a focal point as opposed to the absolute values. The approach was successful, permitting statistical analysis to be conducted for comparing the contemporary interface to GIST (Chapter 9), and GIST to a Mashup display created using a UCD approach (Chapter 10). However, recruitment for the human in the loop study in this thesis took place shortly after the restrictions of the Coronavirus-19 pandemic were lifted, and uptake for a multi-day experiment in a confined space with other participants was not optimal. This affected final participation numbers from the desired total of 60. While there was enough participation from a testing period of nearly three months, the number of participants was still limited overall.

11.3.4 Testing with Novices

The approach of using novices has been followed in other studies conducted in the ComTET simulator. However, they also included one team of submariners to act as the gold standard (Roberts, Stanton and Fay, 2017a; Roberts, Stanton and Fay, 2017b; Stanton, Roberts and Fay, 2017; Roberts, Stanton and Fay, 2018; Stanton and Roberts, 2018; Roberts *et al.*, 2019; Stanton and Roberts, 2019;2020; Stanton *et al.*, 2020b). Most participants responding to the call for participation in the experiment in this thesis were novices, with defence companies also facilitating volunteering efforts of their staff, who had awareness of the domain. This meant that limited insight was gained into how experts would perform using GIST. This was a limitation in the studies conducted, which could be addressed by future studies that seek to only recruit expert participants.

However, getting enough participants for an expert only study could be challenging. The experts required are either working at sea or are shore-based and would be fulfilling obligations associated with this, both personal and professional. This makes recruitment of experts challenging, especially at the number required for a statistically robust study in a reasonable timeframe. While the domain experts have always been extremely generous with their time and contributions to the work conducted in this thesis, an expert only study would likely require an opportunistic approach of when a suitable number of submariners are free to participate, as opposed to responding to a call for participation within a fixed window. This is compounded by other ongoing experiments within the ComTET project, meaning that access to the facility must be coordinated and booked, which might not always be possible.

11.3.5 Method Substitutions

The methods applied in this research were affected by resource constraints, namely time and subject matter expert availability. This was offset by incorporating the outcomes of other methods used in ComTET, discussed in Section 2.3.3 and Section 2.3.4, to make the best use of each. Parallels were drawn between these methods and the methods that they were utilised in lieu of, such as utilising Hierarchical Task Analysis (HTA) and Event Analysis of Systemic Teamwork (EAST) over the stages of Cognitive Work Analysis not explicitly linked to Ecological Interface Design theory. This was successful in that the core theoretical underpinnings were adhered too, although was less than optimal for two reasons.

The first is that while parallels can be drawn between methods, there can be no substitute for the application of the method itself in some circumstances. For example, Salmon *et al.* (2010) conclude their comparison of HTA and CWA by stating that the approaches were complimentary, but are entirely different. The primary differences are the concentration on goals and constraints respectively, and CWA being a formative method. It is argued that these differences are what enable subsequent designs to address unexpected and unfamiliar events (Naikar, Hopcroft and Moylan, 2005). While the remaining stages of CWA are not intrinsically linked to EID, their application is still warranted to account for the constraints they address (McIlroy and Stanton, 2015b). Given that HTA (and EAST) does not operate in this manner, this is a limitation with regards to the completeness of analyses applied. Future work in this area should seek to address this by using a more robust battery of methods that are fully suitable for the problem being investigated, without the use of substitutions.

The second is a missed opportunity to apply a “many models” approach. Salmon and Read (2019) differentiated this approach from a traditional toolkit approach as being the application of multiple methods to the same unit of analysis and behaviour, over investigating something different with each. They proposed using five methods to examine systems and synthesise insights from them, both individually and between them. Demonstrating the approach by utilising it to examine road safety, they showed that insights could be enhanced over each method individually. Two of the five methods are used in this research: CWA, and EAST. However, they were used discretely on different units (the interfaces and control room respectively), albeit with synergistic aims. Consequently, the full extent of possible insights might not have been realised from the analyses completed. Furthermore, as a corollary of the first reason, insights from different perspectives (i.e., a formative CWA vs descriptive/normative HTA) have not been applied to the problem, acting as a further limiting factor. None of the insights derived by Salmon and Read (2019) were incongruent with each other, suggesting that this is a completeness issue, rather than an invalidation issue. However, this is not known to be generalisable, and given the safety-critical nature of Sonar and TMA, future work should seek to evaluate in a complete a fashion as possible, a goal for which the many models approach seems apt.

11.3.6 Real-world Applicability

This research was conducted in line with the aim of ComTET, which was a programme of work designed to understand current ways of working in submarine control rooms and provide evidence-based recommendations for meeting the challenges of future requirements (Roberts, Stanton and Fay, 2015). The results and recommendations from this research have real-world applicability, although not directly to current systems, owing to the level of fidelity used, discussed in Chapter 7. Instead, they are designed as in-feeds to the design process that will be followed for the design of next-generation design submarine platforms.

In this respect, transferability of the as-is experimental results presented in Chapter 9 and Chapter 10 is limited, as the research was not conducted in real-world simulator facility such as Talisman. However, this would apply primarily to the HMIs themselves (i.e., the contemporary interfaces being of low fidelity, and limited in comparison to the higher-fidelity counterparts), as EAST networks from scenarios completed in ComTET have been shown to be comparable to scenarios in Talisman (Roberts, Stanton and Fay, 2017b; Roberts, Stanton and Fay, 2018; Stanton and Roberts, 2020). That is, the work is recognisable in terms of the tasks, information, and communications, although the specifics are different. One such difference is that novices communicated more (Roberts, Stanton and Fay, 2017b). The cited literature observes that the comparability is likely from subject matter input in the simulator design process (Roberts, Stanton and Fay, 2015), and therefore relative differences would likely be seen if conducted in a higher fidelity simulator. Roberts, Stanton and Fay (2017b) argued that this provides tentative initial support for the relative validity of the ComTET experiments. They subsequently nuanced this assertion by recommending that future work compare novices to experts to ascertain exactly which results could be directly applied (absolute validity), and which would require higher-fidelity evaluation (relative fidelity). This has been in part achieved by having a gold-standard comparator currently active Royal Navy submarine command team participate in each study, which saw similar relative results to the novices in experimental cycles (Roberts *et al.*, 2019), although this was a result of observational over statistical derivation.

There are differences between the ComTET experiments and those in this research, namely that no submariners participated, and it was an individual experiment. However, the same work was completed by each operator, sans the communication aspects. This means that the social network could be disregarded, and the information and task networks could be bounded to the two roles the participants were completing (i.e., no visual cuts from periscope). Given that the networks were created from picture compilation command teams, formed of mostly Sonar and TMA related operators, this is unlikely to substantively change the networks over pruning select nodes related to periscope and ship control. These observations suggest that results concerning the work and procedures conducted could be transferrable to real-world submarine control rooms as-is, and HMI-related factors should be referred for higher-fidelity testing. More specifically, improvements to tasks completed (tracker assignment, contact merging, and solution entry) are posited to be readily applicable to the real-world; the tasks completed were representative of real-world interfaces, including HMI aspects, so if an EID interface was deployed, it could reasonably be expected to yield the same direction of benefits. This is likely more accurate for the Sonar component, as the TMA component is a subset of the full functionality used. By contrast, the perceptions of usability and workload are tightly coupled with actual implementations of HMIs, and so would require further evaluation against actual systems before being directly transferable. Of course, the transferability of all results would become more robust with higher-fidelity testing, so these remarks are not intended to preclude comprehensive evaluation of all aspects.

It is hoped that all of this research will be considered for application in submarine control rooms, irrespective of direct or indirect transferability. One route to applicability would be through workshops such as the Tactical Advancements for the Next Generation (TANG) forum, which was an initiative to capture novel ideas arising from design thinking sessions with from individuals not yet fully engrossed into submarine culture (Hall, 2012; Johnston and Featherstone, 2014). Forum attendees were asked to conceptualise ideas that would allow the submarine fleet to keep pace with industry technology, and were instructed to think big (Hall, 2012). There are other similar initiatives being held, such as design sprints (Fay, Roberts and Stanton, 2020). Resultant ideas from the design activities are prototyped and demonstrated to circulate them, and gain feedback. These ideas can then be realised by technology insertion programmes, such as the Acoustic Rapid COTS Insertion (ARCI) programme, which is designed to continuously improve Sonar systems by keeping pace with technology advancements (Guertin and Miller, 1998; Scott, 2006; Johnston and Featherstone, 2014). Johnston and Featherstone (2014) argued that input from the fleet is crucial to drive these insertion programmes, as they inform the changes that should be made to align with end-user needs, over simply guessing where to address development efforts. The GIST HMI presented in this thesis is a practical contribution to this process, and could be put to the submarine community as a base idea, with an updated concept being implemented in the next available technology insertion. While ideas are typically not meant to be seeded, the results from the studies completed as part of this research suggest that this design direction is worth exploring further. Consequently, it could be framed to operators as being a proof of concept that now requires fleet input before being requested as a future capability. This is a realistic possibility given the drive of modern navies to keep pace with the latest technological innovations (Department of National Defence, 2001; Stone, Caird-Daley and Bessell, 2009; Threston, 2009; Stanhope, 2012).

Improving on modern submarine systems, such as Sonar 2076, is no mean feat as they are extremely capable. However, it is believed that GIST advances the state of the art on two fronts. The first is the merging of Sonar and TMA, as discussed in Section 11.2.3.2. Current systems are designed around current roles within the control room, and while they provide integrations with other systems, they offer limited functionality outside of the role(s) they are designed to support. This is by no means a criticism of these systems, rather an observation that multiple products form control room capability; no one system can do everything. GIST advances on this by demonstrating that it is possible to merge roles, or even have all roles use a common core display. Such capability could be realised by ongoing efforts to make submarine combat systems more open and modular, such as the Common Core Combat System (Sea, 2005; Owen *et al.*, 2006; Scott, 2006; Defence Equipment and Support, 2010). While there are no illusions regarding the relative capability of a proof-of-concept display in comparison to deployed at-sea capability, implementing this type of capability in a submarine control room would go above and beyond what is currently known to exist. The second is the way in which data is represented and the workflows implemented. As illustrated in Section 4.1.2.2, submarine interfaces have mostly retained the same "look and feel" since their inception, despite enormous increases in the underpinning technological capability. GIST has demonstrated an advancement on contemporary interfaces representative of those used in current state of the art systems, paving the way for future platforms to make full use of their technological capability.

It is this realisation of the designed interfaces that offers another avenue for real-world exploitation of this work, by utilising the design and evaluation processes in this research. They were designed to bridge the gap between insights from human factors analyses and HMI designs, and the software engineering required to realise these. Submarine programmes are hugely complex, requiring highly skilled personnel (Schank *et al.*, 2011). Each company will have their own methods for conducting human factors and software engineering for submarine projects and would not adopt the processes wholesale. However, an ever-growing culture of agile and continuous improvement means that there is scope for integration in the near future, especially as the methods created are designed to integrate with contemporary software engineering practices (Baxter and Sommerville, 2010). More generally, they could be applied in any company that is delivering complex HMIs based on human factors analyses; this research was focused on Sonar and TMA, although the design process was designed to generic for use across multiple domains, and utilisation a real-world setting.

A key component of this is the focus on integration of the two disciplines, especially given the time and subsequent financial savings that could be achieved by reusing human factors analyses to drive software design. Everything was designed to work as a cohesive whole, but also in a discrete manner to address specific challenges being faced by potential adopters. For example, the taxonomy of constraints is designed to facilitate more comprehensive CWAs, although these do not have to be in the format proposed Section 5.2.1 if an existing HMI should not require a WDA artefact to document it. Similarly, the design process in Chapter 8 does not require a CWA in any particular format (e.g., mapping specific levels to outputs, or adhering to specific schools of thought regarding theory and presentation) to be used.

Given the quality of people required to deliver a submarine successfully, it is realistic that they would be adaptable to the changes and could offer subsequent improvements based on their expertise. This would involve an investment in cross-training staff, although it could be offset against the productivity gains that are achieved by allowing two disciplines to work from one set of analyses, especially if other methods are integrated from the proposed future work in Section 11.4.2, as well as contributing to closing the gap between human factors methods and eventual system design identified by Read *et al.* (2018) when designing the CWA Design Toolkit.

11.4 Future Work

11.4.1 Additional Evaluation of GIST

11.4.1.1 **Improving GIST**

The study and result in Chapter 9 and Chapter 10 focused on individual operators to avoid a confounding effect in the results arising from a team study, as merging the Sonar and TMA roles would have changed communication and might have affected the flow of information for constructing the tactical picture. Using individuals instead ensured that this potential confound was removed. However, it is recognised that this removed the operator and the interface sociotechnical subsystem (Walker *et al.*, 2010b) from the full control room sociotechnical system. Given that individual subsystems and interfaces can affect entire system performance (Meshkati, 1991; Walker *et al.*, 2010b), future research should explore how GIST affects the entire command team sociotechnical system and whether this is congruent with the effects on individual operators. This would provide further evidence on whether an EID approach works for Sonar and TMA, contextualised with an examination of effects on the entire control room sociotechnical system.

However, it would be prudent to apply further human factors methods to improve the interface before conducting a team study. This is because informal feedback from participants indicated there were aspects of usability and performance that could be improved, which is reflected by the usability scores presented in Section 9.3.1 and workload scores presented in Section 9.3.5. A workshop could be held to understand how experts would like to change GIST to better suit their needs, either utilising stages from the Cognitive Work Analysis Design Toolkit (Read *et al.*, 2018) or other recommended human factors interface evaluation methods (Stanton *et al.*, 2013) to increase its usability as a step before the experiment. This is based on the recommendation of including both approaches in future workshops from Chapter 10, and could address the limitation identified in Section 11.3.3, as a smaller cohort of experts could be interviewed to gather user-centred insights to improve GIST. Such an approach would be more formal than was taken throughout the thesis, where existing contact time with experts as part of the project, such as at ComTET project review meetings or the design workshop described in Section 10.2.1, was utilised to elicit feedback on the design of GIST.

The application of the above methods would be oriented around improving the usability and interaction of GIST, using its current state as a base. Depending on the scale of any future research, a “many models” (Salmon and Read, 2019) approach could be utilised to make more foundational changes, or even completely new designs. While this was not possible in the current research due to resource constraints, future projects could utilise the library of analyses conducted as a base to expedite generating a full set of models to inform designs. For example, the remaining stages of CWA could be completed, or EAST could be conducted on the GIST scenario transcriptions to understand the effect it had, if any.

While a more substantial effort over refining GIST, this would be more aligned with one of the key reasons for this research; given that this research was conducted to move away from an iterative approach, it could be counter-intuitive to make an initial step change and then return to iterating. Instead, the “many models” approach could yield alternative designs, or a larger step-change over just tweaking GIST. This could also help to address the notion throughout the literature that EID alone does not drive interface design and is often mixed with other knowledge and methods (McIlroy and Stanton, 2015b), by formalising the pre-cursor CWA analysis’ integration with other methods. In doing so, the gap between analysis and design could be bridged in a more directed fashion. Stated another way, CWA and EID examine the constraints of a work domain, but a “many models” approach could offer a theoretically backed method to constrain the subsequent design options; the design of interfaces will always remain a creative endeavour with near limitless possibilities, so narrowing these down using tried and tested methods would be the next logic steps for a method that concerns itself with constraints.

11.4.1.2 **Experimental Directions**

It would also be of interest to see how the usage of GIST would affect other issues identified as part of the ComTET project, as it merges two operator roles, and this is highly likely to affect the characteristics of the submarine control room sociotechnical system. For example, multiple experiments co-located pairs of Sonar and TMA operators (Roberts *et al.*, 2019; Stanton *et al.*, 2020b) to reduce an identified bottleneck of communications (Stanton, Roberts and Fay, 2017; Roberts, Stanton and Fay, 2018); would merging the co-located roles provide similar benefits, or are there other factors that should be considered? This would also build upon the work of Michailovs *et al.* (2021), who investigated if information integration improved performance in the control room, by investigating if a full integration of the roles has the same outcomes. It would also create a basis for evaluating what aspects of the sociotechnical system would need optimising to ensure joint optimisation, over just updating the interfaces. This is likely to require changing the command team’s work, or restructuring the team itself, as some tasks involving communication between Sonar and TMA will have been completely eliminated by GIST; without accounting for this in the social component of the control room, potential benefits might remain untapped.

It is proposed that a two-part study would be required to achieve these goals. This would ensure that the independent variables, interface change and work structure change, are not confounded. The studies would be based on the final study of ComTET, which saw all operators working in an inwards facing circle (Stanton *et al.*, 2020b). The study was synergistic with the work of Michailovs *et al.* (2021), where multiple streams of information was available to each operator, although the ComTET experiment stopped short of providing copies of the information in each interface.

The first study would examine the effect that an updated version of GIST has on the control room, if any. This would be to assess whether the results from the individual GIST studies are reflected in the team studies, or whether there are deviations. Sonar and TMA operators would be asked to follow current ways of working, but would have access to the capability of their counterpart operators to generate and retrieve information if required. This would build on the work of Michailovs *et al.* (2021) by introducing the ability to fully manipulate the information, such as TMA operators generating a speed estimate using Sonar tools if required. A key area of examination would be ensuring that the amount of information presented is manageable, as integrating information from multiple sources has been shown to be detrimental to OOW performance (Dominguez *et al.*, 2006). This suggests that there is a balance to be struck. While task performance can be improved by providing more information (Michailovs *et al.*, 2021) to address the communications bottleneck (Roberts *et al.*, 2019), too much information would introduce challenges faced by supervisory operators that detrimentally affect performance.

The second study would be based on the application on sociotechnical systems theory to apply joint optimisation to the submarine control room to investigate how to address this. The results of the preceding study would need to be examined using appropriate methods to understand where the command team's procedures could be updated to reflect how they completed work with GIST. While the full changes would be data-driven, changes are likely to be oriented around redefining the command team roles, so that Sonar and TMA operators become "Picture Compilers". While this would not reduce the total amount of information available, it could reduce the amount of information operators are required to process and maintain situational awareness of, if their work is structured appropriately. This would be novel as there is a dearth of submarine control room HMI studies that subsequently examines if further improvements could be elicited through radical social changes over those superficially required to run the study.

The sample sizes required to run these studies with appropriate statistical rigor is highly likely to require novice participants, which can be appropriate for studies with limitations on target demographic recruitment (Walker *et al.*, 2010c; Stanton and Roberts, 2019). However, their sole use creates a limitation in that while they can be trained to perform the experiment, they are not the target demographic, and this stymies transferability and generalisability. To address this, all ComTET team studies have included a currently operational RN submarine command team to act as the ideal comparator team. This approach should be used for future studies where an entire cohort of submariners cannot be obtained. The teams have validated the fidelity of the simulator and tasks (Roberts, Stanton and Fay, 2015; Roberts, Stanton and Fay, 2018), and performance compared to a comparable scenario (Stanton, 2014) completed in a higher-fidelity simulator (Stanton, Roberts and Fay, 2017). Roberts *et al.* (2019) argue that results from these expert teams will exhibit relative validity over absolute validity, with the directions of results being similar to the novice teams. This would need to be validated for each study, but could offer assurances that any results could be generalisable to other operational command teams.

11.4.2 Software Design and Engineering

11.4.2.1 **Further exploring the synergies between CWA, EID, and software engineering**

This thesis has furthered the discourse on the link between Cognitive Work Analysis and software engineering, presenting a different approach to others that have sought to translate Cognitive Work Analysis (Wells *et al.*, 2011; Dhukaram and Baber, 2016; Oosthuizen and Pretorius, 2018). The precis of the argument was that translation to software modelling languages might not be as appropriate as providing software engineers with the Cognitive Work Analysis outputs and using those to follow an agile development approach. However, as the author was the main software engineer implementing their own interface in conjunction with the software provider, the effect of communication and integration with enterprise processes, as advocated for by Viller and Sommerville (2000); and Baxter and Sommerville (2010) could not be adequately evaluated. Therefore, an option for future research would be to study the integration of human factors methods into the software development lifecycle in appropriate detail, expanding on the work of Jamieson and Lau (2010) who explored how multiple distributed teams could work together to create an Ecological Interface Design. The complexity of modern software development activities makes this a significant undertaking to approach, and it is likely that there will not be a one size fits all solution. Another interesting factor to consider is the growing influence of “no code” or “low code” tooling, which could be a prime integration point to explore with Ecological Interface Design. As their naming suggests, these tools facilitate creation of software without substantial underlying code, requiring only the interface to be designed. Consequently, Ecological Interface Designs could be created mostly from their visual form, speeding up their implementation process. This would extend the approach taken by Rechard *et al.* (2015), who created a tool to create Ecological

Interface Designs using a widget-based approach, likely based on the recommendation of a visual thesaurus of forms by Burns and Hajdukiewicz (2004), Hajdukiewicz and Burns (2004), and Jamieson (2003). This could be quite promising for adoption of Ecological Interface Design, as no/low code approaches are a relatively new area of software creation, creating scope to influence any adoption in companies more than that of traditional software, which is likely already fixed in its process, with significant resistance to change (Viller and Sommerville, 2000; Baxter and Sommerville, 2010).

Another option to explore could be the utilisation of artificial intelligence to power ecological interfaces. The abstraction hierarchy could be represented as a neural network to guide actions by a technological agent. Neural networks are machine learning tools inspired by the human brain, mimicking how neurons signal each other to process inputs (IBM, 2021a). Using a neural network based on the abstraction hierarchy of a domain could allow technical agents to gain the same benefits afforded to social agents by an ecological interface design; instead of having specifically programmed behaviour, the technical agent would instead be operating within the bounds of the domain. This could also address the concerns of Dhukaram and Baber (2016) for utilising Unified Modelling Language to specify software design, as it was a fixed representation of the domain. There is literature that demonstrates compatibility between abstraction hierarchies and neural networks, such as an analysis of football commentaries from different sources structured according to an abstraction hierarchy performed by (Silva, Ribeiro and Lopes, 2021).

11.4.2.2 **Integrating other methods into the design approach**

The design approach proposed in this thesis was created to link CWA to EID, and the subsequent software engineering processes required to realise designs. However, it was designed with generalisability in mind, and therefore could readily be adapted to incorporate other Human Factors methods, such as EAST, HTA, the Functional Resonance Analysis Method (FRAM; Hollnagel, 2012), or Accimap (Svedung and Rasmussen, 2002). EAST and HTA are described in Sections 2.3.3 and 2.3.4. FRAM is a method for examining safety-related problems in complex sociotechnical systems by characterising them by their functions, as opposed to their physical structure (Hollnagel, 2012; Patriarca, Di Gravio and Costantino, 2017; Salehi, Veitch and Smith, 2021). While it is primarily designed for safety, it is not limited to this use-case and can be used in general (Patriarca *et al.*, 2020). Accimap is a method for describing accidents using the risk management framework of Rasmussen (1997), mapping contributory factors across six hierarchical levels to understand contributory entities (Svedung and Rasmussen, 2002; Salmon and Read, 2019).

This could be achieved using a variety of options. At the most basic level, the resultant recommendations from conducted analysis could be incorporated into the process by framing them as stories that detail how functionality should be incorporated into the new product. For example, EAST analyses can reveal what information is used in the control room, and the relations between them (Stanton, 2014). This might reveal information pertinent to solutions that TMA operators do not currently have access to in GIST, yielding an associated story (e.g., *"As a TMA operator, I want to access ____, so that the information is readily available"*).

The same logic could be applied to identifying other entities within the work domain that could be incorporated as objects in the software design. An Accimap analysis will reveal several entities that are related to an accident that has occurred, and these could be incorporated into the software design to provide functionality that accounts for them in future versions. For example, McCabe, Baber and Stone (2020) conducted an Accimap analysis on the Karen accident (Marine Accident Investigation Branch, 2016), and recommended that an Artificial Intelligence (AI) support agent be implemented to support pre-mission planning and vessel planning. The AI would be an entity (class) in the code, which would provide the requisite functionality. It is recognised that the AI would be created using several classes, however, for the sake of simplicity it has been described as a monolithic unit.

Moving beyond insights as inputs for the process, it could be adapted to directly integrate a variety of methods. This would be akin to the "many methods" approach advocated for by Salmon and Read (2019), as multiple methods could be applied, and their results integrated into the software process. The exact integration workflow would depend on the method and its outputs. For example, HTA could be well suited to defining the processes that code will need to follow to achieve a goal specified in the user stories, which contain a reference to a related goal in a library of HTA task trees. This would be beneficial in defining exact processes required by the software, addressing concerns about communicating findings of analyses to engineers that have been identified in the literature (Bruseberg, 2008; Baxter and Sommerville, 2010; Wells *et al.*, 2011; Dhukaram and Baber, 2016). Other methods could help to understand and implement processes, such as EAST task networks facilitating an engineer's understanding of what workflows the software will need to support. Alternatively, if designing for future ways of working, CWA's Strategies Analysis could also be employed; as a formative method, it would provide the bounds for engineers to creatively design capability within. Finally, as FRAM uses a functional-based approach, this could be used an initial blueprint of a system that should be created, and how functionality should be linked. This could be bolstered with information networks from EAST, which would describe requisite information, and could therefore inform the data structures used in the software.

Methods of linking and flow in each method could also be used to design resilience, which is a key aspect of modern software. This could be achieved by applying "broken links" approaches to various methods to understand the potential impact and proactively design for the possibility. Stanton and

Harvey (2017) applied this approach to understand possible risks in a RN training activity using EAST, breaking links between nodes in the social and task networks to identify resultant risk and associated mitigation strategies. These strategies could be fed into the process as stories, or they could inform the design additional interface design to support risk reduction. The links that can break would also be of use to engineers, as this would inform them about what aspects of the system require monitoring, and how to determine the health of links between nodes. Other methods could be applied in this way, or they could be used to identify possible failovers that would allow end users to work with the system. Salmon, Carden and Stevens (2018) used link breaking on an Abstraction Hierarchy to identify strategies that would disrupt terrorist cells at the various levels used utilised by WDA. This approach could be used to identify what capability would remain available if a link were to break, allowing designs to account for this. The approach could also be used to identify capability that should never break, to ensure that safety risks are kept as low as possible. Jenkins *et al.* (2010a), citing Hopkins (2000), used link breaking a method for Accimap model validation, stating that breaking any one link could have averted the accident being investigated. This could be used to identify key failure points that could be broken using the software being designed, and account for them in the code and/or interface. For example, the Accimap of McCabe, Baber and Stone (2020) identified that the submarine was operating fast and deep in a busy, crowded area, which led to all vessels without trawl noise being classified as merchant vessels. Knowing this, future Sonar software could incorporate functionality that prompts operators to re-evaluate contacts for trawl noise, or other fishing vessel characteristics, at regular time intervals proportionate to the number of contacts held (more contacts requiring more frequent re-evaluation).

Software requires documenting, and the design process could be expanded to generate this documentation. This could be especially advantageous as it would be based on the same source material the software design is, saving time and financial resources. The different approach for applying CWA to HMIs detailed in Section 5.2.1 is a primary example of this, as the resultant Abstraction Hierarchy reflects the interfaces. Thus, they could be deployed as-is for documentation, alongside a counterpart hierarchy that details the work domain. Consequently, it would be possible for end-users to understand the system and utilise the constraints-based approach of CWA to understand how they might approach tasks differently. The approach could also be used for designing updated procedures for new software, such as using EAST task networks to identify different strategies employed by end-users during testing, and to incorporate effective strategies into procedures. Finally, Accimap could be used to identify applicable legislative requirements for software, which might not be readily apparent, allowing engineers to design for, and demonstrate, compliance.

In summary, there are a variety of ways in which the design process proposed in this research could be extended to incorporate other methods that examine complex socio-technical systems. While there is implicit compatibility with almost all methods by creating user stories from the insights they generate, there is also extensive possibility for more explicit integrations. The focus of these in this section was to promote making facets of creating systems for the systems more readily apparent, so that designed software is appropriately robust. This was to address a potential loss of richness from the various Human Factors methods available (Baxter and Sommerville, 2010; Dhukaram and Baber, 2016), which could see key design requirements omitted. As argued in Section 8.3, directly providing software engineers with human factors outputs could be more beneficial over translating to a software engineering format, such as Unified Modelling Language diagrams. Thus, any future work on the design process should seek to utilise the strengths of each method to directly inform software construction.

11.4.3 Constraint Taxonomy

The constraint taxonomy presented in Chapter 3 was framed as a tool to drive the CWA analysis process, acting as a guide for practitioners to follow when conducting their analysis. However, it is believed that there is potential to use the constraints taxonomy in a wide variety of contexts and applications, which could be explored by future work. A significant proportion of these applications have parallels with the ecological principle of inferring distal variables from proximal variables (King, Read and Salmon, 2022), therefore making visible the invisible (Vicente and Rasmussen, 1987); category linking allows known and observed (proximal) categories to be mapped to those that might not be apparent (distal). This is interesting, as it hints at the possibility for using ecological principles for human factors analyses, especially as King, Read and Salmon (2022) have demonstrated that EID remains effective in a static (i.e., non-interactive) context, which most analysis artefacts take the form of. Though an intriguing possibility, this section will concentrate on solely on how the constraints taxonomy could be utilised.

CWA is iterative, often requiring multiple revisions to arrive at a “completed” artefact (Naikar, Hopcroft and Moylan, 2005; Stanton *et al.*, 2017a). These artefacts represent the constraints in the work domain as they are aware of, requiring validation from subject matter experts. However, knowledge of constraints might be limited by utilisation of familiar routines that mask the underlying justification, and rationalisations or explanations of the system that are incongruent with actual constraints (Rasmussen, Pejtersen and Goodstein, 1994; Naikar, Hopcroft and Moylan, 2005). The taxonomy could be used to address this in existing CWAs by retrospectively coding nodes from each of stage, using the constraint categories as actors, similar to the approach taken by SOCA.

Practitioners could then check off existing categories, and use either of the approaches detailed in Section 3.5 to explore missing categories, contingent on available time and required analysis scope.

The coding could also allow social network analysis to be completed on the artefacts, such as the abstraction hierarchy, allowing comparison with the taxonomy metrics (Section 3.3). In turn, this could be used to identify areas that might need further exploration, such as if a category has a high sociometric status in the taxonomy, but has a much lower status in the abstraction hierarchy. This would be applied with appropriate recognition that domains can differ, and that taxonomy is constructed as a holistic overview of constraints at all points of the continuum of behavioural constraints, which categorises domains as belonging between intent- and law-driven (Rasmussen, Pejtersen and Goodstein, 1994; Bennett and Flach, 2019).

Approaching this from a different standpoint to address this, coded CWA analyses could also be used to create multiple versions of the taxonomy for different points along the continuum, making it more targeted and appropriate for use when the type of domain is already known. This could also provide validation of a domain's categorisation. For example, intent-driven domains might have higher levels of "Decision Making, Goals, and Action" category constraints, as operators can make more decisions when compared to a law-driven domain. Another potential application could be filtering the abstraction hierarchy to understand how much each category constrains the work domain, and where changes could be made to shape remove barriers to effective system performance. For example, if an abstraction hierarchy is predominantly formed of "Technology" category constraints, then this could be indicative that improved technology is required to improve the action space in which users are provided with to operate within their domain.

More generally, the principle of coding by categories could be employed for multiple different methods that examine complex sociotechnical systems, where the composition of constraints would be useful information. However, not all methods deal with constraints, which are a distinguishing factor of CWA. Instead, the notion of constraints could be abstracted into characteristics. Vicente (1999b) and Section 3.2.3 defined constraints as relationships between, or limits on, behaviour; removing the concept of these limitations or relationships would leave the behaviour, which in turn characterises a system.

This could be applied to previous accident reports to categorise the type of accident that occurred from the primary cause and identify other potential causal factors. For example, the USS Greeneville crash (NTSB; 2001), discussed in Section 1.3, could be categorised as a primarily "Actors and Agents" as the primary cause was inadequate interaction between senior command team members. "Actors and Agents" is connected to "Decision Making, Goals, and Action" in the taxonomy, which suggests that factors in this category should be considered. This is confirmed by the NTSB report stating that a further contributory cause was a failure to manage the civilian visitors sufficiently so that they did not impede operations. While identifying additional causal factors in this manner might not be possible in all situations, it could help to direct focus when examining the causal chain(s) of complex accidents. This could also help address instances where accident reports are reductive owing to a specific emphasis, such as road-accident investigations being driver-focused (Newnam *et al.*, 2017; Salmon and Read, 2019), by prompting investigators to consider factors outside the context of the accident's immediate presentation.

Alternatively, investigators could use the taxonomy to examine factors affecting accidents as part of their method from an investigation's beginning, working through the taxonomy to elicit categorical details. Elicited factors could be category coded and added to an AcciMap, or an existing AcciMap could have codes retrospectively added. As with the abstraction hierarchy, this would structure data collection, and add an additional dimension to the completed analysis. Consequently, social network analysis could be applied, and the results used to understand the categories with the most sociometric status as priorities for recommendations aimed at preventing similar future accidents. This could be cross-referenced with their level in the AcciMap to understand what strategies might be most effective in each context. The work of Newnam *et al.* (2017), a combined AcciMap of 21 road freight crashes between 2004 – 2014, also provides an interesting avenue of exploration. With a temporal dimension, there is potential to map the proportion of each category over time to assess whether contributory factors are being reduced overall (i.e., a safer system), or whether causation is being "pushed" around the taxonomy (i.e., an altered risk-profile). That is, evaluating whether the mitigatory measures applied to one category improved safety, or has the risk been pushed to adjacent categories as an emergent property of making the changes.

Such an approach could be applied to safety-critical domains shaped by catastrophic events as a proactive preventative measure. These domains include nuclear power plants (e.g., Three Mile Island), rail travel (e.g., Ladbrooke Grove), and marine travel (e.g., Titanic). Given the extensive focus on analysing these "milestone" accidents in detail, it is unlikely that any new insights could be derived. However, this focus will have created an extensive library of analyses, both during routine and incident situations, that could be analysed to understand how these domains are changing across time. These could be used to understand the changing composition of the domain, which might indicate where routine operation and/or safety improvement efforts should be placed. For example, the working knowledge of a domain might suggest improvements in one category, perhaps due to an overall reductionist approach as identified by Newnam *et al.* (2017), although this might be contradicted by this type of analysis. Consequently, measures can be proactively deployed to holistically enhance work and safety within systems.

However, any wholesale utilisation should be proceeded by validation of the taxonomy as discussed in Section 11.3.2, namely improving the interrater reliability. Categorisation of CWA analyses could also assist with this, as they are the pre-cursors to the EID designs of which the literature was assessed to construct the taxonomy but were not included in the taxonomy version presented in this research due to availability. This is recognised by McIlroy and Stanton (2015b) in that the outcome of CWA is often too voluminous for inclusion in journals. Consequently, incorporating data from categorisation of source CWAs could yield a completely different taxonomy, especially if only the most prevalent constraints are reported by authors in journals. Validation is anticipated to change the social network characteristics of the taxonomy, but the approach of using the taxonomy to drive analysis and yield design insights remains consistent.

11.5 Closing Remarks

Current submarine control rooms are capable, but this does not preclude further improvements. Future requirements, such as more data or new sensors, for platforms are challenging contemporary norms, which are an evolutionary product of historic constraints. These challenges must be addressed to maintain effective performance. An evolutionary approach, combined with Royal Navy professionalism, has consistently yielded adequate performance. However, an increasingly complex global maritime environment and ever-increasing requirements is prompting assessment of where improvements can be made. One area of improvement are the HMIs. Their design is vital for effective interaction between a boats advanced technology and highly trained crew. Should this interaction be hindered, there exists a possibility for adverse situations, such as the USS Greeneville or RN submarine accidents. However, despite their importance, they are still designed with legacy constraint holdovers, a product of their evolutionary lineage. This thesis aimed to understand if EID was a suitable design paradigm, documenting the design and analysis of an EID called GIST to explore this. It is hoped that the demonstration of the benefits GIST provided as an EID, combined with a detailed background on the impetus to change and suggestions on how to enact such change in line with modern software principles, will serve to make change in future submarine control rooms.

References

Abrahamsson, P.a.S., Outi and Ronkainen, Jussi and Warsta, Juhani (2002) *Agile Software Development Methods: Review and Analysis*. Available at: https://arxiv.org/abs/1709.08439.

Abras, C., Maloney-Krichmar, D. and Preece, J. (2004) 'User-Centered Design', *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, 37(4), pp. 445-456.

Agnisarman, S., Madathil, K.C. and Bertrand, J. (2019) 'Context-based visual aids to support the situation awareness of field engineers conducting windstorm risk surveys', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, pp. 1898-1902.

Ahlstrom, U. (2005) 'Work domain analysis for air traffic controller weather displays', *J Safety Res*, 36(2), pp. 159-69.

Ahvenjärvi, S. (2016) 'The human element and autonomous ships', *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, 10(3), pp. 517--521.

Aidala, V.J. (1979) 'Kalman Filter Behavior in Bearings-Only Tracking Applications', *Ieee Transactions on Aerospace and Electronic Systems*, 15(1), pp. 29-39.

Alessi, S.M. (1988) 'Fidelity in the Design of Instructional Simulations', *Journal of Computer-Based Instruction*, 15(2), pp. 40-47.

Alexander, A.L. *et al.* (2005) 'From gaming to training: A review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in pc-based simulations and games', *DARWARS Training Impact Group*, 5, pp. 1-14.

Allen, J.A., Hays, R.T. and Buffardi, L.C. (1986) 'Maintenance Training Simulator Fidelity and Individual-Differences in Transfer of Training', *Human Factors*, 28(5), pp. 497-509.

Amelink, M.H.J. *et al.* (2005) 'Theoretical foundations for a total energy-based perspective flight-path display', *International Journal of Aviation Psychology*, 15(3), pp. 205-231.

Annabattula, J. *et al.* (2015) 'Advanced submarine integrated weapon control system', *Indian Journal of Science and Technology*, 8(35), pp. 1-3.

Annett, J. and Duncan, K.D. (1967) 'Task analysis and training design', *Occup. Psychol.*, 41, pp. 211-221.

References

Annett, J. and Stanton, N.A. (2000) *Task analysis*. CRC Press.

Anokhin, A., Ivkin, A. and Dorokhovich, S. (2018) 'Application of ecological interface design in nuclear power plant (NPP) operator support system', *Nuclear Engineering and Technology*, 50(4), pp. 619-626.

Araujo, D., Davids, K. and Serpa, S. (2005) 'An ecological approach to expertise effects in decision-making in a simulated sailing regatta', *Psychology of Sport and Exercise*, 6(6), pp. 671-692.

Arrabito, G.R., Cooke, B.E. and McFadden, S.M. (2005) 'Recommendations for enhancing the role of the auditory modality for processing sonar data', *Applied Acoustics*, 66(8), pp. 986-1005.

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

Asplin, R.G. and Christensson, C.G. (1988) 'A new generation side scan sonar', *OCEANS'88. A Partnership of Marine Interests. Proceedings*. Baltimore, MD, USA. IEEE, pp. 329-334.

Atlas Elektronik (2016a) *AI&R Sonars Acoustic Intercept and Ranging*. Available at: https://web.archive.org/web/20170703045207/https://www.atlas-elektronik.com/what-we-do/surface-combatant-systems/air-sonars/ (Accessed: 13/11/2023).

Atlas Elektronik (2016b) *ISUS® 100 Combat System*. Available at: https://web.archive.org/web/20160816020815/https://www.atlas-elektronik.com/what-we-do/submarine-systems/isusr-100/ (Accessed: 13/11/2023).

Baber, C., Stanton, N.A. and Houghton, R. (2017) 'Deriving and analysing social networks from SOCA-CAT diagrams', in *Cognitive Work Analysis*. CRC Press, pp. 387-402.

Baddeley, A. (2010) 'Working memory', *Current Biology*, 20(4), pp. R136-R140.

BAE Systems (2015) *Submarine Command System Next Generation*. Available at: http://www.baesystems.com/en/product/submarine-command-system-next-generation (Accessed: 13/11/2023).

BAE Systems and Launderdale, N. (2016) *Artful submarine fires first torpedo using new Common Combat System*. Available at: http://www.baesystems.com/en-uk/article/artful-submarine-fires-first-torpedo-using-new-common-combat-system (Accessed: 13/11/2023).

Baggeroer, A.B. (2005) 'Sonar arrays and array processing', *Review of Progress in Quantitative Nondestructive Evaluation, Vols 24A and 24B*, 760(1), pp. 3-24.

Bainbridge, L. (1983) 'Ironies of automation', *Automatica*, 19(6), pp. 775-779.

Bangor, A., Kortum, P.T. and Miller, J.T. (2008) 'An empirical evaluation of the System Usability Scale', *International Journal of Human-Computer Interaction*, 24(6), pp. 574-594.

Bass, E.J. (2014) 'IEEE Transactions On Human–Machine Systems: Year in Review for 2013', *IEEE Transactions on Human-Machine Systems*, 44(1), pp. 1-12.

Bateman, S. (2011) 'Perils of the Deep: The Dangers of Submarine Proliferation in the Seas of East Asia', *Asian Security*, 7(1), pp. 61-84.

Baxter, G. and Sommerville, I. (2010) 'Socio-technical systems: From design methods to systems engineering', *Interacting with Computers*, 23(1), pp. 4-17.

Beanland, V. *et al.* (2018) 'Challenging conventional rural rail level crossing design: Evaluating three new systems thinking-based designs in a driving simulator', *Safety Science*, 110, pp. 100-114.

Beevis, D., Vicente, K. and Dinadis, N. (1998) 'An exploratory application of ecological interface design to aircraft systems', *Proceedings of the NATO RTO Symposium on Collaborative Crew Performance in Complex Operational Systems*. DTIC Document.

Behymer, K.J. (2017) *Interface Design for the Supervisory Control of Multiple Heterogeneous Unmanned Vehicles*. Wright State University. Available at: http://rave.ohiolink.edu/etdc/view?acc_num=wright1495016420685509.

Bennett, K. *et al.* (2011) *The Dynamics of Situation Awareness*. Display and Interface Design: Subtle Science, Exact Art.

Bennett, K.B. (2017) 'Ecological interface design and system safety: One facet of Rasmussen's legacy', *Applied Ergonomics*, 59(Pt B), pp. 625-636.

Bennett, K.B. (2014) 'Ecological interface design: Military C2 and computer network defense', *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE. doi: 10.1109/smc.2014.6973931. Available at: http://dx.doi.org/10.1109/smc.2014.6973931.

Bennett, K.B., Bryant, A. and Sushereba, C. (2018) 'Ecological Interface Design for Computer Network Defense', *Human Factors*, 60(5), pp. 610-625.

Bennett, K.B. and Flach, J. (2019) 'Ecological Interface Design: Thirty-Plus Years of Refinement, Progress, and Potential', *Human Factors*, 61(4), pp. 513-525.

Bennett, K.B. and Flach, J.M. (2011) *Display and interface design: Subtle science, exact art*. CRC Press.

References

Bennett, K.B. and Flach, J.M. (1992) 'Graphical Displays: Implications for Divided Attention, Focused Attention, and Problem Solving', *Human Factors*, 34.

Bennett, K.B., Payne, M. and Walters, B. (2005) 'An evaluation of a "time tunnel" display format for the presentation of temporal information', *Human Factors*, 47.

Bennett, K.B., Posey, S.M. and Shattuck, L.G. (2008) 'Ecological interface design for military command and control', *Journal of Cognitive Engineering and Decision Making*, 2(4), pp. 349-385.

Benta, M.I. (2003) 'Agna 2.1 User Manual'.

Bisantz, A.M. *et al.* (2001) 'Integrating Cognitive Analyses into a Large Scale System Design Process', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 45(4), pp. 434-438.

Bisantz, A.M. *et al.* (2003) 'Integrating Cognitive Analyses into a Large Scale System Design Process', *International Journal of Human-Computer Studies*, 58(2), pp. 177-206.

Bj, L. (2011) 'Developments in sonar and array technologies', *Underwater Technology (UT), 2011 IEEE Symposium on and 2011 Workshop on Scientific Use of Submarine Cables and Related Technologies (SSC)*. Tokyo, Japan. IEEE, pp. 1-11.

Borst, C. *et al.* (2017) 'Ecological interface design: supporting fault diagnosis of automated advice in a supervisory air traffic control task', *Cognition Technology & Work*, 19(4), pp. 545-560.

Borst, C., Flach, J.M. and Ellerbroek, J. (2015) 'Beyond Ecological Interface Design: Lessons From Concerns and Misconceptions', *IEEE Transactions on Human-Machine Systems*, 45(2), pp. 164-175.

Borst, C. *et al.* (2008) 'Ecological approach to support pilot terrain awareness after total engine failure', *Journal of Aircraft*, 45(1), pp. 159-171.

Borst, C. *et al.* (2006) 'Ecological interface design for terrain awareness', *International Journal of Aviation Psychology*, 16(4), pp. 375-400.

Submarine (2009) Directed by Bosner, S. and Oxley, P. Washington, DC, USA: National Geographic International.

Bourque, P. and Fairley, R.E. (2014) *Guide to the Software Engineering Body of Knowledge (SWEBOK)*. 3.0 edn.: IEEE Computer Society.

Bowden, K. and Grosse, E. (2011) *vVICTORIA Console Development: Design and Fabrication of VICTORIA Console Emulations*. Center, D.T.I. Available at: http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA598423.

Boy, G.A. (2014) 'From automation to tangible interactive objects', *Annual Reviews in Control*, 38(1), pp. 1-11.

Breton, R. and Bossé, É. (2003) *The cognitive costs and benefits of automation*.

Brewer, W.F. (1987) 'Schemas versus mental models in human memory', in Morris, P. (ed.) *Modelling Cognition*. Chichester, UK: Wiley, pp. 187-197.

Brolese, A. (2005) 'Graphic representations of uncertainty within sub-surface environments', *Proceedings of the Annual Meeting of the Cognitive Science Society*, 27(27).

Brooke, J. (1996) 'SUS - A quick and dirty usability scale', *Usability evaluation in industry*, 189(194), pp. 4-7.

Brunswik, E. (1956) *Perception and the Representative Design of Psychological Experiments*. Berkeley, CA, USA: University of California Press.

Brunswik, E. (1957) 'Scope and aspects of the cognitive problem', in Gruber, H., Jessor, R. and Hammond, K.R. (eds.) *Contemporary approaches to cognition*. Cambridge: Harvard University Press, pp. 5-31.

Bruseberg, A. (2008) 'Human Views for MODAF as a Bridge Between Human Factors Integration and Systems Engineering', *Journal of Cognitive Engineering and Decision Making*, 2(3), pp. 220-248.

Bullemer, P.T. and Nimmo, I. (1994) 'Understanding and supporting abnormal situation management in industrial process control environments: a new approach to training', *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*. pp. 391-396 vol.1. doi: 10.1109/ICSMC.1994.399870.

Burcher, R. and Rydill, L.J. (1995) *Concepts in Submarine Design*. Cambridge, UK: Cambridge University Press.

Burns, C.M. (2012) 'Cognitive Work Analysis: New Dimensions', *IFIP Working Conference on Human Work Interaction Design*. Copenhagen, Denmark. Acadmic 2019 05 22: Springer, pp. 1-11. doi: 10.1007/978-3-642-41145-8_1. Available at: https://link.springer.com/chapter/10.1007/978-3-642-41145-8_1#citeas.

Burns, C.M., Bisantz, A.M. and Roth, E.M. (2004) 'Lessons from a comparison of work domain models: representational choices and their implications', *Human Factors*, 46(4), pp. 711-27.

References

Burns, C.M., Bryant, D.J. and Chalmers, B.A. (2005) 'Boundary, purpose, and values in work-domain models: Models of naval command and control', *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 35(5), pp. 603-616.

Burns, C.M., Bryant, D.J. and Chalmers, B.A. (2000) 'A work domain model to support shipboard command and control', *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*. pp. 2228-2233 vol.3. doi: 10.1109/ICSMC.2000.886447.

Burns, C.M. and Hajdukiewicz, J. (2004) *Ecological interface design*. Boca Raton, USA: CRC Press.

Burns, C.M., Kuo, J. and Ng, S. (2003) 'Ecological interface design: a new approach for visualizing network management', *Computer Networks*, 43(3), pp. 369-388.

Burns, C.M. *et al.* (2008) 'Evaluation of ecological interface design for nuclear process control: situation awareness effects', *Human Factors*, 50(4), pp. 663-79.

Calhoun, G. *et al.* (2013) 'Adaptable Automation Interface for Multi-Unmanned Aerial Systems Control: Preliminary Usability Evaluation', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1), pp. 26-30.

Calhoun, G.L. *et al.* (2018) 'Human-autonomy teaming interface design considerations for multi-unmanned vehicle control', *Theoretical Issues in Ergonomics Science*, 19(3), pp. 321-352.

Sonar Operator (2013) Directed by Canadian Armed Forces. Ottawa, Ontario, Canada.

Carrigan, G.P. (2009) *The design of an intelligent decision support tool for submarine commanders*. Massachusetts Institute of Technology.

Chalmers, B.A., Easter, J.R. and Potter, S.S. (2000) 'Decision-centred visualisations for tactical decision support on a modern frigate', *Command and Control Research and Technology Symposium, Making Information Superiority Happen*. Monterey, CA, USA. Citeseer.

Chammas, A., Quaresma, M. and Mont'Alvão, C. (2015) 'A Closer Look on the User Centred Design', *Procedia Manufacturing*, 3, pp. 5397-5404.

Chen, H.-Y. and Burns, C.M. (2007) 'Work Domain Analysis for the Interface Design of a Sonobuoy System', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. pp. 283-287. doi: 10.1177/154193120705100429. Available at: http://journals.sagepub.com/doi/abs/10.1177/154193120705100429.

Chen, H.-Y.W. (2007) *Interface Design for Sonobuoy System*.

Christoffersen, K., Hunter, C.N. and Vicente, K.J. (1996) 'A Longitudinal Study of the Effects of Ecological Interface Design on Skill Acquisition', *Human Factors*, 38(3), pp. 523-541.

Christoffersen, K. *et al.* (1994) 'Ecological Interface Design and Operator Competencies - A (Very) Long-Term Study', *1994 Ieee International Conference on Systems, Man, and Cybernetics - Humans, Information and Technology, Vols 1-3*. pp. 1398-1403. doi: 10.1109/ICSMC.1994.400041. Available at: <Go to ISI>://WOS:A1994BC41S00239.

Clare, A.S. *et al.* (2012) 'Operator Object Function Guidance for a Real-Time Unmanned Vehicle Scheduling Algorithm', *Journal of Aerospace Computing, Information, and Communication*, 9(4), pp. 161-173.

Clarke, J.A. (1999) *A proposed submarine electronic chart display and information system*. Department of Geodesy and Geomatics Engineering, University of New Brunswick.

Clautice, W.G. (1978) 'Submarine Navigation', *Navigation*, 25(1), pp. 34-39.

Coll, P.F. (1994) *Target motion analysis from a diesel submarine's perspective*. Monterey, California. Naval Postgraduate School.

Coll, R. and Wingertsman, J.C. (1990) 'The effect of screen complexity on user preference and performance', *International Journal of Human-Computer Interaction*, 2(3), pp. 255-265.

Cravens, D.G. (2021) *Ecological Interface Design for Flexible Manufacturing Systems: An Empirical Assessment of Direct Perception and Direct Manipulation in the Interface*. Master of Science. Wright State University. Available at: https://corescholar.libraries.wright.edu/etd_all/2552/.

Cullen, W.D. (2001) *The Ladbroke Grove Rail Inquiry: Part 1 Report*. London, UK: Government, U. Available at: https://www.jesip.org.uk/wp-content/uploads/2022/03/Ladbroke-Grove-Rail-Inquiry-Report-Part-1.pdf.

Cunningham, A. and Thomas, B. (2005) 'Target motion analysis visualisation', *proceedings of the 2005 Asia-Pacific symposium on Information visualisation-Volume 45*. Australian Computer Society, Inc., pp. 81-90.

Cushing, S. (1994) *Fatal words: Communication clashes and aircraft crashes*. Chicago, Illinois, United States: University of Chicago Press.

Danczyk, J. *et al.* (2015) 'Visualizing Constraints and Awareness of Submarine Maneuverability and Detectability', *6th International Conference on Applied Human Factors and Ergonomics (Ahfe 2015)* *and the Affiliated Conferences, Ahfe 2015*. pp. 5155-5160. doi: 10.1016/j.promfg.2015.07.548. Available at: <Go to ISI>://WOS:000383740305040.

Davies, A. (2013) *Graph of the week – why (fleet) size matters*. Available at: https://www.aspistrategist.org.au/graph-of-the-week-why-fleet-size-matters/ (Accessed: 13/11/2023).

Davies, T.C., Burns, C.M. and Pinder, S.D. (2006) 'Using ecological interface design to develop an auditory interface for visually impaired travellers', in *Proceedings of the 20th conference of the computer-human interaction special interest group (CHISIG) of Australia on Computer-human interaction: design: activities, artefacts and environments - OZCHI '06*. ACM Press.

Dawe, R.L. and Galbreath, E.R. (1997) 'Detection Enchancement Using Colour For Sonar Displays', *Fifth International Congress on Sound and Vibration*. Adelaide, South Australia. Australian Acoustical Society, p. 8. Available at: https://www.acoustics.asn.au/conference_proceedings/ICSVS-1997/pdf/scan/sv970033.pdf.

de Moura, N.N., de Seixas, J.M. and Ramos, R. (2011) 'Passive Sonar Signal Detection and Classification Based on Independent Component Analysis', in *Soanr Systems*. London, UK: INTECH Open Access Publisher.

DeAngelis, C.M. and Green, R.W. (1992) 'Constructing neural networks for contact tracking', *Neural Networks for Signal Processing [1992] II., Proceedings of the 1992 IEEE-SP Workshop*. Helsingoer, Denmark: IEEE, pp. 560-569. doi: 10.1109/NNSP.1992.253656. Available at: https://ieeexplore.ieee.org/document/253656.

Defence Equipment and Support (2010) 'Combat Systems for the whole fleet', *Desider*(29), p. 2.

Department of National Defence (2001) *Leadmark: The Navy's Strategy for 2020"*. Ottawa, Canada: Department of National Defence.

Dhukaram, A.V. (2016) *CISDA Development Process for decision aids to support self-care decision making*. PhD. University of Birmingham. Available at: https://etheses.bham.ac.uk/id/eprint/6618/.

Dhukaram, A.V. and Baber, C. (2016) 'A Systematic Approach for Developing Decision Aids: From Cognitive Work Analysis to Prototype Design and Development', *Systems Engineering*, 19(2), pp. 79-100.

Diehl, M. and Stroebe, W. (1987) 'Productivity loss in brainstorming groups: Toward the solution of a riddle', *Journal of personality and social psychology*, 53(3), p. 497.

Dinadis, N. and Vicente, K.J. (1999) 'Designing functional visualizations for aircraft systems status displays', *International Journal of Aviation Psychology*, 9(3), pp. 241-269.

Dinadis, N. and Vicente, K.J. (1996) 'Ecological interface design for a power plant feedwater subsystem', *IEEE Transactions on Nuclear Science*, 43(1), pp. 266-277.

Doll, T.J. and Hanna, T.E. (1989) 'Enhanced Detection with Bimodal Sonar Displays', *Human Factors*, 31(5), pp. 539-550.

Dominguez, C. *et al.* (2006) 'Design Directions for Support of Submarine Commanding Officer Decision Making', *Undersea HSI Symposium*. Mystic, CT. p. 8.

Dreyer, L.O. and Oltedal, H.A. (2019) 'Safety challenges for maritime autonomous surface ships: a systematic review', *The Third Conference on Maritime Human Factors. Haugesund*.

Driskell, J.E. and Mullen, B. (2004) 'Social network analysis', in *Handbook of human factors and ergonomics methods*. CRC Press, pp. 565-571.

Drivalou, S. (2005) 'Supporting critical operational conditions in an electricity distribution control room through ecological interfaces', *Proceedings of the 2005 annual conference on European association of cognitive ergonomics*. University of Athens, pp. 263-270.

Drivalou, S. and Marmaras, N. (2009) 'Supporting skill-, rule-, and knowledge-based behaviour through an ecological interface: An industry-scale application', *International Journal of Industrial Ergonomics*, 39(6), pp. 947-965.

Drumheller, K. and Benoit, W.L. (2004) 'USS Greeneville collides with Japan’s Ehime Maru: Cultural issues in image repair discourse', *Public Relations Review*, 30(2), pp. 177-185.

Dry, M. *et al.* (2005) *Psychological Implications for Submarine Display Design*.

Dryer, J.E. (2002) *Robust autonomous vehicle navigation*. Massachusetts Institute of Technology.

Durugbo, C. (2012) 'Work domain analysis for enhancing collaborations: a study of the management of microsystems design', *Ergonomics*, 55(6), pp. 603-20.

Duryea, D.M., Lindstrom, C.E. and Sayegh, R. (2008) 'Submarine imaging systems: developing improved capabilities and technologies', *SPIE Defense and Security Symposium*. Orlando, FL, USA. International Society for Optics and Photonics, pp. 69430X-69430X-8. Available at: http://dx.doi.org/10.1117/12.783814.

References

Effken, J.A. (2006) 'Improving clinical decision making through ecological interfaces', *Ecological Psychology*, 18(4), pp. 283-318.

Effken, J.A. (2001) 'Informational basis for expert intuition', *Journal of Advanced Nursing*, 34(2), pp. 246-55.

Effken, J.A., Kim, N.G. and Shaw, R.E. (1997) 'Making the constraints visible: testing the ecological approach to interface design', *Ergonomics*, 40(1), pp. 1-27.

Ellejmi, M. *et al.* (2018) 'Evaluation of augmented reality tools for the provision of tower air traffic control using an ecological interface design', *Modeling and Simulation Technologies Conference, 2018*. American Institute of Aeronautics and Astronautics Inc, AIAA. doi: 10.2514/6.2018-2939. Available at: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85051666174&doi=10.2514%2f6.2018-2939&partnerID=40&md5=f53b96362e83e1e038f1bfe0cdc268d2 (Accessed: 13/11/2023).

Ellerbroek, J. *et al.* (2013a) 'Experimental Evaluation of a Coplanar Airborne Separation Display', *IEEE Transactions on Human-Machine Systems*, 43(3), pp. 290-301.

Ellerbroek, J. *et al.* (2013b) 'Design of a Coplanar Airborne Separation Display', *IEEE Transactions on Human-Machine Systems*, 43(3), pp. 277-289.

Ellerbroek, J. *et al.* (2011) 'Design of an Airborne Three-Dimensional Separation Assistance Display', *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 41(5), pp. 863-875.

Emery, F.E. *et al.* (1960) 'Socio-Technical Systems', *Management Science Models and Techniques*, 2, pp. 83-97.

Surface Navy Combat Systems Engineering Strategy (2010) Naval Postgraduate School.

Endsley, M.R. (1995) 'Toward a Theory of Situation Awareness in Dynamic-Systems', *Human Factors*, 37(1), pp. 32-64.

Ericsson, K.A. (2009) *Development of professional expertise: Toward measurement of expert performance and design of optimal learning environments*. Cambridge University Press.

Estock, J. *et al.* (2006) 'A model-based approach to simulator fidelity and training effectiveness', *2006 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) Proceedings*.

Fallon, M. (2015) *UK Defence in Numbers*. London, UK: Ministry of Defence. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/467482/20151013-UK_Defence_in_Numbers_screen.pdf (Accessed: 13/11/2023).

Favarò, F.M. *et al.* (2013) 'Software contributions to aircraft adverse events: Case studies and analyses of recurrent accident patterns and failure mechanisms', *Reliability Engineering & System Safety*, 113, pp. 131-142.

Fay, D., Roberts, A.P. and Stanton, N.A. (2019) 'All at Sea with User Interfaces: From Evolutionary to Ecological Design for Submarine Combat Systems', *Theoretical Issues in Ergonomics Science*, 20(5), pp. 632-658.

Fay, D., Roberts, A.P.J. and Stanton, N.A. (2020) 'Interfaces with Legs? Documenting the Design Sprint of Prototype Future Submarine Control Room User Interfaces', *AHFE 2019*. Washington, D.C, USA. Springer International Publishing, pp. 669-680.

Fay, D., Stanton, N. and Roberts, A. (2018) 'Streamlining experimental processes using bespoke software', *Ergonomics & Human Factors 2018*. Birmingham. Available at: https://eprints.soton.ac.uk/429470/.

Fay, D., Stanton, N.A. and Roberts, A. (2017) 'Designing New Interfaces for Submarines: From Cognitive Work Analysis to Ecological Interface Design', *AHFE 2016*. Orlando, FL, USA. Springer International Publishing, pp. 413-425. doi: 10.1007/978-3-319-41682-3_35. Available at: http://dx.doi.org/10.1007/978-3-319-41682-3_35 (Accessed: 13/11/2023).

Feigh, K.M. *et al.* (2018) 'Shifting role for human factors in an 'unmanned' era', *Theoretical Issues in Ergonomics Science*, 19(4), pp. 389-405.

Field, A. (2018) *Discovering Statistics Using IBM SPSS Statistics*. 5th edn. London: SAGE Publications Ltd.

Fillinger, L. *et al.* (2010) 'Towards a passive acoustic underwater system for protecting harbours against intruders', *Waterside Security Conference (WSS), 2010 International*. Carrara, Italy. IEEE, pp. 1-7.

Flach, J. (2017) 'Supporting productive thinking: The semiotic context for Cognitive Systems Engineering (CSE)', *Applied Ergonomics*, 59, pp. 612-624.

Flach, J.M. and Dominguez, C.O. (1995) 'USE - Centered Design: Integrating the User, Instrument, and Goal', *Ergonomics in Design: The Quarterly of Human Factors Applications*, 3(3), pp. 19-24.

Flach, J.M., Stappers, P.J. and Voorhorst, F.A. (2017) 'Beyond Affordances: Closing the Generalization Gap Between Design and Cognitive Science', *Design Issues*, 33(1), pp. 76-89.

Fougnie, D. *et al.* (2015) 'Working memory storage is intrinsically domain specific', *Journal of Experimental Psychology: General*, 144, pp. 30-47.

Genç, H. (2010) 'A new solution approach for the bearing only target tracking problem', *Soft Computing Applications (SOFA), 2010 4th International Workshop on*. IEEE, pp. 95-100.

General Dynamics (2014) *Tactical Control Systems: Delivering Open System Flexibility For Rapid Technology Insertion*. Available at: https://gdmissionsystems.com/maritime-strategic/submarine-systems/tactical-control-systems/ (Accessed: 13/11/2023).

Giang, W.C.W. *et al.* (2010) *Multimodal Interfaces: Literature Review of Ecological Interface Design, Multimodal Perception and Attention, and Intelligent Adaptive Multimodal Interfaces*.

Gibson, J.J. (1979) *The Ecological Approach To Visual Perception*. Boston, MA, USA: Houghton Mifflin.

Gibson, W.H. *et al.* (2005) 'A taxonomy of human communication errors and application to railway track maintenance', *Cognition, Technology & Work*, 8(1), pp. 57-66.

Goettl, B.P. (2006) 'Training and Acquisition of Complex Tasks', in Karwowski, W. (ed.) *International Encyclopedia of Ergonomics and Human Factors, Second Edition - 3 Volume Set*. 2nd edn. Boca Raton, Florida, USA: CRC Press, Inc.

Google Ventures (2016) *The Design Sprint*. Available at: https://www.gv.com/sprint/ (Accessed: 13/11/2023).

Gosling, P. (2008) *Evolutionary Trends In UK Sonar*. London, UK. Available at: https://web.archive.org/web/20141119233616/https://www.rusi.org/downloads/assets/12gosling.pdf (Accessed: 13/11/2023).

Grier, R.A. (2013) 'The Potential Utility of the System Usability Scale in U.S. Military Acquisition', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1), pp. 206-209.

Gross, D. and Freeman, R. (1997) 'Measuring fidelity differentials in HLA simulations', *The 1997 Fall Simulation Interoperability Workshop*. Orlando, USA. Simulation Interoperability Standards Organization.

Grudin, J. (1989) 'The Case against User Interface Consistency', *Communications of the Acm*, 32(10), pp. 1164-1173.

Guertin, N.H. and Miller, R.W. (1998) 'A-RCI-The Right Way to Submarine Superiority', *Naval Engineers Journal*, 110(2), pp. 21-33.

Hajdukiewicz, J. and Burns, C. (2004) 'Strategies for Bridging the Gap between Analysis and Design for Ecological Interface Design', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Inc, pp. 479-483. doi: 10.1177/154193120404800344. Available at: https://doi.org/10.1177/154193120404800344 (Accessed: 13/11/2023).

Hall, D.S. and Miller, N.L. (2009) *Raptor: An empirical evaluation of an ecological interface designed to increase war fighter cognitive performance*. Master of Science in Human Systems Integration. Naval Postgraduate School. Available at: https://calhoun.nps.edu/handle/10945/4726.

Hall, D.S., Shattuck, L.G. and Bennett, K.B. (2012) 'Evaluation of an Ecological Interface Design for Military Command and Control', *Journal of Cognitive Engineering and Decision Making*, 6(2), pp. 165-193.

Hall, T.J. (2012) *A Case Study of Innovation and Change in the U.S Navy Submarine Fleet*. Master of Science in Information Technology Management. Naval Postgraduate School.

Ham, D.H., Yoon, W.C. and Han, B.T. (2008) 'Experimental study on the effects of visualized functionally abstracted information on process control tasks', *Reliability Engineering & System Safety*, 93(2), pp. 254-270.

Hamburger, P., Miskimens, D. and Truver, S. (2011) 'It is Not Just Hardware and Software, Anymore! Human Systems Integration in US Submarines', *Naval Engineers Journal*, 123(4), pp. 41-50.

Hancock, P.A. *et al.* (2008) *Human factors in simulation and training*. Boca Raton, USA: CRC Press.

Hanisch, K.A., Kramer, A.F. and Hulin, C.L. (1991) 'Cognitive representations, control, and understanding of complex systems: a field study focusing on components of users' mental models and expert/novice differences', *Ergonomics*, 34(8), pp. 1129-45.

Harre, M.-C. (2019) *Supporting supervisory control of safety-critical systems with rationally justified information visualizations*. Doctor of Engineering. Universität Oldenburg. Available at: https://oops.uni-oldenburg.de/4033/1/harsup19.pdf.

Harre, M.-C. and Lüdtke, A. (2018) 'Supporting supervisory control of safety-critical systems by optimizing visualizations on monitoring displays for fast and accurate perception', *Human Factors and Ergonomics SocietyAnnual Meeting*. Philadelphia, PA. Human Factors and Ergonomics Society. Available at: https://www.hfes-europe.org/wp-content/uploads/2018/10/Harre2018.pdf.

References

Hart, S.G. (2006) 'Nasa-Task Load Index (NASA-TLX); 20 Years Later', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9), pp. 904-908.

Hart, S.G. and Staveland, L.E. (1988) 'Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research', *Advances in psychology*, 52, pp. 139-183.

Hasani, L.M. *et al.* (2020) 'User-Centered Design of e-Learning User Interfaces: A Survey of the Practices', *2020 3rd International Conference on Computer and Informatics Engineering (IC2IE)*. pp. 1-7. doi: 10.1109/IC2IE50715.2020.9274623. Available at: https://ieeexplore.ieee.org/document/9274623/.

Hautamaki, B.S., Bagnall, T. and Small, R.L. (2005) *Human Interface Evaluation Methods for Submarine Combat Systems*. University Park, PA, USA. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.110.6661&rep=rep1&type=pdf.

Havelsan (2015) *Combat Management System for New Type Submarine*. Available at: https://web.archive.org/web/20171111004345/http://www.havelsan.com.tr/ENG/Main/urun/997/combat-management-system-for-new-type-submarine (Accessed: 13/11/2023).

Henley, P., Schmitt, M. and Huf, S. (2013) 'Should We Design Human-Machine Interfaces for Specifiable Human Tasks or to Afford Human Ingenuity - And How?', *Human Factors & Ergonomics Society of Australia Conference*. Perth, Australia. Human Factors & Ergonomics Society of Australia, p. 8. Available at: https://web.archive.org/web/20160303215700/http://www.ergonomics.org.au/downloads/HFESA_2013_Conference_Proceedings/Should_We_Design_Human-Machine_Interfaces_for_Specifiable_Human_Tasks_or.pdf (Accessed: 13/11/2023).

Hettinger, A.Z., Roth, E.M. and Bisantz, A.M. (2017) 'Cognitive engineering and health informatics: Applications and intersections', *Journal of Biomedical Informatics*, 67, pp. 21-33.

Hilliard, A. and Jamieson, G.A. (2007) 'Ecological interface design for solar car strategy: From state equations to visual relations', *2007 IEEE International Conference on Systems, Man and Cybernetics*. pp. 139-144. doi: 10.1109/ICSMC.2007.4413847.

Hilliard, A. and Jamieson, G.A. (2008) 'Winning solar races with interface design', *Ergonomics in Design*, 16(2), pp. 6-11.

Hingu, D.P. *et al.* (2017) 'A Software Tool for Cognitive Work Analysis', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, pp. 212-216.

Ho, G.A.J.W.H., Dal Vernon, C.R. and Jamieson, G.A. (2003) 'Ecological interface design in practice: A design for petrochemical processing operations', *Human-Computer Interaction: Theory and Practice (part 1), Volume 1*, 1, p. 133.

Hobson, B. (2008) 'Technology Insertion: A Way Ahead', *CANADIAN NAVAL REVIEW*, 4(3), pp. 12-16.

Submarine: Steel Boats - Iron Men (1989) Directed by Hoffman, D. Maine, USA: Varied Directions Pictures.

Hollnagel, E. (2012) *FRAM, the functional resonance analysis method: modelling complex socio-technical systems*. Ashgate Publishing, Ltd.

Hopkins, A. (2000) *Lessons from Longford : the Esso gas plant explosion*. North Ryde, N.S.W: CCH Australia.

Hopkins, A. (2001) 'Was Three Mile Island a ‘Normal Accident’?', *Journal of Contingencies and Crisis Management*, 9(2), pp. 65-72.

Hou, M. *et al.* (2015) *Analytical Techniques for IAS Design*. Intelligent Adaptive Systems: An Interaction-Centered Design Perspective.

Houghton, R.J. *et al.* (2015) 'Combining network analysis with Cognitive Work Analysis: insights into social organisational and cooperation analysis', *Ergonomics*, 58(3), pp. 434-49.

Howie, D.E. and Vicente, K.J. (1998) 'Making the most of ecological interface design: the role of self-explanation', *International Journal of Human-Computer Studies*, 49(5), pp. 651-674.

Huf, S., Arulampalam, S. and Manning, R. (2006) *Virtual Battle Experiments AS-4: Explorations of Networked Maritime Operations*. Edinburgh, Australia. Available at: https://apps.dtic.mil/sti/citations/ADA448283 (Accessed: 13/11/2023).

Huf, S. *et al.* (2009) *Investigating the network enabled conventional submarine II: A summary of Australian simulation research*

Hughes, D.T. *et al.* (2010) 'Passive acoustic monitoring during the SIRENA 10 cetacean survey', *OCEANS 2010*. IEEE, pp. 1-10.

Human Factors Integration Defence Technology Centre (2007) *The CWA Tool*[Computer program].

References

Hunter, A., Hazen, M. and Randall, T. (2014) 'VICTORIA Class Submarine Human-in-the-Loop Experimentation Plan', *19th International Command & Control Research & Technology Symposium (ICCRTS)*. Alexandria, VA. National Technical Reports Library, p. 17.

IBM (2020) *SPSS will not calculate kappa statistic*. Available at: https://www.ibm.com/support/pages/spss-will-not-calculate-kappa-statistic (Accessed: 13/11/2023).

IBM (2021a) *What are neural networks?* Available at: https://www.ibm.com/topics/neural-networks (Accessed: 13/11/2023).

IBM (2021b) *What is automation?* Available at: https://www.ibm.com/topics/automation (Accessed: 13/11/2023).

Ince, L. *et al.* (2009) 'An evolutionary computing approach for the target motion analysis (TMA) problem for underwater tracks', *Expert Systems with Applications*, 36(2), pp. 3866-3879.

International Organization for Standardization (2019) *ISO 9241-210:2019: Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems*. International Organization for Standardization. Available at: https://www.iso.org/standard/77520.html.

Jacobus, P.W., Yan, P.-F. and Barrett, J. (2012) 'Information Management: The Advanced Processor Build (Tactical)', *Johns Hopkins APL Technical Digest*, 23(4), pp. 366-372.

Jamieson, G.A. (2003) 'Bridging the gap between cognitive work analysis and ecological interface design', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Read Import 2019/03/18 Read, pp. 273-277. doi: 10.1177/154193120304700305.

Jamieson, G.A. (2007) 'Ecological interface design for petrochemical process control: An empirical assessment', *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 37(6), pp. 906-920.

Jamieson, G.A. (2002) 'Empirical Evaluation of an Industrial Application of Ecological Interface Design', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Inc, pp. 536-540. doi: 10.1177/154193120204600367. Available at: https://doi.org/10.1177/154193120204600367 (Accessed: 13/11/2023).

Jamieson, G.A. and Lau, N. (2010) 'Scaling Up Ecological Interface Design: Tackling Large Work Domains with International Project Teams', *IFAC Proceedings Volumes*, 43(13), pp. 216-220.

Jamieson, G.A. *et al.* (2007) 'Integrating task- and work domain-based work analyses in ecological interface design: A process control case study', *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 37(6), pp. 887-905.

Jamieson, G.A. and Vicente, K.J. (2001) 'Ecological interface design for petrochemical applications: Supporting operator adaptation, continuous learning, and distributed, collaborative work', *Computers & Chemical Engineering*, 25(7-8), pp. 1055-1074.

Jamson, A.H., Hibberd, D.L. and Merat, N. (2015) 'Interface design considerations for an in-vehicle eco-driving assistance system', *Transportation Research Part C: Emerging Technologies*, 58, pp. 642-656.

Jenkins, D.P. *et al.* (2010a) 'A systemic approach to accident analysis: A case study of the Stockwell shooting', *Ergonomics*, 53(1), pp. 1-17.

Jenkins, D.P. *et al.* (2009) *Cognitive Work Analysis: Coping With Complexity*. Farnham, UK: Ashgate. Human Factors in Defence.

Jenkins, D.P. *et al.* (2008a) 'Using cognitive work analysis to explore activity allocation within military domains', *Ergonomics*, 51(6), pp. 798-815.

Jenkins, D.P. *et al.* (2007) 'The Development of a Cognitive Work Analysis Tool', Berlin, Heidelberg. Springer Berlin Heidelberg, pp. 504-511.

Jenkins, D.P. *et al.* (2008b) 'Applying cognitive work analysis to the design of rapidly reconfigurable interfaces in complex networks', *Theoretical Issues in Ergonomics Science*, 9, pp. 273-295.

Jenkins, D.P. *et al.* (2010b) 'Using cognitive work analysis to explore system flexibility', *Theoretical Issues in Ergonomics Science*, 11, pp. 136-150.

Jiancaro, T., Jamieson, G.A. and Mihailidis, A. (2014) 'Twenty Years of Cognitive Work Analysis in Health Care A Scoping Review', *Journal of Cognitive Engineering and Decision Making*, 8(1), pp. 3-22.

Johnson-Laird, P.N. (1989) 'Mental models.', in Posner, M.I. (ed.) *Foundations of Cognitive Science*. MIT press Cambridge, MA.

Johnson-Laird, P.N. (1983) *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press. 6.

Johnston, K.L. and Featherstone, R.W. (2014) *A case study of introducing innovation through design*. Monterey, California: Naval Postgraduate School.

Jones, B.W. (2009) *USS Toledo SSN-769*. Available at: http://prometheus.med.utah.edu/~bwjones/2009/05/uss-toledo-ssn-769/ (Accessed: 13/11/2023).

Kant, V. and Sudakaran, J.S. (2022) 'Extending the Ecological Interface Design process—Integrated EID', *Human Factors and Ergonomics in Manufacturing & Service Industries*, 32(1), pp. 102-124.

Kilgore, R. and St-Cyr, O. (2006) 'The SRK Inventory: A Tool for Structuring and Capturing a Worker Competencies Analysis', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Inc, pp. 506-509. doi: 10.1177/154193120605000362. Available at: https://doi.org/10.1177/154193120605000362 (Accessed: 13/11/2023).

Kilgore, R. and Voshell, M. (2014) 'Increasing the Transparency of Unmanned Systems: Applications of Ecological Interface Design', *Virtual, Augmented and Mixed Reality. Applications of Virtual and Augmented Reality*. Cham. Springer International Publishing, pp. 378-389.

King, B.J., Read, G.J.M. and Salmon, P.M. (2022) 'Clear and present danger? Applying ecological interface design to develop an aviation risk management interface', *Applied Ergonomics*, 99, p. 103643.

Kirschenbaum, S.S. (2001) 'Analyzing submarine decision making: A question of levels', *Linking expertise and naturalistic decision making*, pp. 189-207.

Kirschenbaum, S.S. *et al.* (2014) 'Visualizing Uncertainty: The Impact on Performance', *Human Factors*, 56(3), pp. 509-520.

Kleiner, B.M. *et al.* (2015) 'Sociotechnical attributes of safe and unsafe work systems', *Ergonomics*, 58(4), pp. 635-649.

Klomp, R. *et al.* (2016) 'Expertise Level, Control Strategies, and Robustness in Future Air Traffic Control Decision Aiding', *IEEE Transactions on Human-Machine Systems*, 46(2), pp. 255-266.

Kortschot, S.W. *et al.* (2017) 'Conducting Cognitive Work Analysis with an Experiment in Mind', in *Cognitive Work Analysis: Applications, Extensions and Future Directions*. CRC Press, pp. 233-248.

Kwok, J. (2007) *Ecological Interface Design for Turbine Secondary Systems in a Nuclear Power Plant: Effects on Operator Situation Awareness*. UWSpace. Available at: http://hdl.handle.net/10012/2987.

Landis, J.R. and Koch, G.G. (1977) 'The measurement of observer agreement for categorical data', *Biometrics*, 33(1), pp. 159-174.

Landsberg, C.R. *et al.* (2012) 'Evaluation of an adaptive training system for submarine periscope operations', *Proceedings of the Human Factors and Ergonomics Society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA, pp. 2422-2426.

Lau, N. and Jamieson, G. (2006) 'Ecological interface design for the condenser subsystems of a boiling water reactor', *Proceedings of the 16th World Congress on Ergonomics*.

Lau, N. *et al.* (2008) 'Ecological Interface Design in the Nuclear Domain: An Application to the Secondary Subsystems of a Boiling Water Reactor Plant Simulator', *IEEE Transactions on Nuclear Science*, 55(6), pp. 3579-3596.

Lee, M.-H. *et al.* (2008) 'Pre-processing faded measurements for bearing-and-frequency target motion analysis', *International Journal of Control, Automation, and Systems*, 6(3), pp. 424-433.

Lee, S.M., Ha, J.S. and Seong, P.H. (2011) 'CREAM-based communication error analysis method (CEAM) for nuclear power plant operators' communication', *Journal of Loss Prevention in the Process Industries*, 24(1), pp. 90-97.

Leveson, N.G. (2000) 'Intent specifications: An approach to building human-centered specifications', *IEEE Transactions on Software Engineering*, 26(1), pp. 15-35.

Lewandowski, L.J. and Kobus, D.A. (1989) 'Bimodal Information Processing in Sonar Performance', *Human Performance*, 2(1), pp. 73-84.

Lintern, G. (2006) 'A functional workspace for military analysis of insurgent operations', *International Journal of Industrial Ergonomics*, 36(5), pp. 409-422.

Joker One: A Tutorial in Cognitive Work Analysis (2013).

Loft, S. *et al.* (2018) 'Using Situation Awareness and Workload to Predict Performance in Submarine Track Management: A Multilevel Approach', *Human Factors*, 60(7), pp. 978-991.

Lützhöft, M.H. and Dekker, S.W.A. (2002) 'On Your Watch: Automation on the Bridge', *Journal of Navigation*, 55(1), pp. 83-96.

Ly, T., Huf, S. and Henley, P. (2007) 'Design for Submarine Command and Control in the 21st Century', *12th International Command and Control Research and Technology Symposum,* . Newport, Rhode Island, United States of America. Newport, Rhode Island, United States of America: DoD CCRP. Available at: http://dodccrp.org/events/12th_ICCRTS/CD/html/papers/228.pdf.

Mack, S. (2003) 'Perisher Submarine ', *Undersea Warfare*, 5(3), p. 36.

Man, Y.M. *et al.* (2018) 'Human factor issues during remote ship monitoring tasks: An ecological lesson for system design in a distributed context', *International Journal of Industrial Ergonomics*, 68, pp. 231-244.

Mansell, T., Tynan, A. and Kershaw, D. (2004) 'Investigating The Network Enabled Conventional Submarine: An Australian Perspective'.

Mantel, B., Hoppenot, P. and Colle, E. (2012) 'Perceiving for Acting With Teleoperated Robots: Ecological Principles to Human-Robot Interaction Design', *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 42(6), pp. 1460-1475.

Maranda, B.H. (2008) 'Passive Sonar', in Havelock, D., Kuwano, S. and Vorländer, M. (eds.) *Handbook of Signal Processing in Acoustics*. New York, NY: Springer New York, pp. 1757-1781.

Marine Accident Investigation Branch (2016) *Report on the investigation of the collision between the stern trawler Karen (B317) and a dived Royal Navy submarine in the Irish Sea on 15 April 2015*. Southampton, United Kingdom.

Masakowski, Y.R. (2000) 'Optimized Manning for the 21st Century', *The 36th International Applied Military Psychology Symposium*. Splits, Croatia. Available at: http://www.iamps.org/_Presentations/2000_526P.pdf (Accessed: 13/11/2023).

Masakowski, Y.R. and Hardinge, N. (2000) 'Comparing the Cognitive Decision-Making Strategies of US and UK Naval Sonar Operators', *The 36th International Applied Military Psychology Symposium*. Split, Croatia. p. 167.

Mason, P. *et al.* (1989) 'Knowledge-Based Interpretation of Passive SONAR Data', in Chan, Y.T. (ed.) *Underwater Acoustic Data Processing*. Dordrecht: Springer Netherlands, pp. 603-607.

Matthews, M.L. *et al.* (2006) *A Predictive Task Network Model for Estimating the Effectiveness of Decision Aids for Sonar Operators*. Toronto, Canada: Center, D.T.I. Available at: https://apps.dtic.mil/sti/citations/ADA634389 (Accessed: 13/11/2023).

Mazaeva, N. and Bisantz, A.M. (2014) 'Ecological Displays, Information Integration, and Display Format', *Journal of Cognitive Engineering and Decision Making*, 8(2), pp. 137-161.

McCabe, F., Baber, C. and Stone, B. (2020) 'Classifying Vessels Using Broadband Sonar: Considerations for Future Autonomous Support', *Contemporary Ergonomics & Human Factors 2020 Conference Proceedings*. pp. 78-85.

McIlroy, R. (2016) *Encouraging eco-driving: the case for vibrotactile information presented through the accelerator pedal*. University of Southampton. Available at: https://eprints.soton.ac.uk/398616/.

McIlroy, R.C. and Stanton, N.A. (2015a) 'A decision ladder analysis of eco-driving: the first step towards fuel-efficient driving behaviour', *Ergonomics*, 58(6), pp. 866-82.

McIlroy, R.C. and Stanton, N.A. (2015b) 'Ecological Interface Design Two Decades On: Whatever Happened to the SRK Taxonomy?', *IEEE Transactions on Human-Machine Systems*, 45(2), pp. 145-163.

McIlroy, R.C. and Stanton, N.A. (2011) 'Getting past first base: Going all the way with Cognitive Work Analysis', *Applied Ergonomics*, 42(2), pp. 358-70.

McIlroy, R.C. and Stanton, N.A. (2012) 'Specifying the requirements for requirements specification: the case for Work Domain and Worker Competencies Analyses', *Theoretical Issues in Ergonomics Science*, 13(4), pp. 450-471.

Mendes, M. and Pala, A. (2003) 'Type I error rate and power of three normality tests', *Pakistan Journal of Information and Technology*, 2(2), pp. 135-139.

Mendoza, P.A., Angelelli, A. and Lindgren, A. (2011) 'Ecological interface design inspired human machine interface for advanced driver assistance systems', *IET Intelligent Transport Systems*, 5(1), pp. 53-59.

Meshkati, N. (1991) 'Integration of workstation, job, and team structure design in complex human-machine systems: A framework', *International Journal of Industrial Ergonomics*, 7(2), pp. 111-122.

Mewett, L.C.R. (2014) 'Pass or Perish: A U.S. Submariner Swaps SSNs for SSKs', *Undersea Warfare*(54).

Michailovs, S. *et al.* (2022) 'The impact of digital image configuration on submarine periscope operator workload, situation awareness, meta-awareness and performance', *Cognitive Systems Research*, 76, pp. 13-25.

Michailovs, S. *et al.* (2021) 'The Impact of Information Integration in a Simulation of Future Submarine Command and Control', *Human Factors*, 0(0), p. 00187208211045872.

Mill, R. and Brown, G. (2005) *Auditory-Based Time-Frequency Rpresentations and Feature Extraction Techniques for Sonar Processing*. Sheffield, UK: Speech and Hearing Research Group, University of Sheffield.

References

Miller, A., Scheinkestel, C. and Steele, C. (2009) 'The effects of clinical information presentation on physicians' and nurses' decision-making in ICUs', *Applied Ergonomics*, 40, pp. 753-761.

Miller, C.A. and Vicente, K.J. (1998) *Abstraction decomposition space analysis for NOVA's E1 acetylene hydrogenation reactor*.

Miller, M.J. and Feigh, K.M. (2019) 'Addressing the envisioned world problem: a case study in human spaceflight operations', *Design Science*, 5.

Ministry of Defence (2022) 'Australian submariners to join Royal Navy crews as UK and Australia deepen defence ties through AUKUS agreement '. Available at: https://www.gov.uk/government/news/australian-submariners-to-join-royal-navy-crews-as-uk-and-australia-deepen-defence-ties-through-aukus-pact.

Mohd Razali, N. and Yap, B. (2011) 'Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests', *J. Stat. Model. Analytics*, 2.

Momtahan, K. and Burns, C. (2004) 'Applications of ecological interface design in supporting the nursing process', *Journal of Healthcare Information Management*, 18(4), pp. 74-82.

Morineau, T. *et al.* (2009) 'Experimental evaluation of a tide prediction display based on the ecological interface design framework', *Cognition, Technology & Work*, 11(2), pp. 119-127.

Mosier, K.L. and Skitka, L.J. (1996) 'Human Decision Makers and Automated Decision Aids', in Parasuraman, R. and Mouloua, S. (eds.) *Automation and human performance: Theory and applications*. Mahwah, NJ: Lawrence Erlbaum, pp. 201–220.

Murphy, M. (2000a) 'Dead-Reckoning Systems', in *Operations Specialist, Volume 1*. Available at: http://www.globalsecurity.org/military/library/policy/navy/nrtc/14308_ch9.pdf (Accessed: 13/11/2023).

Murphy, M. (2000b) 'Internal Communications', in *Operations Specialist, Volume 1*. Available at: http://www.globalsecurity.org/military/library/policy/navy/nrtc/14308_ch3.pdf (Accessed: 13/11/2023).

Murphy, M. (2000c) 'Plotting', in *Operations Specialist, Volume 1*. Available at: http://www.globalsecurity.org/military/library/policy/navy/nrtc/14308_ch10.pdf (Accessed: 13/11/2023).

Nachreiner, F. (1995) 'Standards for ergonomics principles relating to the design of work systems and to mental workload', *Applied Ergonomics*, 26(4), pp. 259-263.

Naikar, N. (2006) 'Beyond interface design: Further applications of cognitive work analysis', *International Journal of Industrial Ergonomics*, 36(5), pp. 423-438.

Naikar, N. (2017) 'Cognitive work analysis: An influential legacy extending beyond human factors and engineering', *Applied Ergonomics*, 59(Pt B), pp. 528-540.

Naikar, N. (2013) *Work domain analysis: Concepts, guidelines, and cases*. Boca Raton, Florida, United States: CRC Press.

Naikar, N. and Elix, B. (2016) 'Reflections on Cognitive Work Analysis and Its Capacity to Support Designing for Adaptation', *Journal of Cognitive Engineering and Decision Making*, 10(2), pp. 123-125.

Naikar, N., Hopcroft, R. and Moylan, A. (2005) *Work domain analysis: Theoretical concepts and methodology*.

Naikar, N., Moylan, A. and Pearce, B. (2006) 'Analysing activity in complex systems with cognitive work analysis: concepts, guidelines and case study for control task analysis', *Theoretical Issues in Ergonomics Science*, 7(4), pp. 371-394.

Naikar, N. *et al.* (2003) 'Designing teams for first-of-a-kind, complex systems using the initial phases of cognitive work analysis: case study', *Human Factors*, 45(2), pp. 202-17.

Naikar, N. and Sanderson, P.M. (1999) 'Work domain analysis for training-system definition and acquisition', *International Journal of Aviation Psychology*, 9, pp. 271-290.

Naito, N. *et al.* (1995) 'An Intelligent Human-Machine System Based on an Ecological Interface Design Concept', *Nuclear Engineering and Design*, 154(2), pp. 97-108.

Nakashima, A., Chow, R. and Wang, W. (2015) 'Research Requirements for Modelling of Auditory Communication in Critical Control Spaces on RCN Platforms', *DRDC Scientific Report DRDC-RDDC-2015-R171*.

National Research Council (1997) *Technology for the United States Navy and Marine Corps, 2000-2035 Becoming a 21st-Century Force*. Washington, DC: The National Academies Press.

National Transportation Safety Board (2001) *Marine Accident Brief: Collision between the U.S. Navy Submarine USS Greeneville and Japanese Motor Vessel Ehime Maru near Oahu, Hawaii* (NTSB/MAB-05/01) (DCA-01-MM-022). Available at: https://www.ntsb.gov/investigations/AccidentReports/Reports/MAB0501.pdf (Accessed: 13/11/2023).

Newnam, S. *et al.* (2017) 'Reforming the road freight transportation system using systems thinking: An investigation of Coronial inquests in Australia', *Accident Analysis & Prevention*, 101, pp. 28-36.

Nielsen, C.W., Goodrich, M.A. and Ricks, R.W. (2007) 'Ecological interfaces for improving mobile robot teleoperation', *IEEE Transactions on Robotics*, 23(5), pp. 927-941.

Niskanen, T. (2018) 'Implementation of a novel taxonomy based on cognitive work analysis in the assessment of safety performance', *International Journal of Occupational Safety and Ergonomics*, 24(4), pp. 624-637.

Norman, D. and Draper, S. (1986) 'User-Centered System Design: New Perspectives on Human-Computer Interaction. Lawrence Earlbaum Associates, Hillsdale, NJ'.

Norman, D.A. (2013) *The Design of Everyday Things*. Revised & Expanded Edition edn. New York: Basic Books.

Norman, D.A. (1988) *The Psychology of Everyday Things*. Basic Books.

Norros, L. (1996) 'System disturbances as springboard for development of operators' expertise', in Middleton, D. and Engeström, Y. (eds.) *Cognition and Communication at Work*. Cambridge: Cambridge University Press, pp. 159-176.

Offshore Systems Ltd (2007) *ECPINS W: The Surface WECDIS Solution*. New Milton, United Kingdom: Ltd, O.S. Available at: https://ecdisregs.com/wp-content/uploads/Offshore-Systems-Ltd-OSI-ECPINS.pdf.

Olsson, G. and Lee, P.L. (1994) 'Effective interfaces for process operators – A prototype', *Journal of Process Control*, 4(2).

Oosthuizen, R. and Pretorius, P. (2018) 'Multimethod approach to modelling of complex sociotechnical systems'. Available at: http://hdl.handle.net/10204/10555.

Owen, L. *et al.* (2006) 'Architectural Description of the UK Common Core Combat System', *International Command and Control Technology Symposium (ICCRTS)*. Cambridge, UK. Department of Defence CCRP. Available at: http://www.dodccrp.org/events/11th_ICCRTS/html/papers/039.pdf.

Parasuraman, R., Sheridan, T.B. and Wickens, C.D. (2008) 'Situation Awareness, Mental Workload, and Trust in Automation: Viable, Empirically Supported Cognitive Engineering Constructs', *Journal of Cognitive Engineering and Decision Making*, 2(2), pp. 140-160.

Patriarca, R., Di Gravio, G. and Costantino, F. (2017) 'A Monte Carlo evolution of the Functional Resonance Analysis Method (FRAM) to assess performance variability in complex systems', *Safety Science*, 91, pp. 49-60.

Patriarca, R. *et al.* (2020) 'Framing the FRAM: A literature review on the functional resonance analysis method', *Safety Science*, 129, p. 104827.

Pawlak, W.S. and Vicente, K.J. (1996) 'Inducing effective operator control through ecological interface design', *International Journal of Human-Computer Studies*, 44(5), pp. 653-688.

Pejtersen, A.M. (1992) 'The book house. An icon based database system for fiction retrieval in public libraries', in *The marketing of library and information services*. 2nd. ed. edn. London: Aslib, The Association for Information Management, pp. 572-591.

Pfautz, S.L. and Pfautz, J. (2008) 'Understanding social and organizational aspects of the work domain using techniques and technologies from intelligence analysis', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, pp. 443-447.

Philippe, R. *et al.* (2016) 'Analysis of submarine steering: effects of cognitive and perceptual–motor requirements on the mental workload and performance of helmsmen', *Cognition, Technology & Work*, 18(4), pp. 657-672.

Plumb, C.M. and Kendrick, A.M. (1981) 'Surface Ship Noise Reduction', *Journal of Naval Engineering*, 26(3), p. 7.

Pope, K., Roberts, A. and Stanton, N. (2019) 'Investigating Temporal Implications of Information Transition in Submarine Command Teams', Cham. Springer International Publishing, pp. 243-253.

Pope, K.A. *et al.* (2020) 'Assessing Situation Awareness Across Different Submarine Control Room Layouts', *Applied Human Factors and Ergonomics 2020*. Online. Springer International Publishing, pp. 475-482. doi: 10.1007/978-3-030-50943-9_60.

Posner, M.I. (1980) 'Orienting of Attention', *Quarterly Journal of Experimental Psychology*, 32(1), pp. 3-25.

Punchihewa, Y. *et al.* (2022) 'Target Motion Analysis via Hard and Soft Data Fusion', *2022 11th International Conference on Control, Automation and Information Sciences (ICCAIS)*. pp. 101-106. doi: 10.1109/ICCAIS56082.2022.9990136. Available at: https://ieeexplore.ieee.org/document/9990136/.

References

Radigan, D. (2019) *Kanban - A brief introduction | Atlassian*. Available at: https://www.atlassian.com/agile/kanban (Accessed: 13/11/2023).

Rasmussen, J. (1999) 'Ecological interface design for reliable human-machine systems', *International Journal of Aviation Psychology*, 9(3), pp. 203-223.

Rasmussen, J. (1974a) 'Human data processor as a system component bits and pieces of a model'.

Rasmussen, J. (1974b) *The human data processor as a system component: Bits and pieces of a model*.

Rasmussen, J. (1997) 'Risk management in a dynamic society: a modelling problem', *Safety Science*, 27(2), pp. 183-213.

Rasmussen, J. (1985) 'The Role of Hierarchical Knowledge Representation in Decision-Making and System Management', *IEEE Transactions on Systems Man and Cybernetics*, 15(2), pp. 234-243.

Rasmussen, J. (1983) 'Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models', *IEEE Transactions on Systems, Man, and Cybernetics, SMC*, 13(3), pp. 257–266.

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

Rasmussen, J. and Vicente, K.J. (1989) 'Coping with Human Errors through System-Design - Implications for Ecological Interface Design', *International Journal of Man-Machine Studies*, 31(5), pp. 517-534.

Rasmussen, J. and Vicente, K.J. (1990) 'Ecological interfaces: A technological imperative in high-tech systems?', *International Journal of Human-Computer Interaction*, 2(2), pp. 93-110.

Rauffet, P. *et al.* (2015) 'Designing sociotechnical systems: a CWA-based method for dynamic function allocation', *Proceedings of the European Conference on Cognitive Ergonomics 2015*. pp. 1-8.

Read, G.J. *et al.* (2015a) 'Designing sociotechnical systems with cognitive work analysis: putting theory back into practice', *Ergonomics*, 58(5), pp. 822-51.

Read, G.J.M. *et al.* (2018) 'A sociotechnical design toolkit for bridging the gap between systems-based analyses and system design', *Human Factors and Ergonomics in Manufacturing & Service Industries*, 28(6), pp. 327-341.

Read, G.J.M., Salmon, P.M. and Lenne, M.G. (2015) 'Cognitive work analysis and design: current practice and future practitioner requirements', *Theoretical Issues in Ergonomics Science*, 16(2), pp. 154-173.

Read, G.J.M., Salmon, P.M. and Lenné, M.G. (2012) 'From work analysis to work design: A review of cognitive work analysis design applications', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. pp. 368-372. doi: 10.1177/1071181312561084. Available at: http://pro.sagepub.com/content/56/1/368.abstract.

Read, G.J.M. *et al.* (2015b) 'Designing a ticket to ride with the Cognitive Work Analysis Design Toolkit', *Ergonomics*, 58(8), pp. 1266-1286.

Read, G.J.M. *et al.* (2016) 'Applying the prompt questions from the Cognitive Work Analysis Design Toolkit: a demonstration in rail level crossing design', *Theoretical Issues in Ergonomics Science*, 17(4), pp. 354-375.

Read, G.J.M. *et al.* (2022) 'Using cognitive work analysis to identify competencies for human factors and ergonomics practitioners', *Ergonomics*, 65(3), pp. 348-361.

Rechard, J. *et al.* (2015) 'Experimental verification of ecological interface prototype issued by an automated generation tool', in Ahram, T., Karwowski, W. and Schmorrow, D. (eds.) *6th International Conference on Applied Human Factors and Ergonomics*. pp. 5747-5754.

Rehak, L.A., Lamoureux, T.M. and Bos, J.C. (2006) 'Communication, Coordination, and Integration of Cognitive Work Analysis Outputs', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(3), pp. 515-519.

Rehkopf, M. (2019a) *Agile epics: definition, examples, and templates*. Available at: https://www.atlassian.com/agile/project-management/epics (Accessed: 13/11/2023).

Rehkopf, M. (2019b) *Kanban cards*. Available at: https://www.atlassian.com/agile/kanban/cards (Accessed: 13/11/2023).

Rehkopf, M. (2018) *User stories with examples and a template* Available at: https://www.atlassian.com/agile/project-management/user-stories (Accessed: 13/11/2023).

Reising, D.V. and Sanderson, P.M. (2004) 'Minimal instrumentation may compromise failure diagnosis with an ecological interface', *Human Factors*, 46(2), pp. 316-33.

Reising, D.V.C. and Sanderson, P.M. (2000) 'Testing the Impact of Instrumentation Location and Reliability on Ecological Interface Design: Control Performance', *Proceedings of the Human Factors* *and Ergonomics Society Annual Meeting*. SAGE Publications Inc, pp. 124-127. doi: 10.1177/154193120004400133. Available at: https://doi.org/10.1177/154193120004400133 (Accessed: 13/11/2023).

Revell, K. *et al.* (2018) 'User centered ecological interface design (UCEID): A novel method applied to the problem of safe and user-friendly interaction between drivers and autonomous vehicles', *1st International Conference on Intelligent Human Systems Integration: Integrating People and Intelligent Systems, IHSI 2018*. Springer Verlag, pp. 495-501. doi: 10.1007/978-3-319-73888-8_77. Available at: https://link.springer.com/content/pdf/10.1007%2F978-3-319-73888-8_77.pdf (Accessed: 13/11/2023).

Revell, K. *et al.* (2019) 'UCEID - The Best of Both Worlds: Combining Ecological Interface Design with User Centered Design in a Novel HF Method Applied to Automated Driving', *AHFE International Conference on Human Factors in Transportation, 2018*. Springer Verlag, pp. 493-501. doi: 10.1007/978-3-319-93885-1_44. Available at: https://link.springer.com/content/pdf/10.1007%2F978-3-319-93885-1_44.pdf (Accessed: 13/11/2023).

Revell, K.M.A. (2015) *Mental models: understanding domestic energy systems and user behaviour*. University of Southampton. Available at: https://eprints.soton.ac.uk/386139/.

Rhie, Y.L. *et al.* (2017) 'Design specifications for Multi-Function Consoles for use in submarines using anthropometric data of South Koreans', *International Journal of Industrial Ergonomics*, 59, pp. 8-19.

Roberts, A., Stanton, N. and Fay, D. (2015) 'The Command Team Experimental Test-bed Stage 1: Design and build of a submarine command room simulator', *6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences*. pp. 2800-2807. doi: 10.1016/j.promfg.2015.07.745. Available at: <Go to ISI>://WOS:000383740302122.

Roberts, A., Stanton, N.A. and Fay, D. (2017a) 'The Command Team Experimental Test-Bed Phase Two: Assessing Cognitive Load and Situation Awareness in a Submarine Control Room', *Advances in Human Aspects of Transportation: Proceedings of the AHFE 2016 International Conference on Human Factors in Transportation*. Florida, USA, July 27-31, 2016. Springer International Publishing, pp. 427-437. doi: 10.1007/978-3-319-41682-3_36. Available at: https://doi.org/10.1007/978-3-319-41682-3_36.

Roberts, A.P.J. and Cole, J.C. (2018) 'Naturalistic Decision Making: Taking a (Cognitive) Step Back to Take Two Steps Forward in Understanding Experience-Based Decisions', *Journal of Applied Research in Memory and Cognition*, 7(1), pp. 70-81.

Roberts, A.P.J. and Stanton, N.A. (2018) 'Macrocognition in Submarine Command and Control: A Comparison of three Simulated Operational Scenarios', *Journal of Applied Research in Memory and Cognition*, 7(1), pp. 92-105.

Roberts, A.P.J., Stanton, N.A. and Fay, D. (2017b) 'Land Ahoy! Understanding Submarine Command and Control During the Completion of Inshore Operations', *Hum Factors*, 59(8), pp. 1263-1288.

Roberts, A.P.J., Stanton, N.A. and Fay, D.T. (2018) 'Go Deeper, Go Deeper: Understanding submarine command and control during the completion of dived tracking operations', *Appl Ergon*, 69, pp. 162-175.

Roberts, A.P.J. *et al.* (2019) 'The effects of team co-location and reduced crewing on team communication characteristics', *Applied Ergonomics*, 81, p. 102875.

Roberts, A.P.J. *et al.* (2020) 'You say it is physical, I say it is functional; let us call the whole thing off! Simulation: an application divided by lack of common language', *Theoretical Issues in Ergonomics Science*, 21(5), pp. 507-536.

Roberts, K.H. and Tadmor, C.T. (2002) 'Lessons learned from non-medical industries: the tragedy of the USS Greeneville', *Qual Saf Health Care*, 11(4), pp. 355-7.

Roscoe, A.H. and Ellis, G.A. (1990) *A subjective rating scale for assessing pilot workload in flight: A decade of practical use*.

Royal Australian Navy (2023) 'The AUKUS Nuclear-Powered Submarine Pathway: a partnership for the future'. Available at: https://www.navy.gov.au/aukus-nuclear-powered-submarine-pathway-partnership-future.

Royal Navy (2023a) 'AUKUS deal delivers new class of submarines for UK and Australia' 2023. Available at: https://www.royalnavy.mod.uk/news-and-latest-activity/news/2023/march/14/20230314-aukus-deal-delivers-new-class-of-submarines-for-uk-and-australia.

Royal Navy (2019) *Home | Royal Navy*. Available at: https://www.royalnavy.mod.uk/ (Accessed: 13/11/2023).

References

Royal Navy (2022) '‘No closer ally’ – head of US Navy praises Royal Navy after talks with First Sea Lord'. Available at: https://www.royalnavy.mod.uk/news-and-latest-activity/news/2022/october/21/20221021-no-closer-ally-head-of-us-navy-praises-royal-navy-after-talks-with-first-sea-lord.

Royal Navy (2020) 'Royal Navy and US Navy to work together to embrace ideas and new tech'. Available at: https://www.royalnavy.mod.uk/news-and-latest-activity/news/2020/december/10/201210-london-tech-bridge-launch.

Royal Navy (2015) *Submariners benefit from realistic sonar training*. Available at: https://web.archive.org/web/20151031043251/http://www.royalnavy.mod.uk/news-and-latest-activity/news/2015/october/28/151028-submariners-realistic-sonar-training (Accessed: 13/11/2023).

Royal Navy (2008) *Trafalgar Class Submarine Updates*. Available at: http://webarchive.nationalarchives.gov.uk/20080530151058/http://www.royal-navy.mod.uk/server.php?show=nav.2553 (Accessed: 13/11/2023).

Royal Navy (2023b) 'UK and US to work more closely on littoral operations'. Available at: https://www.royalnavy.mod.uk/news-and-latest-activity/news/2023/april/05/230405-uk-us-intent-on-littoral-operations.

Saffer, D. (2010) *Designing for interaction: creating innovative applications and devices*. New Riders.

Salehi, V., Veitch, B. and Smith, D. (2021) 'Modeling complex socio-technical systems using the FRAM: A literature review', *Human Factors and Ergonomics in Manufacturing & Service Industries*, 31(1), pp. 118-142.

Salmon, P. *et al.* (2010) 'Hierarchical task analysis vs. cognitive work analysis: comparison of theory, methodology and contribution to system design', *Theoretical Issues in Ergonomics Science*, 11(6), pp. 504-531.

Salmon, P.M., Carden, T. and Stevens, N.J. (2018) 'Breaking bad systems: using work domain analysis to identify strategies for disrupting terrorist cells', in Charles, R. and Wilkinson, J. (eds.) *Contemporary Ergonomics & Human Factors 2018*. Wootton Wawen, UK: Chartered Institute of Ergonomics and Human Factors.

Salmon, P.M. *et al.* (2016) 'More than meets the eye: Using cognitive work analysis to identify design requirements for future rail level crossing systems', *Applied Ergonomics*, 53, pp. 312-322.

Salmon, P.M. *et al.* (2017) 'Cognitive Work Analysis for Systems Analysis and Redesign: rail level crossings case study', in *Cognitive Work Analysis: Applications, Extensions and Future Directions*. CRC Press, pp. 271-284.

Salmon, P.M. and Read, G.J.M. (2019) 'Many model thinking in systems ergonomics: a case study in road safety', *Ergonomics*, 62(5), pp. 612-628.

Salmon, P.M. *et al.* (2007) 'Work domain analysis and intelligent transport systems: implications for vehicle design', *International Journal of Vehicle Design*, 45(3), pp. 426-448.

Sanders, E.B.N. and Stappers, P.J. (2008) 'Co-creation and the new landscapes of design', *CoDesign*, 4(1), pp. 5-18.

Sanderson, P. *et al.* (2004) 'Visual and auditory attention in patient monitoring: A formative analysis', *Cognition, Technology, and Work*, 6.

Sauro, J. and Lewis, J.R. (2012) 'Chapter 8 - Standardized Usability Questionnaires', in Sauro, J. and Lewis, J.R. (eds.) *Quantifying the User Experience*. Boston: Morgan Kaufmann, pp. 185-240.

Schank, J.F. *et al.* (2011) *Learning From Experiance: Lessons from the U.S Navy's Ohio, Seawolf, and Virginia Submarine Programs* (ISBN: 978-0-8330-5896-6, Library of Congress Control Number: 2011939404). Santa Monica, CA, USA: Corporation, R. Available at: http://www.rand.org/content/dam/rand/pubs/monographs/2011/RAND_MG1128.2.pdf.

Schewe, F. and Vollrath, M. (2020) 'Ecological interface design effectively reduces cognitive workload – The example of HMIs for speed control', *Transportation Research Part F: Traffic Psychology and Behaviour*, 72, pp. 155-170.

Schneider, W. and Shiffrin, R.M. (1977) 'Controlled and automatic human information processing: I. Detection, search, and attention', *Psychological Review*, 84, pp. 1-66.

Schroeder, B.L. *et al.* (2020) 'Using a Non-player Character to Improve Training Outcomes for Submarine Electronic Warfare Operators', *Adaptive Instructional Systems*. Cham. Springer International Publishing, pp. 531-542.

Scott, J. *et al.* (2011) 'Software is a Renewable Military Resource', *Software Tech*, 14(1), pp. 4-9.

Scott, R. (2006) 'Open for Business: A New Model for Submarine Sonar', *Jane’s Navy International*, 1.

Sea (2005) 'News', *Sea News*, (Autumn 2005), p. 8. Available at: https://www.sea.co.uk/docs/newsletter/SEA%20NEWS%20AUTUMN%2005.pdf.

Segall, N. *et al.* (2013) 'A Cognitive Modeling Approach to Decision Support Tool Design for Anesthesia Provider Crisis Management', *International Journal of Human-Computer Interaction*, 29(2), pp. 55-66.

Selkowitz, A.R. *et al.* (2017) 'Displaying Information to Support Transparency for Autonomous Platforms', in SavageKnepshield, P. and Chen, J. (eds.) *Advances in Human Factors in Robots and Unmanned Systems*. pp. 161-173.

Seppelt, B.D. and Lee, J.D. (2007) 'Making adaptive cruise control (ACC) limits visible', *International Journal of Human-Computer Studies*, 65(3), pp. 192-205.

Shar, P. and Li, X.R. (2000) 'Passive sonar fusion for submarine C/sup 2/systems', *IEEE Aerospace and Electronic Systems Magazine*, 15(3), pp. 29-34.

Shier, A.P. *et al.* (2018) 'Design and evaluation of a safety-centered user interface for radiation therapy', *Practical Radiation Oncology*, 8(5), pp. e346-e354.

Silva, G., Ribeiro, R. and Lopes, R.J. (2021) '"How Was the Match?": Semantic similarity between electronic media commentary and work domain analysis key phrases', *"How Was the Match?": Semantic similarity between electronic media commentary and work domain analysis key phrases*, pp. 144-150.

Simon, H.A. (2019) *The Sciences of the Artificial, reissue of the third edition with a new introduction by John Laird*. MIT press.

Simon, L. *et al.* (2022) 'Integrating Transparency to Ecological Interface Design', *Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022)*. pp. 3301-3308.

Smith, M.H. *et al.* (2013) *Leading Edge: Combat Systems Engineering & Integration* (February 2013).

Solomon, I.S.D. and Knight, A.J. (2002) 'Spatial processing of signals received by platform mounted sonar', *Ieee Journal of Oceanic Engineering*, 27(1), pp. 57-65.

Sonalysts (2022) *Simulation Engine II*. Available at: https://www.sonalysts.com/what-we-do/simulation-engine-ii/ (Accessed: 13/11/2023).

St-Maurice, J.D. and Burns, C.M. (2018) 'Using cognitive work analysis to compare complex system domains', *Theoretical Issues in Ergonomics Science*, 19(5), pp. 553-577.

Stanhope, M. (2012) 'Future Navy Vision: The Royal Navy Today, Tomorrow and Towards 2025'. Available at: http://www.royalnavy.mod.uk/About-the-Royal-Navy/~/media/Files/Navy-PDFs/About-the-Royal-Navy/Future%20Navy%20Vision.pdf.

Stanton, N., Baber, C. and Harris, D. (2008) *Modelling command and control: Event analysis of systemic teamwork*. Ashgate Publishing, Ltd.

Stanton, N.A. (2015) 'Distributed situation awareness', *Theoretical Issues in Ergonomics Science*, 17(1), pp. 1-7.

Stanton, N.A. (2006) 'Hierarchical task analysis: Developments, applications, and extensions', *Applied Ergonomics*, 37(1), pp. 55-79.

Stanton, N.A. (2014) 'Representing distributed cognition in complex systems: how a submarine returns to periscope depth', *Ergonomics*, 57(3), pp. 403-418.

Stanton, N.A. *et al.* (2003) 'Virtuality in human supervisory control: assessing the effects of psychological and social remoteness', *Ergonomics*, 46(12), pp. 1215-1232.

Stanton, N.A. and Bessell, K. (2014) 'How a submarine returns to periscope depth: analysing complex socio-technical systems using Cognitive Work Analysis', *Appl Ergon*, 45(1), pp. 110-25.

Stanton, N.A. and Harvey, C. (2017) 'Beyond human error taxonomies in assessment of risk in sociotechnical systems: a new paradigm with the EAST 'broken-links' approach', *Ergonomics*, 60(2), pp. 221-233.

Stanton, N.A. and McIlroy, R.C. (2012) 'Designing mission communication planning: The role of Rich Pictures and Cognitive Work Analysis', *Theoretical Issues in Ergonomics Science*, 13, pp. 146-168.

Stanton, N.A. *et al.* (2020a) *Task Analysis of RESM Operator's Workstation and Activities with Recommendations for Reducing Potential Human Failures* (1000133785-2b.2.1).

Stanton, N.A. and Roberts, A.P.J. (2020) 'Better together? Investigating new control room configurations and reduced crew size in submarine command and control', *Ergonomics*, 63(3), pp. 307-323.

Stanton, N.A. and Roberts, A.P.J. (2019) 'Block off: an examination of new control room configurations and reduced crew sizes examining engineered production blocking', *Cognition, Technology & Work*.

Stanton, N.A. and Roberts, A.P.J. (2018) 'Examining Social, Information, and Task Networks in Submarine Command and Control', *IEEE Transactions on Human-Machine Systems*, 48(3), pp. 252-265.

References

Stanton, N.A., Roberts, A.P.J. and Fay, D.T. (2017) 'Up periscope: understanding submarine command and control teamwork during a simulated return to periscope depth', *Cognition Technology & Work*, 19(2-3), pp. 399-417.

Stanton, N.A. *et al.* (2020b) 'Returning to periscope depth in a circular control room configuration', *Cognition Technology & Work*.

Stanton, N.A. *et al.* (2013) *Human Factors Methods: A Practical Guide for Engineering and Design*. 2 edn. Farnham, England: Ashgate.

Stanton, N.A. *et al.* (2017a) *Cognitive Work Analysis: Applications, Extensions and Future Directions*. Boca Raton, Florida, United States: CRC Press.

Stanton, N.A. *et al.* (2017b) 'State-of-science: situation awareness in individuals, teams and systems', *Ergonomics*, 60(4), pp. 449-466.

Stanton, N.A. *et al.* (2006) 'Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology', *Ergonomics*, 49(12-13), pp. 1288-311.

Stanton, N.A. and Young, M.S. (2000) 'A proposed psychological model of driving automation', *Theoretical Issues in Ergonomics Science*, 1(4), pp. 315-331.

Stassen, H.G. (1989) 'On the Modeling of Manual Control Tasks', in McMillan, G.R. *et al.* (eds.) *Applications of Human Performance Models to System Design*. Boston, MA: Springer US, pp. 107-122.

Sterman, J. (2000) *Business dynamics: systems thinking and modeling for a complex world*. New York City, NY, United States of America: McGraw-Hill Higher Education.

Stillion, J. and Clark, B. (2015) *What it takes to win: Succeeding in 21st century battle network competitions*. Washington, D.C., US: Assessments, C.f.S.a.B. Available at: https://web.archive.org/web/20160410040054/http://csbaonline.org/wp-content/uploads/2015/07/What-it-Takes-to-Win.pdf.

Stone, R., Caird-Daley, A. and Bessell, K. (2009) 'SubSafe: a games-based training system for submarine safety and spatial awareness (Part 1)', *Virtual Reality*, 13(1), pp. 3-12.

Strauch, B. (2017) 'The Automation-by-Expertise-by-Training Interaction: Why Automation-Related Accidents Continue to Occur in Sociotechnical Systems', *Human Factors*, 59(2), pp. 204-228.

Svedung, I. and Rasmussen, J. (2002) 'Graphic representation of accidentscenarios: mapping system structure and the causation of accidents', *Safety Science*, 40(5), pp. 397-417.

Talcott, C.R., Martinez, S.G. and Stansifer, C. (2007) 'Perception-action icons: an interface design strategy for intermediate domains', *Human Factors*, 49(1), pp. 120-35.

Tamber-Rosenau, B.J. and Marois, R. (2015) 'Attentional Capacity and Limitations', in Toga, A.W. (ed.) *Brain Mapping*. Waltham: Academic Press, pp. 281-288.

Tatasciore, M. *et al.* (2018) 'The Effect of the Degree of Static Automation in Simulated Submarine Track Management', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 62(1), pp. 1668-1668.

Tatasciore, M. *et al.* (2022) 'Should We Just Let the Machines Do It? The Benefit and Cost of Action Recommendation and Action Implementation Automation', *Human Factors*, 64(7), pp. 1121-1136.

Tatasciore, M. *et al.* (2020) 'The Benefits and Costs of Low and High Degree of Automation', *Human Factors*, 62(6), pp. 874-896.

Threston, J.T. (2009) 'The AEGIS Combat System', *Naval Engineers Journal*, 121(3), pp. 109-132.

Tong, F. (2004) 'Splitting the Spotlight of Visual Attention', *Neuron*, 42(4), pp. 524-526.

Torenvliet, G.L., Jamieson, G.A. and Vicente, K.J. (2000) 'Making the most of ecological interface design: the role of individual differences', *Applied Ergonomics*, 31(4), pp. 395-408.

Tractinsky, N., Katz, A.S. and Ikar, D. (2000) 'What is beautiful is usable', *Interacting with Computers*, 13(2), pp. 127-145.

Tran, F.F., Hilliard, A. and Jamieson, G.A. (2017) 'Keeping the Lights On Across the Continent', *Ergonomics in Design*, 25(4), pp. 10-22.

Trentesaux, D., Moray, N. and Tahon, C. (1998) 'Integration of the human operator into responsive discrete production management systems', *European Journal of Operational Research*, 109(2), pp. 342-361.

Turner, D.I.E. (2017) *A Case Study in Design Thinking Applied Through Aviation Mission Support Tactical Advancements for the Next Generation (TANG)*. Master of Science In Information Technology Management. Naval Postgraduate School. Available at: https://apps.dtic.mil/docs/citations/AD1053495.

Upton, C. and Doherty, G. (2008) 'Extending Ecological Interface Design principles: A manufacturing case study', *International Journal of Human-Computer Studies*, 66(4), pp. 271-286.

Van Buskirk, W.L. *et al.* (2019) 'Application of Theory to the Development of an Adaptive Training System for a Submarine Electronic Warfare Task', *Adaptive Instructional Systems*. Springer International Publishing, pp. 352-362.

Van Dam, S.B.J. (2014) *Ecological Approach to Pilot Traffic Awareness*. Delft University of Technology.

Van Dam, S.B.J., Mulder, M. and van Paassen, M.M. (2008) 'Ecological Interface Design of a Tactical Airborne Separation Assistance Tool', *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 38(6), pp. 1221-1233.

van Marwijk, B.J. *et al.* (2011) 'Supporting 4D trajectory revisions on the flight deck: Design of a human–machine interface', *International Journal of Aviation Psychology*, 21(1), pp. 35-61.

van Paassen, M.M. *et al.* (2018) 'Ecological Interface Design for Vehicle Locomotion Control', *IEEE Transactions on Human-Machine Systems*, 48(5), pp. 541-555.

van Paassen, M.M. *et al.* (2013) 'Designing for shared cognition in air traffic management', *Journal of Aerospace Operations*, 2(1-2), pp. 39-51.

van Velsen, L., Ludden, G. and Grünloh, C. (2022) 'The Limitations of User-and Human-Centered Design in an eHealth Context and How to Move Beyond Them', *J Med Internet Res*, 24(10), p. e37341.

Varga, M., Winkelholz, C. and Traeber-Burdin, S. (2017) 'Exploration of User Centered and System Based Approaches to Cyber Situation Awareness', *environment*, 1, p. 2.

Ven, S. (1956) 'The atomic submarine "Nautilus"', *The Soviet Journal of Atomic Energy*, 1(1), pp. 117-118.

Vernon, D., Reising, C. and Sanderson, P.M. (2002) 'Ecological interface design for Pasteurizer II: A process description of semantic mapping', *Human Factors*, 44(2), pp. 222-247.

Vicente, K. (1999a) 'Ecological Interface Design: Supporting operator adaptation, continuous learning, distributed, collaborative work', *Proceedings of the Human Centered Processes Conference*. Brest, France. pp. 93-97.

Vicente, K.J. (2001) 'Cognitive engineering research at Risø from 1962-1979', in Salas, E. (ed.) *Advances in human performance and cognitive engineering research*. New York, NY, USA: Elsevier, pp. 1–57.

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

Vicente, K.J. (2002) 'Ecological interface design: Progress and Challenges', *Human Factors*, 44(1), pp. 62-78.

Vicente, K.J. (1996) 'Improving dynamic decision making in complex systems through ecological interface design: A research overview', *System Dynamics Review*, 12(4), pp. 251-279.

Vicente, K.J. (1992) 'Memory recall in a process control system: a measure of expertise and display effectiveness', *Memory & Cognition*, 20(4), pp. 356-73.

Vicente, K.J. (1991a) 'Representation aiding for problem solving in process control systems', *Conference Proceedings 1991 IEEE International Conference on Systems, Man, and Cybernetics*. pp. 1189-1194 vol.2. doi: 10.1109/ICSMC.1991.169849. Available at: https://ieeexplore.ieee.org/document/169849/.

Vicente, K.J. (1991b) *Supporting knowledge-based behavior through ecological interface design*. PhD. University of Illinois. Available at: https://www.ideals.illinois.edu/handle/2142/20135.

Vicente, K.J. (1995) 'Task analysis, cognitive task analysis, cognitive work analysis: what's the difference?', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, pp. 534-537.

Vicente, K.J., Christoffersen, K. and Pereklita, A. (1995) 'Supporting Operator Problem Solving through Ecological Interface Design', *IEEE Transactions on Systems Man and Cybernetics*, 25(4), pp. 529-545.

Vicente, K.J. and Ethier, C.R. (2000) 'Why fluid dynamics matters for display design in process control: commentary on Bennett and Malek', *Human Factors*, 42(3), pp. 451-4.

Vicente, K.J. *et al.* (1996) 'Evaluation of a rankine cycle display for nuclear power plant monitoring and diagnosis', *Human Factors*, 38(3), pp. 506-521.

Vicente, K.J. and Rasmussen, J. (1987) *Cognitive Control of Human Activities and Errors: Implication for Ecological Interface Design* (87-550-1345-7).

References

Vicente, K.J. and Rasmussen, J. (1992) 'Ecological Interface Design: Theoretical Foundations', *IEEE Transactions on Systems Man and Cybernetics*, 22(4), pp. 589-606.

Vicente, K.J. and Rasmussen, J. (1990) 'The Ecology of Human-Machine Systems II: Mediating 'Direct Perception' in Complex Work Domains', *Ecological Psychology*, 2(3), pp. 207-249.

Vicente, K.J. and Rasmussen, J. (1988) *A Theoretical Framework for Ecological Interface Design*. Kongens Lyngby, Kingdom of Denmark: Denmark, T.U.o. Available at: http://orbit.dtu.dk/en/publications/a-theoretical-framework-for-ecological-interface-design(fefa7bd8-fbf5-4407-8fe5-06051aa68006).html.

Vicente, K.J. and Wang, J.H. (1998) 'An ecological theory of expertise effects in memory recall', *Psychological review*, 105(1), p. 33.

Viller, S. and Sommerville, I.A.N. (2000) 'Ethnographically informed analysis for software engineers', *International Journal of Human-Computer Studies*, 53(1), pp. 169-196.

von Bertalanffy, L. (1950) 'The Theory of Open Systems in Physics and Biology', *Science*, 111(2872), pp. 23-29.

Vredenburg, K. *et al.* (2002) 'A survey of user-centered design practice', *Proceedings of the SIGCHI conference on Human factors in computing systems Changing our world, changing ourselves - CHI '02*. Minneapolis, Minnesota, USA. 503460: ACM, pp. 471-478. doi: 10.1145/503376.503460.

Waite, A.D. (2002) *Sonar for practising engineers*. Wiley New York.

Walker, G.H. *et al.* (2006) 'Event Analysis of Systemic Teamwork (EAST): a novel integration of ergonomics methods to analyse C4i activity', *Ergonomics*, 49(12-13), pp. 1345-69.

Walker, G.H. *et al.* (2010a) 'From ethnography to the EAST method: A tractable approach for representing distributed cognition in Air Traffic Control', *Ergonomics*, 53(2), pp. 184-197.

Walker, G.H. *et al.* (2010b) 'From the 6 Ps of Planning to the 4 Ds of Digitization: Difficulties, Dilemmas, and Defective Decision Making', *International Journal of Human-Computer Interaction*, 26(2-3), pp. 173-188.

Walker, G.H. *et al.* (2008) 'A review of sociotechnical systems theory: a classic concept for new command and control paradigms', *Theoretical Issues in Ergonomics Science*, 9(6), pp. 479-499.

Walker, G.H. *et al.* (2010c) 'Same or different? Generalising from novices to experts in military command and control studies', *International Journal of Industrial Ergonomics*, 40(5), pp. 473-483.

Wang, W. (2016) 'Development of a Baseline Workload Model for Future Submarine Sonar Crewing Analysis', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Washington, District of Columbia, United States of America. Human Factors and Ergonomics Society, pp. 670-674. doi: 10.1177/1541931213601154. Available at: http://pro.sagepub.com/content/60/1/670.abstract.

Wang, W. *et al.* (2003) 'Ecological driver-vehicle interface design based on situation awareness', *Intelligent Vehicle Symposium, 2002. IEEE*. IEEE. doi: 10.1109/ivs.2002.1187977. Available at: http://dx.doi.org/10.1109/ivs.2002.1187977.

Waterson, P., Le Coze, J.C. and Andersen, H.B. (2017) 'Recurring themes in the legacy of Jens Rasmussen', *Applied Ergonomics*, 59(Pt B), pp. 471-482.

Watson, M.O. and Sanderson, P. (2007) 'Designing for attention with sound: Challenges and extensions to ecological interface design', *Human Factors*, 49(2), pp. 331-46.

Webber, J. (2015) *HMCS VICTORIA page 2*. Available at: https://web.archive.org/web/20180531120225/http://webberswarships.ca/styled-65/index.html (Accessed: 13/11/2023).

Wells, W.H. (2011) *Development of a Cognitive Work Analysis Framework Tutorial Using Systems Modeling Language*. University of Central Florida Orlando, Florida.

Wells, W.H. *et al.* (2011) 'Application of Systems Modeling Language (SySML) for Cognitive Work Analysis in Systems Engineering Design Process', *Journal of Universal Computer Science*, 17(9), pp. 1261-1280.

Wickens, C.D. *et al.* (2015) 'Complacency and Automation Bias in the Use of Imperfect Automation', *Human Factors*, 57(5), pp. 728-739.

Williams, A. (2009) 'User-Centered Design, Activity-Centered Design, and Goal-Directed Design: A Review of Three Methods for Designing Web Applications', *Proceedings of the 27th ACM international conference on Design of communication*. Bloomington, Indiana, USA. Association for Computing Machinery, pp. 1–8. doi: 10.1145/1621995.1621997. Available at: https://doi.org/10.1145/1621995.1621997.

Winchell, J.D., Panell, R.C. and Pickering, E.J. (1976) *A Personnel Readiness Training Program: Operation of the AN/BQR-20A* (ADA033435). Available at: https://apps.dtic.mil/sti/citations/ADA033435.

Womble, B. *et al.* (2011) *Delivering savings with open architecture and product lines* (NPS-AM-11-C8P02R01-022). Available at: https://calhoun.nps.edu/handle/10945/33605.

Woods, D.D., Patterson, E.S. and Roth, E.M. (2002) 'Can We Ever Escape from Data Overload? A Cognitive Systems Diagnosis', *Cognition, Technology & Work*, 4(1), pp. 22-36.

Wright, H., Mathers, C. and Walton, J.P.R.B. (2013) 'Using visualization for visualization: An ecological interface design approach to inputting data', *Computers & Graphics*, 37(3), pp. 202-213.

Wróbel, K., Montewka, J. and Kujala, P. (2017) 'Towards the assessment of potential impact of unmanned vessels on maritime transportation safety', *Reliability Engineering & System Safety*, 165, pp. 155-169.

Wróbel, K., Montewka, J. and Kujala, P. (2018) 'Towards the development of a system-theoretic model for safety assessment of autonomous merchant vessels', *Reliability Engineering & System Safety*, 178, pp. 209-224.

Wu, L. *et al.* (2016) 'EID vs UCD: A Comparative Study on User Interface Design in Complex Electronics Manufacturing Systems', *Engineering Psychology and Cognitive Ergonomics*. Springer International Publishing, pp. 354-362.

Xu, W., Dainoff, M.J. and Mark, L.S. (1999) 'Facilitate complex search tasks in hypertext by externalizing functional properties of a work domain', *International Journal of Human-Computer Interaction*, 11(3), pp. 201-229.

Yan, S. *et al.* (2022) 'Constructing a Model to Discriminate the Workload Level of Ship Interface Operators', *Journal of Marine Science and Engineering*, 10(8), p. 1098.

Young, M.S. and Birrell, S.A. (2012) 'Ecological IVIS design: Using EID to develop a novel in-vehicle information system', *Theoretical Issues in Ergonomics Science*, 13(2), pp. 225-239.

Young, M.S. *et al.* (2007) 'Mental Workload in Command and Control Teams: Musings on the Outputs of EAST and WESTT', *Engineering Psychology and Cognitive Ergonomics*. Berlin, Heidelberg. Springer Berlin Heidelberg, pp. 455-464.

Zarnich, R.E. (1999) 'A fresh look at broadband passive sonar processing', *Proceedings of the 1999 Adaptive Sensor Array Processing Workshop (ASAP '99), MIT Lincoln Laboratory, Lexington, MA*. pp. 99-104.

Zestic, J. *et al.* (2019) 'A Novel Auditory Display for Neonatal Resuscitation: Laboratory Studies Simulating Pulse Oximetry in the First 10 Minutes After Birth', *Human Factors*, 61(1), pp. 119-138.

Zhiyin, T. and Lin, H. (2009) 'Research on submarine counter-detection probability by threat surface ship sonar', *4th International Conference on Computer Science & Education*. Nanning, China. IEEE, pp. 81-85.