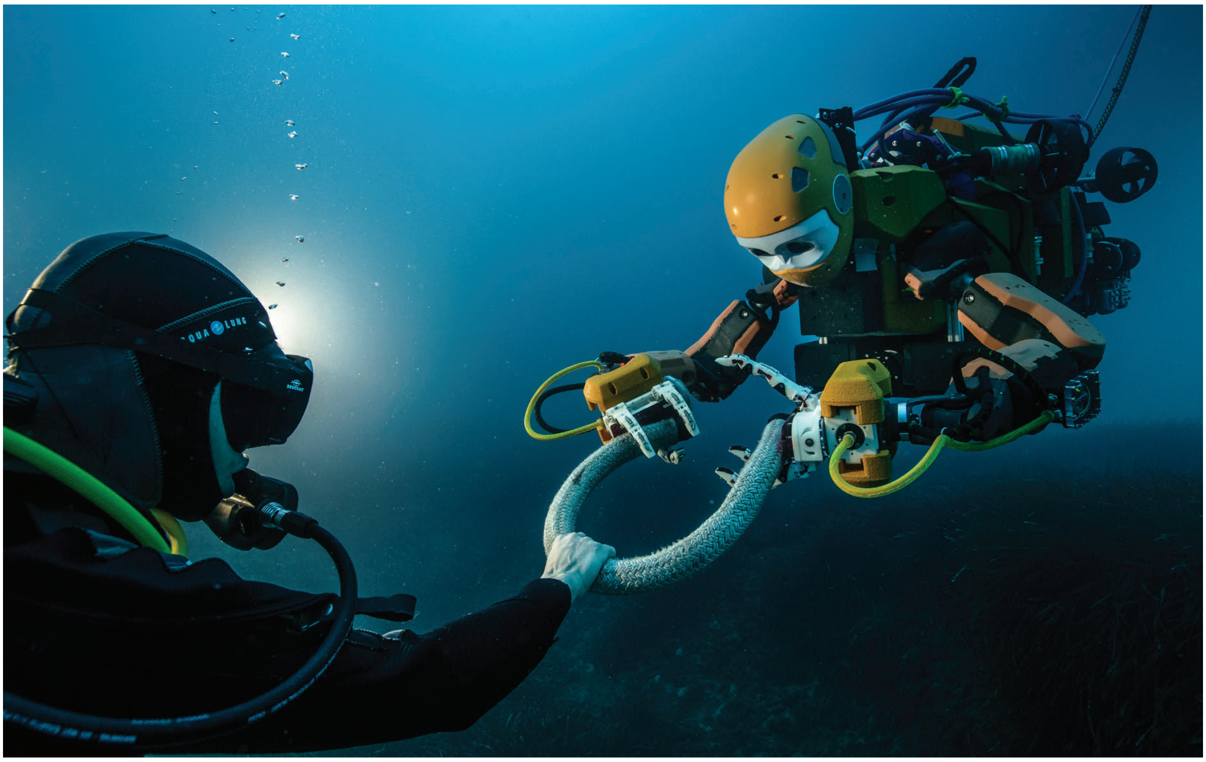
RAND AUSTRALIA

Research Report

MARIGOLD BLACK, LINDA SLAPAKOVA, PAOLA FUSARO, JAMES BLACK

Supporting the Royal Australian Navy's Campaign Plan for Robotics and Autonomous Systems

Human-Machine Teaming and the Future Workforce



For more information on this publication, visit **www.rand.org/t/RRA1377-2**.

Published by the RAND Corporation, Santa Monica, Calif., and Canberra, Australia

Cover: T. Seguin / F. Osada, Stanford University / DRASSM.

Cover image design: Rick Penn-Kraus

About This Report

The Royal Australian Navy (RAN) is in the process of modernising its forces to better address the growing challenges being faced by Australia in the Indo-Pacific region. Increased state competition in the region combined with accelerated military modernisation have caused the RAN to consider the use of robotics, autonomous systems, and artificial intelligence (RAS-AI) in maritime roles and missions. Innovative technologies and new ways of conducting military operations will be needed to deliver such systems and, as is highlighted in the *2020 Defence Strategic Update*, to 'shape Australia's strategic environment, deliver credible deterrence and respond to challenges to our interests' (Department of Defence, 2020, p. 7).

This report is a continuation of work conducted in support of the RAN's *RAS-AI Strategy 2040*, released in 2020. RAND Australia was asked to provide policy analysis and advice to support development of an actionable RAS-AI Campaign Plan that could assist RAS-AI implementation efforts. Our research team examined three specific areas to support development of an actionable plan: military innovation, a missions and technology assessment for maritime RAS-AI, and human-machine teaming (HMT). This document focuses on the final of the three listed areas. This work should inform the RAN, other Australian Defence services, and Defence more broadly about the implications of HMT for the RAN's future workforce.

This work was conducted by RAND Australia. RAND Australia is RAND's Canberra-based subsidiary that analyses defence, national security, economic, and social issues for Australian clients. With a commitment to core values of quality and objectivity, RAND Australia combines local research talent with world-class experts from across RAND's global presence to solve complex Australian public policy problems.

For more information on RAND Australia or to contact our director, please visit www.rand.org/australia.

Contents

Table

Summary

Issue

In 2020 the Australian Department of Defence released the *2020 Defence Strategic Update*, which outlined security challenges it expects Australia to face in the Indo-Pacific in coming decades. To adjust to the new strategic and security environment, the Australian military is engaging in a modernisation process, inherent to which are technological advancements, innovative thinking, and strong engagements with industry and academia. The Royal Australian Navy (RAN) intends to invest in an array of new capabilities to meet this developing threat environment. In 2020, the RAN released its *RAS-AI [Robotics, Autonomous Systems, and Artificial Intelligence] Strategy 2040*, which examines the challenges and potential capabilities that RAS-AI systems could bring to the maritime domain in air, surface and undersea operations. As part of an effort to develop a Campaign Plan to implement the strategy, RAND Australia conducted research to study the implications of RAS-AI and, more specifically, human-machine teaming (HMT) on the RAN's future workforce.

Approach

The report draws on a review of relevant open-source academic and grey literature, with a focus on identifying possible lessons for the RAN. The analysis concentrated on the overall RAS-AI needs of Defence and the impacts on workforce and skills, with a particular focus on the implications of HMT for the Defence workforce, broadly defined.

The first chapter develops the HMT concept with a view to its specific relevance for the future workforce. The concept of HMT captures the effective integration of humans and RAS-AI into warfighting systems to gain an operational advantage over adversaries (UK Ministry of Defence [MOD], 2018).

The second chapter builds on the broad consensus in existing literature that RAS-AI, and HMT specifically, is likely to significantly change the nature of many military functions as well as the requisite skills (Ministère des Armées, 2019; Morgan et al., 2020). These effects are likely to be felt across different parts of Defence that interact with RAS-AI and may result in a growing need to adapt workforce management, recruitment and retention efforts. This report looks at the implications of HMT for the future workforce, and specifically in relation to skills—both technical and non-technical—training, and organisational learning.

Key Findings

The following findings underscore the importance of understanding that HMT requires a fundamentally different, and decidedly novel, way of working:

- Normalisation of HMT in the RAN will require flexible management of the Defence workforce, and continuous adaptation of existing structures and concepts.

- HMT necessitates a shift in cognition as much as training and perception.
- The goal in HMT is to optimise the interaction and leverage the strengths of both human and machine.
- HMT encompasses broad and complex issues and defies categorisation into an ineluctable laundry list of principles, activities and resources.
- Understanding and considering the spectrum of human-machine interaction (HMI) is integral.
- New technological developments should be aligned with actual problems/needs.
- The skills/attributes required for HMT may not traditionally be prized by Defence. Organisational learning must be embraced to deliver this challenging capability.
- HMT should be leveraged, not only as an effective military fighting capability but also as a concurrent training capability.
- Success in HMT requires significant uplift, across the breadth of the workforce.
- HMT requires identifying the strengths and weaknesses of humans and machines and capitalising on advantages, so they become greater than the sum of their parts.
- HMT focus should be on problems to be solved, the appropriate ratio and mode of HMI and finding the sweet spot in terms of effort payoff.
- RAS-AI must perform such that they instil a sense of trust, safety and reliability in those who use them. This includes transparency across interface design, strong user focus, and clear compliance with international humanitarian law.
- There must be proactive attention to the conceptual and ethical complexities of the HMT paradigm, starting from the moment of design.

Acknowledgements

This work was sponsored by CAPT Nicholas Trongale, the inaugural Director of Robotics, Autonomous Systems and Artificial Intelligence (RAS-AI) in Warfare Innovation Navy. Key players on the team included Andrew Kirby, Ryan Pope, CMDR Felicity Petrie and Daniel Livingstone. Discussions between RAND Australia researchers, Nicholas, and his team helped to improve the work.

Conversations with Margaret Sorial and Natalene Di Pietro from the Future Navy Warfare Branch helped to improve and focus the effort on those areas that could most directly help the Royal Australia Navy.

We would like to thank the reviewers, Tim Marler from the RAND Corporation and Carol Smith from Carnegie Mellon University, for their thoughtful and useful comments that helped to improve this document.

Abbreviations

AI	artificial intelligence
HMI	human-machine interaction
HMT	human-machine teaming
LVC	live, virtual and constructive
MOD	Ministry of Defence [UK]
NSAIC	National Security Commission on Artificial Intelligence
R&D	research and development
RAN	Royal Australian Navy
RAS-AI	robotics, autonomous systems, and artificial intelligence
STEM	science, technology, engineering and mathematics

1. Introduction

Challenges for the Integration of Human-Machine Teaming into the Royal Australian Navy

Rapid technological change provides various opportunities and challenges for defence organisations, with corresponding implications for their workforce and personnel policies. Effective exploitation of new and emerging technologies therefore requires defence organisations to adapt their workforce to reflect the changing nature of various military functions as well as the military profession overall. It also involves adapting models for recruitment, learning, development and retention to ensure sufficient quantity (capacity) and quality (capability) of personnel (Winkler et al., 2019). This holds for robotics, autonomous systems, and artificial intelligence (RAS-AI)—a cluster of new and emerging technologies whose exploitation is at the core of many efforts to innovate in defence, including in Australia and the United Kingdom.

As the strategic and operational landscape shifts, the technologies associated with RAS-AI provide an 'inflection point' in the delivery of military advantage and transformation. The Royal Australian Navy (RAN) needs to capitalise on this by understanding how to best integrate technologies into the workforce to unlock this advantage. Indeed, this becomes an urgent consideration as adversaries seek to unlock the same advantage, and as new technologies and capabilities accelerate and proliferate (UK Ministry of Defence [UK MOD], 2018, p. iii).

To inform the RAN's ongoing efforts to facilitate RAS-AI integration, including through the development of a Campaign Plan, this report provides an overview of the various impacts of RAS-AI on the Defence workforce as well as the corresponding implications for workforce structure, skills requirements, and workforce management. In particular, it considers where there are specific challenges and opportunities associated with developing the optimal combination of crewed and uncrewed systems and teams, and effectively balancing human and machine cognition across a variety of tasks and functions (UK MOD, 2018, p. 44).

The RAN's *RAS-AI Strategy 2040* recognises that 'realising Navy's RAS-AI vision relies on its people. The ability to plan for, acquire, operate and maintain RAS-AI capabilities is contingent on growing a workforce who have a deep understanding of the Australian maritime operating environment and RAS-AI technologies' (RAN, 2020, p. 19). In alignment with the work of the RAN's Future Navy Workforce, integration of RAS-AI into the RAN means identifying and optimising the skills, knowledge, attributes and proficiencies required for an RAS-enabled workforce and battlespace. This may manifest as greater situational awareness; technological skills and know-how to enable lighter physical and cognitive loads; sustainment with increased anticipation and efficiency; greater force protection; and superior manoeuvre possibilities across all domains, at the very least (UK MOD, 2018, p. 1). A recent report published by the European Commission on defence-related skills for the contemporary context and the future (Galai et al., 2019) highlighted skills 'as a key ingredient of European defence industrial competitiveness and resilience in supplying the capabilities needed for safeguarding Europe's strategic interest and autonomy' (Galai et al., 2019, p. vii).

A key area for attention in the future trajectory of the RAN workforce is developing greater understanding of the opportunities afforded by RAS-AI in human-machine interactions (HMIs), and specifically human-machine teaming (HMT). How the RAN approaches the interactions its people have with RAS-AI may well be fundamental to the futurity of the force. The critical consideration is the advantage to be gained from adopting and incorporating autonomous systems as it is weighed against the impact on the RAN's human workforce. In the RAS-AI context, that process of optimisation for crewed and uncrewed platforms should be seen as a rich opportunity space ripe for positive exploitation. It offers an actionable means by which to modernise and enrich the human capabilities of the existing workforce and amend its orientation such that it recognises the connective tissue that now exists between military skills and those in the civil sector.

The *RAS-AI Strategy 2040* articulates the need to 'normalise human-machine teaming', stating that because 'RAS capabilities are integrated into the maritime force, these systems will need to work seamlessly with existing crewed and additional RAS systems' (RAN, 2020, p. 19). This process of 'normalisation' will be 'developed through *ab initio* training and revisited periodically throughout career courses' (RAN, 2020, p. 19). In practice, however, the normalisation of HMT necessitates a shift in cognition as much as training and perception. Ultimately, it is the human's internal acceptance of the robot into the workforce and warfighting systems, broadly defined, that will determine the nature, and indeed the success, of HMI. This represents a completely different way of working, both cognitively and physically.

The benefits of a RAS-AI–enabled workforce may well go beyond force protection and force-projection priorities. Through HMT and augmentation there is an opportunity to enhance human cognitive attributes and abilities, to see and establish new avenues for development and training, to construct novel human-computer interfaces and modes of interaction, and to identify new operating concepts that will capitalise on the qualities of the human-machine team. Successfully exploited, these opportunities may speak to fundamental force culture transformation, and may generate an evolving catalogue of new capabilities and relationships. As the American psychologist B. F. Skinner (1969, p. 302) famously argued, 'The real problem is not whether machines think but whether men do.'

Research Method

The report draws on a review of relevant open-source academic and grey literature, with a focus on identifying possible lessons for the RAN. The analysis is concentrated on the overall impacts of RAS-AI on the Defence workforce and skills, with a particular focus on the implications of HMT and RAS-AI for the back-of-house workforce. The report constitutes an unstructured desktop literature review. The adopted structure was devised to parse the literature, breaking it up into several key areas for analysis. It concentrates on articles, reports and accounts that identify the implications of HMT on workforce development, noting many critical challenges as well as suggested courses of action. To bridge the methodological divide that often exists between the technical and the social dimensions of HMT, this review is purposefully selective, prioritising the integration of a diversity of sources from across the field rather than pursuing an exhaustive catalogue of sources.

In our initial scoping of the problem space, the research team identified and extracted sources via RAND Australia’s access to various academic and scholarly databases, screening article titles and abstracts to determine whether they met the primary inclusion criteria. Sources were to include academic studies, government publications, and documents detailing, in the broadest sense, considerations concerning HMT and its implications for workforce development. While we prioritised books, articles and reports published after 2010, we also included works that demonstrate the depth and breadth of the problem space, address issues of relevance not captured by recent studies, or deal with topics identified as of specific importance to the RAN’s development of a RAS-AI Campaign Plan.

To build on specific areas of interest, we performed additional, refined searches in the databases. We did not include relevant literature at the classified or restricted level or material that is otherwise publicly unavailable. We grouped the findings from articles and sources into subcategories that allowed us to look at HMT from a number of dimensions. This approach allowed us to cover a broad range of phenomena, and to conduct a comprehensive and integrative analysis of the issues.

Structure of This Report

The report is broken into two main sections. Chapter Two develops the HMT concept with a view to its specific relevance for the future workforce. The concept of HMT captures the effective integration of humans and RAS-AI into warfighting systems to gain an operational advantage over adversaries (UK MOD, 2018). Due to the need to ensure meaningful human oversight of autonomous systems, continue leveraging inherent and unique strengths of the human mind (e.g., intuitive judgement), and facilitate effective collaboration between humans and systems, HMT is a significant element in the integration of RAS-AI into Defence.

HMT assumes several conditions to be present for true human-machine collaboration to be achieved. These may include bounded autonomy (i.e., autonomy within defined bounds of social contracts, organisational structure, or situational constraints), shared knowledge and awareness that facilitates learning and adaptation, and benevolence to collaboratively advance defined goals (Brill et al., 2018). However, the manner in which a fully seamless blend of human and machine activity is achieved under these conditions remains an open question for Defence.

Chapter Three builds on the broad consensus in existing literature that RAS-AI, and specifically HMT, is likely to significantly change the nature of many military functions as well as the requisite skills (Ministère des Armées, 2019). These effects are likely to be felt across the different parts of Defence that interact with RAS-AI and may result in a growing need to adapt workforce management, recruitment and retention efforts.

This report looks at the implications of HMT for the future workforce, including its impact on skills—both technical and non-technical—training, and organisational learning. Noting the breadth of personnel functions that are likely to be affected by RAS-AI integration, it is important to recognise that the exploitation of RAS-AI requires upskilling of personnel who interact with or use RAS-AI, as well as adaptation to the broader impacts of RAS-AI on the nature of work and careers. This will occur at the individual level and will affect the size, shape and management of the wider Defence workforce, at the aggregate level. It may also include the creation of new workforce roles within Defence, changes in the definition and scope of existing roles, new non-specialist technical and non-technical skills, and workforce management processes to strengthen organisational capacity and culture and facilitate the continuous innovation associated with integration of RAS-AI.

In Chapter Four, we leveraged the findings to highlight items that are especially relevant to HMT, acknowledging that some of the workforce-related issues featured in this report are more general and could potentially apply to any emerging technology.

2. Human-Machine Teaming

The primary objective of this chapter is to define and characterise HMT, to identify challenges with HMT more broadly, and to identify issues relevant to the RAN's RAS-AI Campaign Plan. While it draws in questions that elucidate the breadth and depth of issues informing contemporary understanding of HMT, it selectively elaborates on those that may facilitate RAS-AI integration into the RAN and that demonstrate the potential impacts of RAS-AI on the Defence workforce.

Approaches to Human-Machine Teaming

We need to address, at the outset, the inherent definitional ambiguity and complexity of the term 'human-machine teaming'. Despite the frequency with which it is 'thrown around in the contemporary zeitgeist of human factors', the term remains 'an elusive psychological construct' (Brill et al., 2018, p. 457). HMT is an evolution of the idea of HMI, an idea popularised in the 1980s that referred to how people interact with computers, and theorised that humans and computers 'engage in a communicative dialogue whose purpose is the accomplishment of some task' (Card, Newell and Moran, 1983, p. 4). HMT, which is also referred to as human-autonomy teaming, may be defined as 'interdependence in activity and outcomes involving one or more humans and one or more autonomous agents, wherein each human and autonomous agent is recognized as a unique team member occupying a distinct role on the team, and in which the members strive to achieve a common goal as a collective' (O'Neill et al., 2020, p. 22).

What that looks like in practice can vary greatly and may be determined by how one decides a team is constituted. Here, Brill et al. (2018, p. 455) point to how we identify the qualities of human teams as a useful analogue:

> Teammates typically perform complementary, largely non-redundant functions. Teammates have bounded autonomy, freedom to act according to one's judgment, but within the limits of social contracts, organizational structure, or situational constraints. Autonomy mandates non-deterministic behavior; however, the aforementioned constraints should limit unpredictability of actions and outcomes. Teammates have shared knowledge and shared awareness . . . and as such, they learn and change. With sufficient experience and a positive history, teammates will learn to trust one another, and their collective experiences will lead to enhanced transparency. Lastly, teammates act with benevolence to help one another and to further the team's goals.

Brill et al. (2018, p. 455) suggest that for true HMT, these conditions must be met, with the most critical element being the presence of 'true autonomy (even if bounded)'. Without this autonomy, the technological artefact, even if it is an advanced form of automation, is not a teammate but merely a tool. Moreover, in most cases these machines have been utilised and conceptualised as entities that are separate to the human. Modern interactions are moving increasingly towards the human developing a *relationship* with the autonomous entity, however (de Visser, Pak and Shaw, 2018).

As such, there has been increasing interest among researchers in understanding human-robot interactions by using a more social perspective or orientation, and considering robotic platforms as teammates, or partners. This shift comes from the awareness that increased complexity in autonomous systems has also resulted in systems that are more brittle, less transparent, and that require a greater degree of human oversight. Advances in automation have generated an environment whereby human oversight itself is becoming more cognitively onerous, and many of the intended benefits of automation are negated (Madni and Madni, 2018, p. 4). Continuing to reap the benefits from emerging technologies in this context necessitates the development of a new cognitive paradigm that allows for optimisation of both the human and the machine.

The HMT paradigm rethinks the nature of the interaction between the human and machine, such that it allows for a more coactive interaction—that is, the human and robotic agents are acting in concurrence or together. This type of interaction may well facilitate greater opportunity and augmented capability. As was articulated in a recent report published by the National Security Commission on Artificial Intelligence (NSCAI, 2021, p. 34), 'Synergy between humans and AI holds the promise of a whole greater than the sum of its parts'. What the HMT does is try to retain 'the benefits of automation—mainly efficiency—while minimizing two of its chief costs and hazards—especially brittleness (the inability to adapt to new situations and contexts) and alienation of the operator' (Cesafsky, Stayton and Cefkin, 2019, p. 67).

But 'what does it mean to "partner" with technology and what are the catalysts that move an autonomous system from being viewed as a tool to the perception that it is, in fact, a teammate?' (Brill et al., 2018, pp. 457–458). Determining what HMT looks like is highly complex, and somewhat divisive. Building a team from humans and robots requires engagement with a suite of practical and technical considerations, as well as a dynamic schema of research on the philosophical dimensions of having machines as teammates—from the robot side as much as the human (Cesafsky, Stayton and Cefkin, 2019, p. 67; see also Schaefer et al., 2017).

The underpinning logic of HMT is in understanding the elements and characteristics of what it means to be an effective team player in a human team and translating some of the features to the multiagent or human-machine team. There is some research to suggest that HMIs should, by design, emulate those rich interactions that define human relationships, and look to human-human models as the standard, at least in the near term. As de Visser, Pak and Shaw (2018, p. 1410) have observed, 'Autonomy should be able to take advantage of extant human capabilities of detecting incidental information from others. With that in mind, and inspired by models of human–human interaction, we propose that future design of autonomy ought to take cues from the social sciences'.

Against this approach, RAS-AI systems would be strategically designed so that they have styles/methods of interaction that leverage the nuances of human teaming relationships (Lyons and Havig, 2014, p. 182). This requires identification of some of the more intangible qualities that define human-to-human teams, which may include

- effective methods of communication
- trust
- group morale
- leadership
- organisation
- monitoring and provision of feedback on individual and collective performance
- individual values
- mental and physical aptitude
- external factors
- idiosyncratic and often flawed mental models developed by humans
- difficulty in communicating those models consistently across all agents.

In theory, when this kind of conceptual approach is adopted, the interface will follow logically. The associated transparency will enable shared intent/awareness between teammates (Lyons and Havig, 2014, p. 189).

Moreover, there are studies that show that fashioning HMIs such that they mimic or emulate human relationships can improve HMT dynamics in other ways. This approach may assist in producing behavioural trust that is informed by neurological processes most closely associated with human-human interactions. The design of automated systems would incorporate consideration of the task, the individual, and the level of anthropomorphism required to achieve a certain result (de Visser et al., 2017).

However, this does not necessarily mean that machines in the HMT environment should necessarily look like, and mirror, humans. There is empirical evidence to suggest that there are broad differences between individuals in how the likeness of robotic partners is perceived, no matter the level of sophistication and resemblance (Lyons and Wynne, 2021). While it is vitally important to understand, as has long been recognised in computers-as-social-actors research (Nass and Moon, 2000), that humans have a strong predisposition to respond to technology in social ways (Lyons and Wynne, 2021), it is also important to recognise that many human characteristics that enrich human-to-human team dynamics are not easily applicable to an automated agent (Joe et al., 2014). As Haimson et al. (2019, p. 1690) ask, 'Is a more human-like machine teammate necessarily a better machine teammate?' Clifford Nass and Victoria Groom argue that machines cannot become full-fledged teammates. Because they are 'lacking humanlike mental models and a sense of self, robots may prove untrustworthy and will be rejected from human teams'. (Groom and Nass, 2007, p. 483).

In the contemporary academic discourse, the way to navigate this complex paradigm of HMI is to understand and identify the strengths and weaknesses of humans and machines in their interactions, but also to capitalise on the advantages of each so that they become greater than the sum of their parts. Developed in 1951, the Fitts list identified that humans are better at perceiving patterns; improvising and using flexible procedures; timely recall of relevant facts; reasoning inductively; and exercising judgement. Machines are better at responding quickly to control tasks; performing repetitive and routine tasks; reasoning deductively; handling simultaneous complex tasks; and fast and accurate computation (Fitts, 1951, p. 10). While there has been considerable analysis of the role/function allocation of humans and machines since the development of the Fitts list, it is important to understand that the 'teaming' paradigm is somewhat removed from the role/function allocation approach. The intent is to 'design a seamless integration of human and technological capabilities into a well-functioning sociotechnical system' (Behymer and Flach, 2016, p. 105).

The point here is 'to think "combine and succeed" not "divide and conquer"' (Johnson et al., 2014, p. 77). This requires considered thought on how to use the machine to best effect but also, much more importantly and significantly, how to use the human to best effect. Fully exploiting RAS-AI is not just about 'machines learning autonomously' or 'humans teaching machines' or even 'machines teaching humans' (Ransbotham et al., 2020, p. 8). It involves understanding how to tailor the interdependence relationship between human and machine and assume a synergic relationship as part of the design. And this extends to decisionmaking in HMT.

Paul Scharre argues that while much of the discourse around autonomy suggests a choice between human versus autonomous decisionmaking, 'Hybrid human-machine cognitive architectures will be able to leverage the precision and reliability of automation without sacrificing the robustness and flexibility of human intelligence' (Scharre, 2016, p. 151). This approach to HMT is demonstrated in what has come to be known as 'Centaur Chess' where humans and AI play on the same team. The AI identifies possible moves and vulnerabilities that the human partner may not have noticed, allowing the human player to 'manage strategy, prune AI searches to focus on the most promising areas, and manage differences between multiple AI' (Scharre, 2016, p. 153). Under these circumstances, the AI can drive players toward excellence by charting a series of faultless opening moves, and generate recommendations as the game continues (Cassidy, 2014). For example, AI may make recommendations or illuminate problems or solutions that the human may never have identified or considered. The human is called on to act against these new problems or ideas. The most effective HMT accents the empowerment of the human.[1]

Modes of Interaction

One of the key areas to be attended in the HMT paradigm is the development of common goals and understandings in HMT, including the spectrum of HMI. In this context, it is important to focus on the problems to be solved, identify the ratio and mode of HMI that will most effectively address those problems, and 'assess the sweet spot in developing effort payoff' (Johnson et al., 2014, p. 76). Understanding context and criticality and aligning technologies and the mode of HMT with actual problems or needs, is imperative. We know that in many areas, machines are smarter than humans; they can augment our cognitive strengths, conduct

[1] The Chinese military appear to be taking the teaming or, more accurately, 'hybridisation' concept further. As part of its military modernisation programs, the People's Liberation Army is conceptualising a new style of warfare which greater emphasises cognitive dimensions. It is investing in strategies that encourage 'mental dominance' and which improve '"the field of decision-making through reasoning," as the speed and complexity of conflict continue to increase'. China is planning to progress HMT to new level by 'blending human and machine intelligence, including through leveraging insights from brain science and such techniques as the use of brain-computer interfaces'. Kania, 2020, pp. 86 and 84.

higher-level tasks, and replicate human skills to extend our physical capabilities. But the 'combination of human and machine abilities is most likely to get the job done reliably and cost-effectively. . . . These are important characteristics for any work system that expects to accomplish complex work in the real world' (Johnson et al., 2014, p. 75).

To effectively respond to a specific problem, understanding the spectrum of HMI is important. Drawn from a recent Massachusetts Institute of Technology (MIT) report looking at expanding the impact of smart machines in organisational learning, the following five modes of HMI may offer a means by which to identify not merely what HMT could look like, that is, the level of machine vs human dominance, but perhaps more importantly, why, and where to team. According to the report (Ransbotham et al., 2020, pp. 11–12), the five modes are

- Automator: The machine has most of the context and the ability to make decisions. Human involvement would hinder processes.
- Decider: The machine can capture much of the context and make decisions, but humans implement solutions.
- Recommender: When organisations make large numbers of decisions repeatedly, the machine incorporates much, but not all of, the business context.
- Illuminator: During creative work that requires human thought, the machine provides insights to inform the process.
- Evaluator: When multiple hypothetical solutions are developed by the human, the machine assesses the many complex dependencies.

What these modes offer is the means for personnel within an organisation to work together as well as learn from each other over time, and to do so in the right way, and within optimal contexts (Ransbotham et al., 2020, p. 1). As the MIT report states, 'Matching the right mode of human-machine interaction to the situation affects whether feedback advances or limits organisational learning' (Ransbotham et al., 2020, p. 10). When these contexts are understood, and organisations are able to define multiple, effective ways for humans and robotics systems to work and learn together, significant benefits and advances may be generated. The higher the number of modes that are used in an area, the more significant the change and the higher the returns in HMT. As the report concludes, 'Broader competencies allow organizations to fit a wider variety of interaction modes to a wider variety of situations' (Ransbotham et al., 2020, p. 11).

But while these modes offer a point of clarity, assessing the appropriate HMT model should not necessarily be a straightforward task. As in any team, recognition of the dynamic elements is integral to engaging in successful teamwork. As scholars from the Alliance Innovation Laboratory in Silicon Valley observe, effective teamwork is an intricate engineering challenge that requires generating 'actual coordination of complex activities such as communication, joint action, and human-aware execution to successfully complete a task, with potentially shifting goals, in varying environmental conditions mired in uncertainty' (Cesafsky, Stayton and Cefkin, 2019, p. 67; see also Seeber et al., 2019, p. 3). Developing a suitable, effective, and optimal mode of HMT should be seen as an experiential issue that is unique and native to a specific system. It therefore requires, and allows for, more inventive and dynamic agency and trust calibrations than other automation paradigms (Cesafsky, Stayton and Cefkin, 2019).

HMT also requires transparency in terms of the intent or objective of the human and the machine in the partnership. And here, the development of ‘a method to establish shared intent and shared awareness between a human and a machine’ is seen as a ‘way to reduce uncertainty regarding the performance and or behavior of an automated tool’ (Lyons and Havig, 2014). In HMT, transparency is required across multiple dimensions including (but not limited to) machine operations, decisionmaking, evolving priorities, human intervention and revision, goal setting, and constraints (Madni and Madni, 2018, p. 4). The goal in system transparency is to understand the behaviour, intentions and future objectives of a system. Precisely the level of transparency required is still a matter of scholarly debate, however, as are determinations on the kinds of data, information and signals to be communicated (de Visser, Pak and Shaw, 2018, p. 1412).

Cognitive Load

One of the primary considerations for ensuring that the RAS-AI workforce is in an optimal state for performing their role is developing a keen understanding of cognitive load and span of command—that is, the number of active elements an individual can control at the same time without undermining their ability to undertake higher-level thinking (UK MOD, 2018). For example, while in combat, soldiers may have to manage several uncrewed air, ground and underwater vehicles, while also having to operate alongside other humans. This will require strong intellectual capacity and resilience.

These considerations are becoming increasingly important as more conventional platform-centric warfare is being superseded by a network-centric model. In the latter, advantage is seen to be delivered by superior capacity for strategic utilisation of information, and the ability to coordinate kinetic and non-kinetic effects across various systems (Galai et al., 2019, pp. 18–19). HMT is fundamentally about using that partnership to get the very best from the operator with regard to both mental and physical capacities so that opponents can be outmanoeuvred and outsmarted. It follows, then, that RAS-AI should be used to free up, maximise and elasticise mental capacity.

However, this is not merely about downsizing human involvement (Johnson et al., 2014, p. 79). Mission command is made sensible by optimising independence of subordinate action so that greater initiative and tempo is generated but is also balanced by unifying measures and risk management. In the RAS-AI context, mission command is applied through the dynamic management of levels and types of automated processes, balancing the risk with the advantage gained from machine capability within changing contexts (UK MOD, 2018, p. 44).

As articulated by the UK MOD (2018, p. 44), the idea of an optimal span of command is largely ‘driven by human cognitive loading and how many active elements an individual can control, even where the interpersonal demands like leadership are absent’. For example, in cases where the human operator is ‘task-saturated’ due to managing automated systems and unanticipated behaviours in environments that are increasingly technologically complex, there may be limited cognitive bandwidth for maintaining complex multitasking or high-level thinking (UK MOD, 2018, p. 44). Humans are increasingly facing a deluge of data in conflict situations, and ambiguity in data and decisionmaking constitutes a major challenge for computer algorithms (UK MOD, 2018, p. 51). Given that the complexity and tempi of future warfare will increase, it may be beneficial—perhaps even essential—to enhance cognitive human performance to deal with data overload, remaining attuned to the suite of legal, ethical and social complexities it brings with it (Scharre, 2015).

Yet, to assume that RAS-AI will simplify decisionmaking processes or increase the span of command is to disregard the fact that autonomous systems tend to have harder limits than humans and are typically less able to operate beyond their design parameters. The consequences of this are that machines will suffer catastrophic failure or will identify system limitations and the need for human intervention at moments of acute stress. This puts the problem in the hands of a human when they may not be sufficiently attentive, or when there is limited opportunity to adequately understand the problem and act to prevent failure (UK MOD, 2018, pp. 31–32). Paul Schutte, a senior researcher in human factors and human-computer interaction at the NASA Langley Research Center, observes, 'In these times, the human must perform a task that he or she rarely performs—a task that is complex or unusual'; they are the 'last resort and the final fail-safe' (Schutte, 2017, p. 237).

The problem in the human taking on this role of 'exception handler' is that as autonomous systems become more reliable and effective, the frequency of situations requiring the intervention of the pilot will be reduced. While this might be seen to be, overall, a positive development, when we consider what this requires of the human, we can see a significant problem. As Schutte (2017, p. 238) argues, the irony here is that 'because these cases are rare, the human has less experience and skills for handling them' and therefore has a higher propensity for failure. Accordingly, some level of human involvement in tasks that are being allocated to machines may be required in order to ensure continued engagement and maintenance of skills and competencies.[2] The natural limits of human mental capacity also highlight the need to be able to dynamically alter the level and kind of control exercised by human operators over the systems to develop and improve team effectiveness (UK MOD, 2018, p. 45).

Trust

Based in the School of Global, Urban and Social Studies at the Royal Melbourne Institute of Technology University, Aiden Warren and Alek Hillas (2020, p. 824) argue that 'humans use trust as a substitute for total reliability', but trust issues in HMI, and specifically teaming, extend well beyond questions of reliability. Ensuring an optimal measure of trust is integral

[2] 'Furthermore, human attention is neither constant nor consistent. Simply monitoring systems holds people's attention poorly. It is often very difficult for a previously unengaged person to be able to ramp up their mental alertness at a point of crisis or orient themselves sufficiently quickly to the key variables and context in time to act' (UK MOD, 2018, p. 32).

to the success of human-machine teams, but it is also vital to enculturating RAS-AI into the broader organisation. End-user trust in the performance and reliability of RAS-AI systems will be critical to generating acceptance among military personnel and enabling effective adoption and exploitation of RAS-AI in different mission contexts. But what is trust? What does it mean to have trust in a machine/robot or an AI, and what are the implications of that understanding for the development of effective and optimal HMT? Thomas B. Sheridan, pioneer of robotics and remote-control technology, distinguished between 'trust as an effect or outcome of certain automation characteristics (e.g., reliability) and trust as a cause of operators' behaviour when utilizing automation'. He argues that 'human operator trust in automation is now a major topic of interest because it significantly affects whether and how automation is used' (Sheridan, 2002; see also Madhavan and Wiegmann, 2007).

There are important variations in the ways humans perceive and react to non-human agents compared with their human teammates. While humans can, and often do, treat machines in a social manner, they often come to different decisions and exhibit different patterns of brain activation in decisionmaking when working with a machine than they do when collaborating with other humans. Recent research has shown that 'people experienced less emotion and spent less effort inferring mental states with machines than with humans. These findings are compatible with research showing that people perceive, by default, less mind in machines than in humans'. The implications of this are that the denial of mind or perception of inferiority, in general terms, can result in prejudice or discrimination (Haslam, 2006) and can form the basis for out-group discrimination or segregation (Crisp and Hewstone, 2007; Tajfel and Turner, 1986). This is important for HMT because any loss of 'trust' and collaborative impetus will undermine the human-machine partnership.

But as HMT becomes more common, and as technologies allow for new kinds of interactions, we may also see other issues around trust emerge. For example, preliminary research has shown that when using technologies such as Brain-Computer Interface, an expanded form of HMT where technologies allow the human brain to communicate directly with machines, strong human-robot attachment may develop, and sometimes even a feeling of 'self-extension into the robot' (Binnendijk, Marler and Bartels, 2020, p. 23). While this suggests that human-machine trust may be enhanced in these contexts, what impact might it have on how humans determine appropriate levels of trust? And to consider the issue from a slightly different angle, what effect could it have on interpersonal relationships and broader unit cohesion? What may be as important to consider in HMT is not just trust dynamics between one human and one machine, but how trust manifests and effects dynamics when there is a multiplicity of human and robotic agents working together.

There are also factors beyond the social dynamics of HMT that may influence the human's willingness to work in partnership with RAS-AI and may result in institutional vulnerability. These include, but are not limited to, uncertainties over how RAS-AI technologies function, the quality of inputs and outputs for RAS-AI (e.g., the impact of algorithmic biases on how RAS-AI perform), and the broader organisational impacts of RAS-AI (e.g., uncertainties over the impact of automation on jobs and employment). According to Jai Galliot (2018, p. 124), the notion of trust in the RAS-AI context is 'defined by a power process, the capacity of soldiers to endure subjection to technology and the extent to which automation impinges upon one's autonomy or otherwise impacts the soldier's wellbeing'. An increase in the delegation of tasks to RAS-AI, particularly in terms of decisionmaking is likely to raise a raft of trust-related anxieties among warfighters (Binnendijk, Marler and Bartels, 2020).

To address some of these issues, there are ongoing efforts to define how trust-building can be effectively integrated into the design of RAS-AI and relevant validation and testing processes. Trust-building will also require efforts at the research and development (R&D) level to detect and mitigate algorithmic biases, ensure transparency, safety, interpretability, and legal compliance of RAS-AI technologies to ensure reliability, ease of use, and familiarity of RAS-AI (Brill et al., 2018). Some of these reservations may be offset by the growth of popular trust in comparable technologies in the civilian sector (Binnendijk, Marler and Bartels, 2020, p. 23).

Of particular importance is the need for trust to be calibrated. Trust calibration occurs when individual team members adjust their attitudes and/or expectancy of favourable responses from other teammates or the collective team (Lee and See, 2004). In essence it is about whether users are adopting an *appropriat*e level of trust in the automated agent. As McDermott and Brink (2019, p. 362) observe, 'If users understand the strengths and weaknesses of the automation, they will rely on it appropriately and use it more effectively'. Calibrated trust tends to include the following four dimensions:

- Belief: Does the user believe the machine will achieve what it is designed to do?
- Understanding: Does the user understand system errors and limitations?
- Intent: Does the user intend to use the automated system?
- Reliance: Does the user actually use the automated system?

Trust calibration has been described as a sensemaking process wherein 'operators must make sense of the automation's functioning to determine when they can depend on the automation and under what circumstances the automation is unreliable' (Fallon et al. 2010, p. 390). There is evidence to suggest that too little trust in robotic systems may lead operators to disengage the autonomous capabilities of a system, which may result in damage to both the robot and its environment. Conversely, too much trust may result in dire consequences to the extent that the operator may find that it puts themselves or others in harm's way (Galliott, 2018, p. 124). Moreover, trust calibration must also be dynamic. There are calibration points, including (but not limited to) environmental conditions, system conditions, and mission or task phase, that may degrade or improve the performance of the automated system, and where trust should be either reduced or heightened. For example, the ability for an autonomous aircraft to effectively follow terrain in the mountains deteriorates and the aircraft needs to fly at a higher altitude (McDermott and Brink, 2019, p. 364). Operators need to be informed and practiced in identifying when they should rely on automation, when additional oversight is required, and when performance is not satisfactory.

For calibrated trust to be sufficiently supported, there must be ready access to, and provision of data on the source of diagnostic information and signal reliability (McDermott et al., 2017). As articulated by the UK MOD (2018, p. 54), 'The more capable our AI systems are, the greater their ability to conduct local processing and respond to more abstract, higher level commands. The more we trust the AI, the lower the level of digital connectivity we will demand to maintain system control'. But this also necessitates the development of the requisite standards, assurance, and certification processes coupled with mechanisms to ensure demonstrable and meaningful human accountability and responsibility.

Legal and Ethical Considerations

The issue of meaningful human involvement and accountability is of much broader significance than merely supporting the effectiveness of the HMT. Such questions draw in complex, and sometimes confronting, legal and ethical considerations. The presence of smart machines in human affairs, and especially in military domains where there may be targeting or fires, has meant engagement with multifaceted and ethically confronting ideas about what constitutes meaningful human control. Indeed, the constancy of these questions in the scholarship attests to their extraordinary complexity and importance, as well as their propensity to confound.

For example, in 1960, Norbert Wiener, one of the seminal thinkers of early cybernetic systems, keenly articulated the risks he saw in having automated machines execute functions in human affairs. He suggested that the result was likely to be that humans would not have the capacity to understand such machines' logic or have adequate control over their actions. Ultimately, this coupling of 'two agencies essentially foreign to each other', which operate against very different measures, may well result in disaster (Schwarz, 2021, p. 53; see also Wiener, 1960).

Wiener was concerned with the problems of his own time, but his concerns certainly resonate in the current technological environment. We continue to be confronted by the idea that 'when human atoms are knit into an organisation in which they are used, not in their full rights as responsibly human beings, but as cogs and levers and rods, it matters little that their raw material is flesh and blood' (Wiener, 1954, p. 181). As Elke Schwarz (2021, p. 69) has highlighted, against an increasingly technological military landscape 'the capacity to take responsibility and feel the weight of a morally complex decision becomes more difficult as these decisions are distributed and mediated through technological interfaces, nodes, and various system components'. Human reliance on machines in complex military environments should prompt solemn questions as to how augmented levels of machine autonomy might affect human control over that technology.

There have been manifold efforts to delineate the fundamentals of meaningful human control, most tending to include the following:

- An operator, or human 'on the loop', should be in a position to make a conscious and informed decision about the appropriate use of the system.
- He or she should have adequate information about the target, the system itself, and the context within which it will be used.
- The system should be predictable, transparent, and reliable.
- The system has been adequately tested to operate as intended.

- Operators have been trained.
- There is the potential for timely human action and intervention, such that control over the system is facilitated to the best possible degree.[3]

However, as contexts change, and as systems advance, the extent to which these elements are enduring or applicable is uncertain. Even in the contemporary context, these fundamentals must contend with a suite of evolving technical, moral and legal complexities that make the space dynamic. Because of the intangible and moving contours—complex and ever-progressing technological features and the underlying logics of RAS-AI—some scholars cast doubt over whether we can maintain meaningful human control. Schwarz (2021, p. 55) argues that ultimately 'the human operator embedded in complex digital ecologies encounters limits that hamper moral agency and decision-making capacities considerably and in turn complicate aspirations of human control'.

In the HMT paradigm, machines participate in how actions and processes are shaped, which allows for the development of a 'co-constitutive relation' between technological systems and human frames for decisionmaking (Schwarz, 2021, p. 56). David Gunkel (2018, p. 54) argues, 'We now have things that are deliberately designed to exceed our control and our ability to respond or answer for them'. This gulf is contemporarily bridged by adopting a strong user-focus in technological configuration. Dialogue protocols and interface design are incorporated to mediate high-level cognitive processes, communication, decisionmaking and competence, and situational awareness (Schwarz, 2021, p. 56).

Given the cognitive and technical complexity of the contemporary legal and ethical HMT landscape, and its evolving nature, legal obligations need to be understood to ensure against impracticability or risk exposing legal avenues that are vulnerable to misuse (UK MOD, 2018), especially as they try to pre-empt an unknowable technological future. RAS-AI must perform such that they instil a sense of trust, safety and reliability in those who use them. While this includes transparency across interface design and strong user-focus, those who are involved in the development of RAS-AI in these contexts should also ensure the systems clearly comply with international humanitarian law. There must be proactive attention to the conceptual and ethical complexities of the HMT paradigm, starting from the moment of design.

As legislation is developed that may foster societal adoption of technology, lines of accountability and responsibility will become increasingly significant, and potentially difficult to navigate. Heightened awareness of the evolving moral-technical landscape that informs and shapes the pursuit of meaningful human control will be critical. Legal and ethical questions will likely grow in complexity, particularly as collaborations between humans and machines become closer and even extend into the realms of hybridisation in combat situations. It may become harder to discern what meaningful human control, or appropriate human judgement, looks like. In the traditional military kill-chain, it is possible to isolate each stage so as to determine legal culpability; it may become much more difficult to assign culpability when there is a kill-web, and when the separation point between the human and machine is less obvious (Binnendijk, Marler and Bartels, 2020, p. 26).

[3] On these points see, for example, Roff and Moyes, 2016; and Horowitz and Scharre, 2015.

The presence and uses of RAS-AI will continue to expand in society, and as there are increased efforts and successes in making such technologies safe and reliable, we are likely to see greater public appetite for them, and extended uses within the military space. Indeed, it may be that we start to see automated weapon systems designed such that they have better capacity to comply with the principles of proportionality and distinction required under the Laws of Armed Conflict. Under these circumstances, it will become considerably easier for states to justify their use (UK MOD, 2018, p. 50). There must be commensurate attentions to the ethical and legal dimensions as this landscape continues to evolve.

3. Implications for Future Workforce

The primary objective of this chapter is to explore the implications of HMT for the RAN's future workforce. There is relative consensus within the existing literature that RAS-AI, and specifically HMT, is likely to significantly change the nature of many military functions as well as the requisite skills. These effects will be felt across different parts of Defence that interact with RAS-AI and may result in a growing need to adapt workforce management, recruitment and retention efforts.

Technical Skills Requirements

For the defence sector, core existing capability areas may well continue to form the basis of the future workforce, but they will do so against a changing technological landscape and an increasingly challenging operational context. The continuing relevance of existing capability areas will require continued upkeep of traditional skills such as mechanics and electronics. However, throughout the defence equipment life cycle, rapid technological advancements, potentially disruptive innovation, and the insertion of emerging technologies into areas including cyber, RAS-AI, quantum technologies, and hypersonics will demand transformation across capability planning as well as training and skill development (Galai et al., 2019, p. 22).

As military equipment continues to become a complex system of systems, supported and underpinned by a variety of enablers and non-combat equipment, the defence sector will need to develop new ways of thinking about defence industrial skills because they will not, as has been the convention in the past, be restricted to a specific platform. Instead, they must be understood as they relate to, and are affected by, broader interoperability demands, dual-use technologies and equipment, and increasing digitalisation and automation. In some parts of the Defence force, understanding, using and supervising automated technologies may require highly-skilled AI-literate personnel with vocational or higher-education training spanning topics such as sensor fusion, electronics, cybersecurity, big data, programming and coding, modelling, and simulation (Galai et al., 2019). End users should, for example, be able to interrogate and de-risk RAS systems' design—and potentially define the desired functionality of the RAS depending on the objectives or the mission—requiring a higher degree of technical skill and knowledge (UK MOD, 2018).

In relation to the 'back of house' workforce—personnel involved in R&D, support and enabling functions, and strategic decisionmaking roles—RAS-AI technologies may also require personnel with a strong background in science, technology, engineering and mathematics (STEM) to develop effective HMT solutions. Foundational STEM skills are considered key in the development of RAS-AI. Solid STEM skills enable military professionals to grasp RAS-AI functionalities and to train in the higher RAS-AI skills, while also being able to work in adjacent roles in and around AI that may be useful for Defence, such as data preparation, curation and protection (Hall and Pesenti, 2017). For example, foundational mathematics skills are important when designing machine learning systems involving statistics and data science. Increasing quantities and complexities of data, as well as more opportunities for Defence to utilise remote and autonomous systems, will require a range of specialist skills for their development and adoption. Specialist personnel may be required to enable the effective organisation, classification and analysis of complex data sets. The need for specialist skills may increase as RAS-AI mature sufficiently to take over more repetitive or dull tasks and personnel shift into new roles requiring greater aptitude, cognitive ability, and expertise.

The UK MOD has placed significant emphasis on STEM recruitment. In 2019, it announced the creation of the STEM Graduate Inflow Scheme, which is expected to cover three years of university tuition fees, as well as an annual bursary (Lancaster, 2019). In addition, the Defence STEM Engagement connects with national initiatives from the Department for Education and the Department for Business, Energy, and Industrial Strategy to promote the uptake of STEM subjects in schools (UK MOD, 2019).

A recent report published by the NSCAI (2021, p. 10) has similarly argued that the talent deficit in the U.S. Department of Defense 'represents the greatest impediment to being AI-ready by 2025'. It argues that the United States needs to 'scale up digital talent in government. National security agencies need more digital experts now or they will remain unprepared to buy, build, and use AI and associated technologies' (NSCAI, 2021, p. 10). The report recommends the development of

> new talent pipelines, including a U.S. Digital Service Academy to train current and future employees. It needs a civilian National Digital Reserve Corps to recruit people with the right skills—including industry experts, academics, and recent college graduates. And it needs a Digital Corps, modeled on the Army Medical Corps, to organize technologists already serving in government. (NSCAI, 2021, p. 10)

The NSAIC (2021, p. 174) recommends that to become leaders in the RAS-AI context, the United States must attend to the following four key talent areas:

- *Researchers*, who would focus on the R&D of technologies that would advance semi- and fully autonomous systems. They would be across the most up-do-date RAS-AI research and help to progress ideas, be involved in their inception and development, and drive research through to testing of prototypes for major projects.
- *Implementers*, who would look after the cleaning of data, feature extraction and selection, and data analysis. They would also conduct model training and tuning, foster collaborations with domain experts and end users, and locate advantageous opportunities.
- *End users*, who would have their usual activities augmented by RAS-AI. RAS-AI–enabled systems would largely behave and look like existing software, requiring some system-specific training, but otherwise not require specific expertise in AI (with the exception of some data-handling roles).
- *Informed consumers*, who would make improved consumer choices when purchasing technology and have awareness of how their actions affect market dynamics.

There will also be a need for government and other officials to oversee systems, as well as people to make purchasing decisions.

While the RAN need not dominate this space, attending to these archetypes may also be valuable for successful integration of RAS-AI into the workforce. Attracting RAS-AI experts and retaining specialist skills requires proactively embracing innovation and new ways of working where needed to enhance organisational efficiency and capability (Ministère des Armées, 2019; British Army, 2022).

A flexible risk management approach will promote more agile delivery of Defence's capabilities using RAS-AI. Moreover, this flexibility will allow the RAN to more effectively implement the corresponding 'fail fast/fail early' approach that is integral to innovation. In this context, Defence will need to adapt its incentives for personnel to make nuanced risk assessments rather than focusing on compliance with strict guidelines that may be overly risk-averse and prevent experimentation and innovation (U.S. Department of Defense, 2018). This may require revision of traditional governance principles and processes, making them flexible and tailored to the specifics of a program, continual skills development, a work environment that gives talented individuals the space to innovate, and exchange of knowledge through collaboration with industry and academia (Freeman et al., 2015).

Attributes and Non-Technical Skills

A nuanced and comprehensive understanding of the dynamics of HMIs, now and into the future, is key to understanding the kinds of skills and training that will best take advantage of those relationships. HMT requires a new way of working physically and, perhaps more significantly, cognitively, and raises challenges that will require continual attention in order to deliver the best mix of humans and machines. In complementarity with attentions to the technical skills required in mastery of RAS-AI, Industry 4.0, a shorthand term for the movement towards the digitisation, mechanisation and automation of the manufacturing environment, will also require investment in generalist skills, as well as what has come to be termed 'soft skills' among existing and future service personnel.[1]

Generalist and multidisciplinary skills that will facilitate the development and exploitation of RAS-AI may include agile contracting and contract management skills, defence-related program support and management, cost modelling, and data protection law (reporting, accountability and liability). Indeed, at the contracting stage, considerations will need to be made about who owns data, models, and so on. The process of RAS-AI development and exploitation relies on a range of services provided by the Defence 'back of house' workforce. Personnel in these roles may support the integration and use of RAS-AI indirectly (e.g., by supporting acquisition processes) but also serve as end users (e.g., if AI models are used by support personnel in repair and mechanical roles to monitor potential system vulnerabilities).

There may also be a need for professionals to design, manufacture, produce, and constantly update different systems and subsystems across domains, as well as to be involved at other critical stages of an equipment life, such as at the management and

[1] For a more detailed definitional discussion of Industry 4.0, see Lasi et al., 2014.

contracting levels. This will also require the monitoring that is necessary for RAS-AI systems to maintain operationality. Because of the rapidity of information exchange, and the various associated communication management and production systems, military organisations will need to develop a commensurate culture of adaptivity and flexibility around their processes, which are acknowledged to traditionally be quite rigid.

The RAN will need to seek and foster attributes in their employees that allow for cognitive adeptness, agile learning, innovative and creative thinking, high-level interpersonal and communication skills, and the capacity for managing complex and evolving problems (Galai et al., 2019). This will need to happen across all levels. From the end user perspective, dealing with large flows of data may require skills such as organisation (e.g., time management, data classification and analysis); innovative thinking to solve complex problems; and advanced interpersonal, social communication and networking skills in order to share important technical insights across an increasingly dispersed chain of command (Burmaoglu and Saritas, 2017). At the strategic level, decisionmakers may be required to think across disciplines and consider a wide range of data sources in the context of policy and operational planning (Winkler et al., 2019).

Moreover, the ethical dimensions of RAS-AI may also require new specialist roles for personnel with expertise across human needs and behaviours, such as cognitive psychologists, digital anthropologists, HMI professionals, and user-experience researchers and designers to drive the design of RAS-AI systems that benefit end users through increased useability, transparency and interpretability (Smith, 2020). Indeed, there may be significant benefit in fostering diversity in teams with regard to education, thought process, overall experience, as well as race and culture. Diverse teams are better able to support the speculative work necessary to make systems that are robust.

Access to External Skills

The RAN's *RAS-AI Strategy 2040* states that 'identifying, developing, and sustaining the future workforce' will not only necessitate the acquisition of new personnel in highly sought-after disciplines but will also require 'accessing a range of skills outside of Navy' (RAN, 2020, p. 19). The RAN has recognised that in order to have 'suitably qualified and experienced personnel to develop and employ RAS-AI', it will need to commit to 'workforce transformation, training system adaptation and a framework for human machine teaming' (RAN, 2020, p. 19). Certainly, as there is a notable shift from the defence platform to an ecosystem comprising systems, technologies, and professional services, defence skills are no longer developed in a vacuum but are becoming increasingly connected to those that may be more comfortably located in the civil and dual-use sectors (Galai et al., 2019, p. 5). As Anton Shufutinsky et al. (2020, p. 73) observed in a recent paper on the implications of the changing technology landscape on talent acquisition, we appear to be at 'an inflection point in which industry is continuously and rapidly transformed by technological advancement, and not only are these technologies digital, smart, and capable, but it appears that the largest technological gains lie ahead and we must be prepared to tackle the changes that come with these advancements'.

In the development of RAS-AI capability, this will be keenly felt in the skills, aptitudes, and cognitive demands it requires of RAN personnel, many of whom may not be most immediately and comfortably native to the RAN's training and skill base and may incur non-traditional cognitive demands. Under these circumstances, a more proactively productive relationship with industry and academia and greater acceptance of the immediate and integral role that those sectors have in the RAN's RAS-AI enterprise is required. It is those sectors that will inform technology and research trajectories and will be called on to complement and/or supplement the RAN's skill base in those RAS-AI–related areas in which the organisation cannot readily develop skills or knowledge. This means that attentions must project beyond those skills that will deliver current and planned defence equipment programs, and a RAS-AI–literate/enculturated workforce. There must be development of channels and personnel to readily access and take advantage of emerging technologies, and foster innovation and creativity to remain continually competitive (Galai et al., 2019, p. viii). On this point, Paul Scharre, author of *Robotics on the Battlefield*, has argued, 'The winner of this revolution will not be who develops these technologies first, or even who has the best technologies, but who figures out how to best use them' (Scharre, 2014a, p. 9).

Even though there are characteristics that are unique to the defence sector, the RAN may well find useful insights on workforce development strategies from the commercial sector, particularly the high-technology and information technology (IT) industries, as these industries share some technological similarities and skills requirements and challenges (Galai et al., 2019, p. 67). Understanding, and perhaps even adopting, some of the practices of the civil sector makes sense because RAS-AI technologies require transformed levels and types of engagement with that sector. Many players in these industries have a competitive incentive to drive innovation; they accept and often foster overlapping areas of activity with the defence industry and other civil sector players; and they tend to demonstrate similar skills competencies to those required in the defence industry (Galai et al., 2019, p. 109).

Moreover, these industries tend to face some similar challenges (e.g., in skills retention), but adopt different mitigation strategies. In this specific context, IT and high-technology industries tend to develop specialised, or targeted, retention strategies for high-performing talent (Galai et al., 2019, p. 67). The RAN would do well to understand how these industries recruit and retain skills native to this field, including 'engineering, software development, software engineering and cyber skills' as it will be competing with these industries for superior talent (Galai et al., 2019, p. 109).

Moreover, many essential skills and competencies needed for RAS-AI employment do not currently exist in traditional defence forces. As is acknowledged in a recent report published by the European Commission, 'There is an inherent unpredictability in terms of the future demand for defence capabilities as well as the supply of skills for the defence industry' (Galai et al., 2019, p. 22). Organisations outside of Defence may include repositories of specialist and non-specialist skills relevant to RAS-AI integration. In the case of industry, this includes, for example, strong skills in portfolio management, risk management, and project delivery that could be highly beneficial to Defence. The UK MOD has been particularly active in developing closer links with industry and academia through a so-called Whole of Force approach that aims to ensure continuous capabilities delivery relying on industry and civilian support, including through its Defence Enterprise approach. Table 3.1 provides an overview of selected examples of collaboration between the UK defence, academia, industry and wider government sectors on the human component of RAS-AI.

While only a snapshot of the landscape, these examples demonstrate what purposeful collaboration and engagement can look like between these entities and sectors, where there may be points of convergent interest or advantage, and some of the activities that might be conducive to continuous capabilities delivery.

Table 3.1. Examples of UK 'Whole of Force' Collaboration in the Context of RAS-AI

Name	Participants	Description
MOD Autonomy Research Programme (2020–present)	Defence Science and Technology Laboratory (Dstl), MOD, industry, academia	Analyses multiple aspects of RAS. Focuses on areas such as the 'optimisation of human autonomy teaming' as well as research into 'legal certification and regulatory requirements' and 'identifying ways to remove the human from harm'. Also includes work on the ethics of HMT, including protection and disruption of such teams.
Future Workforce and Human Performance Programme (2018–present)	Dstl, MOD, industry, academia	Program initiated to develop approaches to secure MOD's human capability in a time of significant technological change. Includes research into areas such as integration of humans into systems, optimising human performance through innovative tools and novel concepts, and approaches to training.
Project Nelson (2018–present)	Royal Navy, industry	Aims to enable faster development and deployment of AI across the Royal Navy through a user-centric approach. Incorporates the understanding of service personnel about processes within the Royal Navy to evaluate possible technologies. Works with the user community to build products and services, including man-machine interfaces that best suit users' needs.
Project Wilton Maritime Autonomous System Team (2020–present)	Royal Navy, industry	Part of the 1st Mine Countermeasures Squadron. The team received formal ownership of their autonomous mine countermeasures vessel, RNMB *Harrier*, in August 2020. The introduction of RNMB *Harrier* included a week-long familiarisation and training package delivered by the prime contractor, Atlas Elektronik UK.
Semi-autonomous reconnaissance vehicles for the Army (2019–2020)	Defence and Security Accelerator, Army, industry	Asked for innovative solutions to the challenges posed by HMT in the utilisation of uncrewed systems. Demonstrations occurred at the Army warfighting experiment in March 2020.
Serapis framework (2019–present)	Dstl, industry	Scope includes contracts looking at RAS as well as AI. Lot 3 in the Serapis framework, 'Decide', explicitly calls for proposals addressing command, control, communications, computers, intelligence, surveillance, and reconnaissance systems that might improve decisionmaking among human and autonomous systems. Lot 6, 'Understand', covers proposals on the application of autonomy and AI and machine learning to defence and security problems as well as man-machine interface development.

While there are opportunities for a ‘Whole of Defence’ approach to strengthen the Defence skills base for RAS-AI integration, there are also various barriers that can constrain collaboration and limit opportunities to leverage the approach for new personnel policies (Gearson et al., 2020). These can include cultural misconceptions because the attachment to military culture and values may disincentivise collaboration with external parties. Outsourcing to the private sector can be perceived as a factor that could undermine in-house military capability and skills. Tensions can also arise due to misperceptions concerning employment conditions (e.g., that contractors are better paid and less exposed to operational risks than regular military personnel) and differences in employment and professional development incentives (e.g., the academic drive for publishing) (Galai et al., 2019).

Recruitment and Skills Retention

RAS-AI integration is likely to have various impacts on the Defence workforce through changing personnel roles and HMT solutions, requirements for attracting and retaining new skills, and a changing structure and culture. In addressing these impacts, Defence must consider a range of requirements for how it manages its workforce, from recruitment and retention to active upskilling through training and education. The rapid pace of technology advances in RAS-AI requires, first and foremost, a comprehensive understanding of the Defence RAS-AI workforce as well as the trends that may change requirements for its management. This understanding should include establishing what skills and competences already exist in Defence and how these may need to be updated or complemented with new ones arising out of development and adoption of new RAS-AI technologies. Where Defence identifies the need for new roles or positions, it should also understand how these should be staffed (e.g., by civilian or military staff, through secondments or outsourcing options) and establish minimum requirements (e.g., in relation to professional experience, security certifications, and standards) (U.S. Department of Defense, 2018; Ministère des Armées, 2019).

In this space, it is important to recognise that major shifts in the workforce may engender concerns about the displacement of existing personnel (Galliott, 2018, p. 124). A key theme in this problem space is likely to be around the issue of identity. For example, during the transition to steam-powered ships in navies, it was not just that a fundamental technology was being superseded but also that the intrinsic nature of what it was to be a sailor changed. Sailors no longer climbed the mast and worked the rigging; they were engineers in the engine room (Scharre, 2014b). This threat arguably appears much greater in the context of RAS-AI. From the perspective of the user, RAS-AI technologies will fundamentally alter the human-machine relationship (UK MOD, 2018).

Until recently, new technologies were principally deployed in place of, or to augment, the physical rather than cognitive (Packer, 2019). RAS-AI may threaten both, which is a much more unnerving proposition. This perceived threat is compounded by the humans’ need to protect their identity. Recent psychological studies have found that people readily ignore data that may challenge their core identity, in military contexts and more broadly. This is simply a function of the human condition and is part of a process of ‘identity-protective cognition’, a self-defence mechanism that guides the human away from ideas and beliefs that might result in alienation from the supports they have adopted for countless dimensions of their everyday existence. This kind of cognitive process may make incorporation of new technologies challenging for the military, especially if they are seen to undermine personnel's specific occupational skill (Scharre, 2014b).

A perennial challenge for Defence will be in maintaining the relevant skills base for RAS-AI. This is an acknowledged challenge across militaries looking to future-proof the workforce. The European Commission report cited earlier highlighted that there is a clear 'overlap between current and near-term skills mismatches indicating a potential chronic difficulty in sourcing and retaining skills' (Galai et al., 2019, p. viii). This comes from the difficulties Defence has in competing in the RAS-AI recruiting market and remaining innovative to secure the right subject-matter expertise. The recruiting challenge extends, similarly, to the ability to 'nurture sufficient in-house knowledge and understanding to generate intelligent customer capabilities'. As is noted by the UK MOD (2018, p. 7), 'This will be essential to understand where AI should be exploited; translate operational requirements and constraints to non-military AI experts; and support the generation of effective military industrial teams for innovation, problem solving and force development'.

Military recruitment processes may have to be adapted to the increasing need for skilled and high-aptitude personnel. This need may be met by the adaptation of entry requirements and selection methodologies. Processes used to select candidates for Defence roles may need to be reassessed in response to the changing workforce and skills requirements driven by RAS-AI integration. Defence may therefore benefit from reassessing the relative weight of various mental and physical selection standards to increase the pool of available specialist candidates and go beyond its traditional recruitment pool to attract a more diverse workforce. Given the increased emphasis on high aptitude across the Defence workforce, tests and standards emphasising cognitive performance may be emphasised over traditional physical selection standards, for example (Winkler et al., 2019).

There may also be a need to adapt career pathways. Defence organisations frequently rely on outdated 'base-fed workforce models', meaning that the armed forces recruit at entry level and develop service personnel's skills over the course of their careers (Gearson et al., 2020). However, the rapid pace of technological development and current skills shortfalls may necessitate bringing in talents from civilian organisations and the private sector through shorter-term contracts and innovative recruitment approaches that may not correspond with such traditional career pathways. For example, in the United States, lateral-entry programs, which allow industry or private sector representatives to join the military in mid-level positions, as well as a reconsideration of the military up-or-out promotion system, represent some options for Defence to adapt existing structure to rapidly recruit personnel with in-demand skills (Winkler et al., 2019). Similarly, in the U.S. context, relying on the reserves has been a means by which to mobilise niche skills when needed.

Moreover, attracting personnel with in-demand skills from outside of Defence may be achieved by developing corresponding education programs and drawing in candidates through educational opportunities. In relation to RAS-AI specifically, this could include increased funding for PhD placements in big data, AI and robotics, and machine learning, as well as improving educational programs in relevant fields across universities from which Defence can subsequently recruit. While domestic AI talents will address long-term skills shortfalls, Defence can also attract overseas (to the extent possible in light of any nationality-based restrictions) to address potential long-term skills shortages (e.g., by increasing the number of annual visas delivered to exceptional talents and through academic partnerships) (Select Committee on Artificial Intelligence, 2018).

Training and Organisational Learning

RAS-AI integration may require Defence to develop new frameworks and approaches for continuous upskilling of the workforce through training and education. While attracting specialists through financial incentives may offer short-term solutions to the fast-changing pace of skills requirements, organisations in Defence or the commercial sector are unlikely to meet skills demand by relying only on hiring new employees (Galai et al., 2019). Notwithstanding the financial and material cost of reskilling employees, recruitment of specialists and enhancement of the Defence skills base to integrate RAS-AI may therefore be more easily achieved internally than externally. This is due to the significant competition across various sectors for RAS-AI specialists and the challenges Defence faces in competing with private sector employers for new talent. Skills training systems should be adapted to changing demand and new technological developments. As such, a continuous assessment of potential skills gaps and opportunities for strengthening the Defence skills base through new training opportunities is required.

Military educational institutions will have to equip junior ranks, non-commissioned officers, and officers with the necessary technical knowledge through courses in, for example, technological literacy and AI ethics. As RAS-AI becomes increasingly mature and integrated across Defence, lifelong in-house training and education opportunities for military specialisation will be required. Reading lists, residential programs, AI-driven intelligent tutoring systems, and academic partnerships and conferences may be useful to help military professionals develop the necessary technical and soft skills (Ryan, 2018).

Of key importance in upskilling in HMT will be innovative and interactive forms of learning such as live and synthetic training with RAS-AI. To develop comprehensive understanding of function allocation and optimised roles, especially as RAS-AI continues to advance, 'an aggressive strategy of iterative experimentation, prototyping, concept and technology development, and organizational refinement' should be adopted (UK MOD (2018, p. 47). As is acknowledged by the UK MOD (2018, p. 47),

> High quality live and synthetic collective training and experimentation with AI systems will be essential to optimise our ability to create effective human-machine teams. Training and experimentation with real users will be vital for operators to understand the strengths, weaknesses and critical limitations of such AI systems while also providing vital data to improve AI responses, including about the human behaviours in the team.

Aspects of live, virtual and constructive (LVC) training could be especially helpful in supporting workforce development with RAS-AI. Among many other uses, LVC training may be used to conduct scenario-based assessments and evaluations of HMT; understand and practise integration of RAS-AI into expeditionary military operations; and build competence, expertise and capability through rehearsal and testing in virtual environments. As with human partners, there must be a process of training and growing alongside autonomous and robotic partners to ensure that interfacing is tailored both to individuals within the team and the broader collective.

For this training to be most effective, it must be dynamic, diverse and realistic, and create a surrogate battle landscape that allows for experimentation and delivers and builds trust and confidence (UK MOD, 2018, p. 47). To these ends, there must be fidelity across the simulation, physical, functional and psychological dimensions. *Simulation fidelity* refers to the extent to which there is a true replication of tasks, equipment and environments. *Physical fidelity* refers to the accuracy of the features of the tools and systems (e.g., size and weight, appearance of controls, and cues or feedback). *Functional fidelity* refers to the extent to which the way a simulator responds to user commands is realistic or replicates actual system functions. *Psychological fidelity* refers to how accurately the simulator can prompt cognitive, emotional and behavioural responses that are appropriate to performance or action when operating equipment in a specific context or environment (Straus et al., 2019, pp. xii–xiii).[2]

Incorporation of these elements will ensure high-quality LVC that will have utility in training and developing HMT as well as in establishing and identifying future HMT requirements for the RAN because the environment, contexts and threat profile will likely evolve across the breadth of RAN's lines of development (UK MOD, 2018, p. 47).

Moreover, once autonomous systems are seen increasingly as assistants, collaborators or partners rather than tools, and collective trust is built and fostered, their presence in teams will grow and become expected, and their independent use will increase. This is likely to result in deeper and more high-volume attachments to, and engagements with, these kinds of technologies (de Visser, Pak and Shaw, 2018, p. 1410). As the use of RAS-AI grows, the RAN may also benefit from greater investment in in-house training programs or invest in training initiatives delivered by external providers. In this context, Defence can learn from initiatives implemented across the private sector, including longer-term training in the form of technical research placements coupled with specialist mentoring (Galai et al., 2019).

In implementing these approaches, Defence can foster a continuous learning culture as an enabler for retention as well as upskilling. It is therefore important to note that training and learning opportunities should be formalised and provided continuously to incentivise their uptake. As such, there must also be an understanding of how to build organisational capacity for innovation using bottom-up practices that promote a culture of innovation and ensure buy-in into new ways of working. Notably, reflecting on best practices from civilian employers, peer-to-peer recognition may be an effective remedy for defence organisations to drive engagement and productivity (Suchland and Brown, 2021). Defence organisations can also adopt new models for monitoring ongoing evaluation of organisational culture and its conduciveness to innovation, such as through the adoption of cultural maturity indices to assess change in behaviour and outcomes (National Audit Office, 2012).

---

[2] For further RAND analysis of the workforce implications of LVC training, see, for example, Marler et al., 2021.

4. Final Observations

Summary

This report has reviewed relevant open-source academic and grey literature, with a focus on identifying possible lessons for the RAN. The report has considered the challenges and impacts of RAS-AI integration, with a particular focus on the implications of HMT for the Defence workforce, broadly defined. The report highlights that HMT requires a fundamentally different, and decidedly novel, way of working. In contrast with approaches to HMI where the human is learning to leverage the machine, the goal in HMT is to optimise the interaction and leverage the strengths of both sides: human *and* machine. This is especially urgent as there is a global and adversarial race among militaries to unlock this advantage, and new technology capabilities continue to accelerate.

Limitations and Future Work

This review of the literature was geared toward how HMT might affect the future workforce of the RAN, and therefore some literatures were excluded to the extent that they did not speak to the specific concerns of the RAN. As such, this review was purposefully selective, prioritising the integration of a diversity of sources from across the field rather than pursuing an exhaustive catalogue. The review, largely limited to secondary analyses and grey material, critically integrated received wisdom on the necessary considerations concerning HMT and its implications for workforce development, noting many critical challenges as well as suggested courses of action. The review also resulted in findings that may be applicable to other emerging technologies in the defence sector.

However, we were also not able to pursue, in any detail, the impacts of integrating particular techniques of HMT on the workforce. The scope for this project was largely limited to HMT assuming the level of partnership or collaboration. But there is important and growing scholarship on an extended form of HMT where there is a blending or hybridisation of human and machine intelligence. These techniques involve technologies that allow the human brain to communicate directly with machines. The implications of the use of these techniques for the future workforce, in terms of the shape of the force, trust dynamics between multiagent HMTs, requisite skill sets and attributes, and ethical and legal issues (among other considerations) may be strikingly different to those outlined in this report. Because of the dynamic sociotechnical nature of HMT, it may well be difficult to forecast the extent to which the integration, and effects, of broad-based use of teams comprising humans and machines will change the future workforce.

Primary Findings and Observations

Our key findings and observations are the following:

- Normalisation of HMT in the RAN will require flexible management of the Defence workforce and continuous adaptation of existing structures and concepts.
- HMT necessitates a shift in cognition as much as training and perception.
- The goal in HMT is to optimise the interaction and leverage the strengths of both human and machine.
- HMT encompasses broad and complex issues and defies categorisation into an ineluctable laundry list of principles, activities, and resources.
- Understanding and considering the spectrum of HMI is integral.
- New technological developments should be aligned with actual problems/needs
- The skills/attributes required for HMT may not traditionally be prized by Defence. Organisational learning must be embraced to deliver this challenging capability.
- HMT should be leveraged, not only as an effective military fighting capability but also as a concurrent training capability.
- Success in HMT requires significant uplift across the breadth of the workforce.
- HMT requires identifying the strengths and weaknesses of human and machines and capitalising on advantages so that they become greater than the sum of their parts.
- HMT focus should be on problems to be solved, the appropriate ratio and mode of HMI and finding the sweet spot in terms of effort payoff.
- RAS-AI must perform such that they instil a sense of trust, safety and reliability in those who use them. This includes transparency across interface design, strong user focus, and clear compliance with international humanitarian law.
- There must be proactive attention to the conceptual and ethical complexities of the HMT paradigm, starting from the moment of design.

Broader and Pervasive Issues

There are a number of pervasive issues that must be recognised and considered in the HMT context. The complexity of the topic of HMT, and the necessary investments and inputs it entails—physical, material, organisational and, most significantly, intellectual/cognitive—requires the RAN to take a comprehensive and self-critical approach to its implementation. HMT is not merely a matter of integrating the most recent technological developments into the force in a ready manner; it requires deep engagement with some fundamental questions about the human condition. HMT demands in-depth engagement with the multifaceted issues of trust, but it also draws in questions on what it means to be human and non-human, and the relationship between those entities. Moreover, 'the relationship between autonomy and the human is constantly changing, reflecting the adaptivity of autonomy itself and human perceptions' (de Visser, Pak and Shaw, 2018, p. 1410), especially as the cognitive abilities of machines are progressively advanced. The mentality and approach required for effective HMT is completely unlike that required to deploy machines as tools. HMT does not merely mean downsizing human involvement (Johnson et al., 2014, p. 79).

HMT encompasses broad and complex issues and, as such, largely defies categorisation into an ineluctable laundry list of principles, activities and resources because it is a fundamentally different, and decidedly novel, way of working. Consequently, the implications of adopting HMT for the future workforce are substantial and wide-ranging. RAS-AI will require flexible management of the Defence workforce, with workforce management increasingly denoting the management of complex personnel networks; it will also require a continuous adaptation of existing structures and concepts. It is only in understanding the deep complexities of HMT that an appropriate approach to constructing the future workforce can be formulated, not just in terms of training and skills (technical and non-technical), but also the attributes and qualities that might be required of personnel and the organisational learning that must be embraced to deliver this new and challenging capability. The changes required across the force should not be underestimated. Success in HMT requires significant uplift across the breadth of the workforce. HMT will be most effective when recognising the limits of humans to interact with and control multiple entities, with or without automated assistance.

The demands of HMT, particularly in terms of creating interactions, will have a broad-based benefit in that it will require designers to 'consider human well-being as central to the operational success of the overall human-machine system that is being designed' (Cesafsky, Stayton and Cefkin, 2019, p. 65). In this sense, designing for HMT goes much deeper than simply the user interface, and means thinking beyond maximising autonomous capabilities to the more important goal of improving mission performance of the work system (Johnson et al., 2014, p. 75). The true value and power in developing an effective approach to HMT is that it will encourage virtues to emerge, proliferate and multiply. Perhaps most importantly, it will also grow a workforce that knows how to work smarter (Johnson et al., 2014, p. 79). HMT should be leveraged not only as an effective military fighting capability but also as a concurrent training capability. The opportunity is in purposefully extending the human cognitive capability such that the exploitation of smart technology is as comprehensive and creative as possible, and in building human capacities such that they are equipped to deal with the impending landscape of RAS-AI and the exponential growth of future opportunities (Lyons and Havig, 2014, p. 182).

References

Behymer, K. J., and J. M. Flach, ‘From Autonomous Systems to Sociotechnical Systems: Designing Effective Collaborations’, *She Ji: The Journal of Design, Economics, and Innovation*, Vol. 2, No. 2, 2016, pp. 105–114.

Binnendijk, Anika, Timothy Marler and Elizabeth M. Bartels, *Brain-Computer Interfaces: U.S. Military Applications and Implications, An Initial Assessment*, Santa Monica, Calif.: RAND Corporation, RR-2996-RC, 2020. As of 7 March 2022: https://www.rand.org/pubs/research_reports/RR2996.html

Brill, Christopher, M. L. Cummings, A. W. Evans, Peter A. Hancock, Joseph B. Lyons and Kevin Oden, ‘Navigating the Advent of Human-Machine Teaming’, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 62, No. 1, 2018, pp. 455–459.

British Army, *British Army Approach to Robotics and Autonomous Systems: Generating Human-Machine Teams*, 2022. As of 29 April 2022: https://www.army.mod.uk/media/15790/20220126_army-approach-to-ras_final.pdf

Burmaoglu, Serhat, and Ozcan Saritas, ‘Changing Characteristics of Warfare and the Future of Military R&D’, *Technological Forecasting and Social Change*, Vol. 116, 2017, pp. 151–161.

Card, S. K., A. Newell, and T. P. Moran, *The Psychology of Human-Computer Interaction*, Hillsdale, N.J.: Lawrence Erlbaum, 1983.

Cassidy, Mike, ‘Centaur Chess Brings Out the Best in Humans and Machines’, *The Bloomreach Blog*, 30 December 2014.

Cesafsky, Laura, Erik Stayton, and Melissa Cefkin, ‘Calibrating Agency: Human-Autonomy Teaming and the Future of Work amid Highly Automated Systems’, *Ethnographic Praxis in Industry Conference Proceedings (EPIC)*, 2019, pp. 65–82.

Crisp, R., and M. Hewstone, ‘Multiple Social Categorization’, *Advances in Experimental Social Psychology*, Vol. 39, 2007, pp. 163–254.

Department of Defence, *2020 Defence Strategic Update*, Commonwealth of Australia, 2020.

de Visser, Ewart, S. S. Monfort, K. Goodyear, L. Lu, M. O’Hara, M. R. Lee, R. Parasuraman and F. Krueger, ‘A Little Anthropomorphism Goes a Long Way’, *Human Factors*, Vol. 59, No. 1, 2017, pp. 116–133.

de Visser, Ewart, Richard Pak, and Tyler H. Shaw, ‘From “Automation” to “Autonomy”: The Importance of Trust Repair in Human–Machine Interaction’, *Ergonomics*, Vol. 61, No. 10, 2018, pp. 1409–1427.

Fallon, C. K., A. K. Murphy, L. Zimmerman and S. T. Mueller, ‘The Calibration of Trust in an Automated System: A Sensemaking Process’, in *2010 International Symposium on Collaborative Technologies and Systems*, 2010, pp. 390–395.

Fitts, P. M., ed., *Human Engineering for an Effective Air Navigation and Traffic Control System*, Washington, D.C.: National Research Council, 1951.

Freeman, Jon, Tess Hellgren, Michele Mastroeni, Giacomo Persi Paoli, Kate Cox and James Black, *Innovation Models: Enabling New Defence Solutions and Enhanced Benefits from Science and Technology*, Santa Monica, Calif.: RAND Corporation, RR-840-MOD, 2015. As of 7 March 2022: https://www.rand.org/pubs/research_reports/RR840.html

Galai, Katerina, Lucia Retter, Julia Muravska, Marta Kepe, Alice Lynch, Anna Knack, Jacopo Bellasio, Antonia Ward, Sofia Meranta and Davide Maistro, *Vision on Defence-Related Skills for Europe Today and Tomorrow*, Brussels: European Commission, 2019.

Galliott, Jai, ‘The Soldier’s Tolerance for Autonomous Systems’, *Paladyn: Journal of Behavioral Robotics*, Vol. 9, No. 1, 2018, pp. 124–136.

Gearson, John, Philip Berry, Joe Devanny and Nina Musgrave, *The Whole Force by Design: Optimising Defence to Meet Future Challenges*, London: Serco Institute, 2020.

Groom, V., and C. Nass, ‘Can robots be teammates?: Benchmarks in human–robot teams’, *Interaction Studies*, Vol. 8, No. 3, 2007, pp. 483–500.

Gunkel, David, ‘Other Things: AI, Robots and Society’, in Ziziz Papacharissi, ed., *A Networked Self and Human Augmentics, Artificial Intelligence, Sentience*, New York: Routledge, 2018, pp. 51–68.

Haimson, C., C. L. Paul, S. Joseph, R. Rohrer and B. Nebesh, ‘Do We Need “Teaming” to Team with a Machine?’ in *HCI International Conference Proceedings*, 2019, pp. 169–178.

Hall, Wendy, and Jerome Pesenti, *Growing the Artificial Intelligence Industry in the UK*, London: UK Government, 2017.

Haslam, N., ‘Dehumanization: An Integrative Review’, *Personality and Social Psychology Review*, Vol. 10, No. 3, 2006, pp. 252–264.

Horowitz, Michael C., and Paul Scharre, ‘Meaningful Human Control in Weapon Systems: A Primer’, working paper, Washington, D.C.: Center for a New American Security, March 2015.

Joe, Jeffrey, John Ohara, H. D. Medema and Johanna Oxstrand, ‘Identifying Requirements for Effective Human-Automation Teamwork’, paper presented at Probabilistic Safety Assessment and Management (PSAM) 12, Honolulu, Hawaii, June 2014.

Johnson, M., J. M. Bradshaw, R. R. Hoffman, P. J. Feltovich and D. D. Woods, ‘Seven Cardinal Virtues of Human-Machine Teamwork: Examples from the DARPA Robotic Challenge’, *IEEE Intelligent Systems*, Vol. 29, No. 6, 2014, pp. 74–80.

Kania, Elsa, ‘Minds at War: China’s Pursuit of Military Advantage Through Cognitive Science and Biotechnology’, *PRISM*, Vol. 8, No. 3, 2020, pp. 83–101.

Lancaster, Mark, ‘Recruiting STEM Graduates into Defence’, UK Parliament Hansard, Vol. 656, 11 March 2019.

Lasi, H., P. Fettke, H. G. Kemper, T. Feld and M. Hoffmann, ‘Industry 4.0’, *Business & Information Systems Engineering*, Vol. 6, No. 4, 2014, pp. 239–242.

Lee, J. D., and K. A. See, ‘Trust in Automation: Designing for Appropriate Reliance’, *Human Factors*, Vol. 46, No. 1, 2004, pp. 50–80.

Lyons, Joseph B., and Paul R. Havig, ‘Transparency in a Human-Machine Context: Approaches for Fostering Shared Awareness/Intent’, in R. Shumaker and S. Lackey, eds, *Virtual, Augmented and Mixed Reality: Designing and Developing Virtual and Augmented Environments* (*Lecture Notes in Computer Science*, Vol. 8525), 2014, pp. 181–190.

Lyons, Joseph B., and Kevin T. Wynne, ‘Human-Machine Teaming: Evaluating Dimensions Using Narratives’, *Human-Intelligent Systems Integration*, Vol. 3, 2021, pp. 129–137.

Madhavan, P., and D. A. Wiegmann, ‘Similarities and Differences Between Human–Human and Human–Automation Trust: An Integrative Review’, *Theoretical Issues in Ergonomics Science*, Vol. 8, No. 4, 2007, pp. 277–301.

Madni, Azad M., and Carla C. Madni, ‘Architectural Framework for Exploring Adaptive Human-Machine Teaming Options in Simulated Dynamic Environments’, *Systems*, Vol. 6, No. 4, Article 44, 2018.

Marler, Timothy, Matthew W. Lewis, Mark Toukan, Ryan Haberman, Ajay K. Kochhar, Bryce Downing, Graham Andrews and Rick Eden, *Supporting Joint Warfighter Readiness: Opportunities and Incentives for Interservice and Intraservice Coordination with Training-Simulator Acquisition and Use*, Santa Monica, Calif.: RAND Corporation, RR-A159-1, 2021. As of 3 May 2022: https://www.rand.org/pubs/research_reports/RRA159-1.html

McDermott, Patricia L., and Ronna T. Brink, ‘Practical Guidance for Evaluating Calibrated Trust’, in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 63, No. 1, November 2019, pp. 362–366.

McDermott, Patricia L., Katherine E. Walker, Cynthia O. Dominguez, Alex Nelson and Nicholas Kasdaglis, ‘Quenching the Thirst for Human-Machine Teaming Guidance: Helping Military Systems Acquisition Leverage Cognitive Engineering Research’, paper presented at the 13th International Conference on Naturalistic Decision Making, Bath, UK, 2017.

Ministère des Armées, *Artificial Intelligence in Support of Defence: Report of the AI Task Force*, 2019. As of 29 April 2022: https://www.defense.gouv.fr/sites/default/files/aid/Report%20of%20the%20AI%20Task%20Force%20September%202019.pdf

Morgan, Forrest E., Benjamin Boudreaux, Andrew J. Lohn, Mark Ashby, Christian Curriden, Kelly Klima and Derek Grossman, *Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World*, Santa Monica, Calif.: RAND Corporation, RR-3139-1-AF, 2020. As of 29 April 2022: https://www.rand.org/pubs/research_reports/RR3139-1.html

Nass, C., and Y. Moon, 'Machines and Mindlessness: Social Responses to Computers', *Journal of Social Issues*, Vol. 56, No. 1, 2000, pp. 81–103.

National Audit Office, *Governance for Agile Delivery*, 2012. As of 29 April 2022: https://www.nao.org.uk/wp-content/uploads/2012/07/governance_agile_delivery.pdf

National Security Commission on Artificial Intelligence, *Final Report*, 2021. As of 24 June 2021: https://irp.fas.org/offdocs/ai-commission.pdf

O'Neill, T., N. McNeese, A. Barron and B. Scholle, 'Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature', *Human Factors*, 22 October 2020. As of 3 May 2022: https://journals.sagepub.com/doi/pdf/10.1177/0018720820960865

Packer, Jonathan, *The Ethics of Labor Automation: AI, Technological Disenfranchisement, and Libertarian Paternalism*, thesis, Cambridge, Mass.: MIT Sloan School of Management, 2019. As of 27 May 2021: https://dspace.mit.edu/handle/1721.1/122399

RAN—*See* Royal Australian Navy.

Ransbotham, Sam, Shervin Khodabandeh, David Kiron, François Candelon, Michael Chu, and Burt LaFountain, 'Expanding AI's Impact with Organizational Learning, Findings from the 2020 Artificial Intelligence Global Executive Study and Research Project', *MITSloan Management Review*, 20 October 2020. As of 3 May 2022: https://sloanreview.mit.edu/projects/expanding-ais-impact-with-organizational-learning/

Roff, Heather, and Richard Moyes, 'Meaningful Human Control, Artificial Intelligence and Autonomous Weapons Systems', briefing paper prepared for the Informal Meeting of Experts on Lethal Autonomous Weapons Systems, UN Convention on Certain Conventional Weapons, Geneva, 11–15 April 2016.

Royal Australian Navy, *RAS-AI Strategy 2040*, 2020. As of 3 May 2022: https://www.navy.gov.au/sites/default/files/documents/RAN_WIN_RASAI_Strategy_2040f2_hi.pdf

Ryan, Mick, 'Intellectual Preparation for Future War: How Artificial Intelligence Will Change Professional Military Education', *War on the Rocks*, 3 July 2018. As of 3 May 2022: https://warontherocks.com/2018/07/intellectual-preparation-for-future-war-how-artificial-intelligence-will-change-professional-military-education/

Schaefer, Kristin E., Edward R. Straub, Jesse Y. C. Chen, Joe Putney and Arthur Evans III, ‘Communicating Intent to Develop Shared Situation Awareness and Engender Trust in Human Agent Teams’, *Cognitive Systems Research*, Vol. 46, 2017, pp. 26–39.

Scharre, Paul, *Robotics on the Battlefield Part 1: Range, Persistence and Daring*, Washington, D.C.: Center for a New American Security, 21 May 2014a.

———, ‘How to Lose the Robotics Revolution’, *War on the Rocks*, 29 July 2014b. As of 3 May 2022: https://warontherocks.com/2014/07/how-to-lose-the-robotics-revolution/

———, ‘The Human Element in Robotic Warfare’, *War on the Rocks* (Special Series: Beyond Offset), 11 March 2015. As of 3 May 2022: https://warontherocks.com/2015/03/the-human-element-in-robotic-warfare/

———, ‘Centaur Warfighting: The False Choice of Humans vs. Automation’, *Temple International and Comparative Law Journal*, Vol. 30, No. 1, 2016, pp. 151–165.

Schutte, Paul C., ‘How to Make the Most of Your Human: Design Considerations for Human–Machine Interactions’, *Cognition, Technology and Work*, Vol. 19, Nos. 2–3, 2017, pp. 233–249.

Schwarz, Elke, ‘Autonomous Weapons Systems, Artificial Intelligence, and the Problem of Meaningful Human Control’, *Philosophical Journal of Conflict and Violence*, Vol. 5, No. 1, 2021.

Seeber, Isabella, Eva Bittner, Robert O. Briggs, Triparna de Vreede, Gert-Jan de Vreede, Aaron Elkins, Ronald Maier, Alexander B. Merz, Sarah Oeste-Reiß, Nils Randrup, Gerhard Schwabe and Matthias Söllner, ‘Machines as Teammates: A Research Agenda on AI in Team Collaboration’, *Information and Management*, Vol. 57, No. 2, 2019, Article 103174.

Select Committee on Artificial Intelligence, *AI in the UK: Ready, Willing and Able? Report of Session 2017–19*, London: House of Lords, 2018.

Sheridan, T. B., *Humans and Automation: System Design and Research Issues*, Santa Monica, Calif.: Wiley Interscience, 2002.

Shufutinsky, Anton, James Sibel, Anselm Beach and Anthony Saraceno, ‘OD for Robots? Implications of Industry 4.0 on Talent Acquisition and Development’, *Organization Development Journal*, Vol. 38, No. 3, 2020, pp. 59–76.

Skinner, B. F., *Contingencies of Reinforcement: A Theoretical Analysis*, New York: Meredith, 1969.

Smith, Carol, ‘Creating a Curious, Ethical, and Diverse AI Workforce’, *War on the Rocks*, 5 March 2020. As of 3 May 2022: https://warontherocks.com/2020/03/creating-a-curious-ethical-and-diverse-ai-workforce/

Straus, Susan G., Matthew W. Lewis, Kathryn Connor, Rick Eden, Matthew E. Boyer, Timothy Marler, Christopher M. Carson, Geoffrey E. Grimm and Heather Smigowski, *Collective Simulation-Based Training in the U.S. Army: User Interface Fidelity, Costs, and Training Effectiveness*, Santa Monica, Calif.: RAND Corporation, RR-2250-A, 2019. As of 3 May 2022:
https://www.rand.org/pubs/research_reports/RR2250.html

Suchland, Emma, and Peter Brown, 'Is Your Workforce Fit for the Future?', webpage, PWC, 2021. As of 3 May 2022:
https://www.pwc.com/gx/en/services/entrepreneurial-private-business/fit-for-the-future.html

Tajfel, H., and J. Turner, 'The Social Identity Theory of Intergroup Behavior', in S. Worchel and W. Austin, eds, *Psychology of Intergroup Relations*, Chicago: Nelson-Hall, 1986, pp. 7–24.

UK Ministry of Defence, *Joint Concept Note 1/18, Human-Machine Teaming*, 2018. As of 3 May 2022:
https://www.gov.uk/government/publications/human-machine-teaming-jcn-118

———, 'Defence STEM Engagement', webpage, 17 January 2019. As of 3 May 2022:
https://www.gov.uk/guidance/defence-stem-engagement

UK MOD—*See* UK Ministry of Defence.

U.S. Department of Defense, *Harnessing AI to Advance Our Security and Prosperity: Summary of the 2018 Department of Defense Artificial Intelligence Strategy*, 2019. As of 29 April 2022:
https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF

Warren, Aiden, and Alek Hillas, 'Friend or Frenemy? The Role of Trust in Human-Machine Teaming and Lethal Autonomous Weapons Systems', *Small Wars & Insurgencies*, Vol. 31, No. 4, 2020, pp. 822–850.

Wiener, Norbert, *The Human Use of Human Beings: Cybernetics and Society*, New York: Da Capo Press, 1954.

———, 'Some Moral and Technical Consequences of Automation', *Science*, Vol. 131, 1960, pp. 1355–1358.

Winkler, John D., Timothy Marler, Marek N. Posard, Raphael S. Cohen and Meagan L. Smith, *Reflections on the Future of Warfare and Implications for Personnel Policies of the U.S. Department of Defense*, Santa Monica, Calif.: RAND Corporation, PE-324-OSD, 2019. As of 3 May 2022:
https://www.rand.org/pubs/perspectives/PE324.html

RAND AUSTRALIA

The Royal Australian Navy (RAN) is modernising its forces to better address the growing challenges faced by Australia in the Indo-Pacific region. This report provides an overview of the various impacts of robotics, autonomous systems, and artificial intelligence (RAS-AI) on the Defence workforce to inform the RAN's ongoing efforts to facilitate RAS-AI integration.

The authors draw on a review of relevant open-source academic and grey literature, with a focus on identifying possible lessons for the RAN. The analysis concentrated on the overall impacts of RAS-AI on the Defence workforce and skills, with a particular focus on the implications of human-machine teaming (HMT) for the Defence workforce.

The findings underscore the fundamentally different and novel way of working required to effectively adopt HMT. Integration of HMT into the workforce will require flexible management of complex personnel networks and continuous adaptation of existing structures and concepts.

This report is a continuation of work conducted in support of the RAN's RAS-AI Strategy 2040, released in 2020. RAND Australia was asked to provide policy analysis and advice to support development of an actionable RAS-AI Campaign Plan that could assist RAS-AI implementation efforts. The research team has examined three specific areas to support development of an actionable plan: military innovation, a missions and technology assessment for maritime RAS-AI, and HMT. This work should inform the RAN, other Australian Defence services, and Defence more broadly about the implications of HMT for the RAN's future workforce.

www.rand.org/australia

RR-A1377-2