Inquiry into the Department of Defence Annual Report 2022–23

Submitted to

Joint Standing Committee on Foreign Affairs, Defence and Trade
Foreign Affairs and Aid Subcommittee

Submitted by

RAND Australia
Suite 24, M Centre
11 Palmerston Lane
Manuka ACT 2603

23 January 2024

Contact: Dr Andrew Dowse AO
[redacted]

BACKGROUND

Following an invitation from the Joint Standing Committee, RAND Australia provides this submission in relation to the Inquiry into the Department of Defence Annual Report 2022–23. The Inquiry seeks advice on five topics: Assistance to Ukraine, Defence Health System, Capability Assurance Mechanism, Artificial Intelligence and Autonomous Weapons related issues, and Armaments manufacture, procurement and inventory. Although the RAND Corporation has produced analysis relevant to all of these topics, our submission addresses three of them that relate to recent research undertaken in Australia.

RAND is a respected, nonpartisan entity with a mission to improve policy and decision-making through evidence-based research and analysis. RAND Australia manages multidisciplinary research teams that combine local talent from our small Australian office with over 1,000 research experts from RAND offices in the United States and Europe. Those teams address a variety of defence, national security, social and economic well-being, education, labour and health topics.

CAPABILITY ASSURANCE

Defence acquisition has been criticised for poor performance, with the Defence Strategic Review (DSR) noting it is not fit for purpose and needs to focus on delivery of timely and relevant capability.[1] Of most concern is poor schedule performance[2], which may be caused by a combination of optimism bias in project management, and a generally held belief that schedule slippage is an acceptable outcome in a stable environment in which efficiency trumps effectiveness. With erosion of strategic warning time, schedule can no longer be the poorer cousin in future capability acquisitions.

Defence capability should be delivered within defined requirements of scope, budget and schedule, which are established for each project prior to the acquisition phase.[3] Arguably this practice has not provided clarity as to which of these outcomes should take precedence if they become conflicted. Such capability program guidance of the relative importance of these requirements should be established at Gate 0 and managed throughout the acquisition process.

In some projects, schedule and budget may be less important than developing capability advantage: resulting in ‘first-of-class’ developments that often translate to higher risk. However, the essence of our strategic circumstances and consequent direction is that acquisition projects should be more disciplined, delivering within planned schedule and budget, with scope as a dependent variable[4] – as contrast with the traditional approach of prioritising higher levels of capability regardless of cost and schedule implications (which the DSR characterises as ‘pursuit of the perfect solution’). The DSR also emphasised top-down direction of future capability, a focus on capability integration, and use of off-the-shelf procurements. A further consideration is the temporal nature of advantage as threats evolve, which lends greater weight to an evolutionary approach to capability upgrades.

Thus, greater speed to capability is a key element of the DSR. This has been described in terms of Minimum Viable Capability, which represents a threshold level of capability that offers operational benefit in the shortest time - with the subsequent ability to apply evolutionary upgrades to improve capability. Fundamentally, the process needs to establish up front (i.e., at Gate 0) the critical capability priorities, and then structure the acquisition to deliver optimal outcomes accordingly. This process requires an evidence-based approach to decision-making about project priorities and risks to achievability of outcomes.

[1] Australian Government, National Defence: Defence Strategic Review, 2023, https://www.defence.gov.au/about/reviews-inquiries/defence-strategic-review, p 91.
[2] The ANAO 2021-22 Major Projects Report identified that 34 projects that had exited the MPR had an aggregate slippage of 1363 months. See https://www.anao.gov.au/work/major-projects-report/2021-22-major-projects-report
[3] Department of Defence, Defence Capability Manual, Canberra, https://www.defence.gov.au/sites/default/files/2022-07/Defence-Capability-Manual.pdf, 2022.
[4] Attributed to DEPSEC CASG, see Ben Felton, 20th ADM Congress tackles post-DSR challenges, ADM, 22 June 2023, https://www.australiandefence.com.au/events/event-reporting/20th-adm-congress-tackles-post-dsr-challenges

The Minimum Viable Capability approach focuses on delivery of a threshold level of capability within major acquisition projects, based upon a Pareto principle that the last 20% of capability may come at a premium in terms of additional cost and schedule. While this could be true for many projects, the circumstances will vary and thus each project should be considered in its own context and using the Capability Program Directive as guidance. If pursuing capability as a dependent variable, early analysis needs to be undertaken to determine what level of capability can be achieved within constrained budget and schedule – and whether that level of capability meets the guidance.

Analytical studies undertaken early in the capability process provide robust evidence to support capability decisions[5], helping Defence appreciate the available capability options, fundamental input to capability (FIC) considerations, costs, timeframes, and how risks to achievability should be reduced. There is a higher prospect of poor performance in projects that:

- do not undertake such studies, or
- commission such studies and do not heed the recommendations[6], or
- do not revisit those studies when the circumstances of a project changes.

The Minimum Viable Capability approach represents a new epoch in which capability assurance is relevant throughout the lifecycle. Rather than achieving an optimal level of capability at introduction into service that thereafter degrades comparative to threats, there is an increased expectation for many defence systems that they will be incrementally upgraded at points throughout their life to maintain comparative advantage. Thus, Defence processes should shift from a primary focus on accepting new capability into service, to an ongoing monitoring of capabilities as their effectiveness against emerging threats diminish.

Such a program of evolutionary upgrades during the lifecycle is not well represented in defence capability processes, but may be reflected in future updates of the Defence Capability Manual. There are potential mechanisms for refresh arrangements in current processes, but not a structured approach.[7] An adjustment to the One Defence Capability System might include implementation of a more structured process in which the effectiveness of in-service capabilities against emerging threats and the context of extant capability update plans are regularly assessed by Capability Managers, with changing priorities fed into the Defence Capability Assessment Program (DCAP). This would provide an evidence-driven basis for capability decision-making, as well as the increased agility that Defence requires to effectively deter and respond in our deteriorating strategic environment.

While neither the Defence Annual Report 2022-23 nor the Defence Capability Manual specifically refer to ‘capability assurance’, interest in this term arises from the Defence Capability Assurance and Oversight bill, which promotes the establishment of a Defence Capability Assurance Agency. That agency would be charged with responsibility for assessing the complex risk associated with materiel procurement and sustainment, but focused primarily on test and evaluation.[8]

Test and evaluation is an important activity for Defence, supporting the objective assessment of a system with known confidence.[9] T&E is relevant across the One Defence Capability System, from developmental to in-service phases.[10] The criticality of the T&E function is correlated with the level of risk in the related capability activity, and is especially relevant to developmental and integration risk. With the DSR promoting greater speed of acquisition and off-the-shelf capabilities, care must be exercised so that T&E isn’t associated with the quest for perfectionism, but is undertaken consistent

[5] Defence Capability Manual, p 8.
[6] For example, see Department of Defence, Hunter Class Frigate Procurement Review, 2023, https://www.aph.gov.au/Parliamentary_Business/Tabled_Documents/4366, p 31.
[7] Defence Capability Manual, p 49.
[8] The Senate, Foreign Affairs, Defence and Trade Legislation Committee, Defence Capability Assurance and Oversight Bill 2023, November 2023, https://parlinfo.aph.gov.au/parlInfo/download/committees/reportsen/RB000148/toc_pdf/DefenceCapabilityAssuranceandOversightBill2023.pdf
[9] Department of Defence, Defence Test and Evaluation Strategy, 2021, https://www.defence.gov.au/about/strategic-planning/defence-test-and-evaluation-strategy
[10] Department of Defence, Defence Test & Evaluation (T&E) Strategy, 2021, https://www.defence.gov.au/sites/default/files/2021-09/Defence-Test-and-Evaluation-Strategy.pdf, p 10.

with the level of risk.[11] Indeed, T&E may be more relevant to the conduct of incremental upgrades than the traditional view of T&E typically focused on the acquisition process. Care should also be exercised in relation to the source of capability risk, in that broader FIC or strategic alignment issues may not be addressed through a focus on T&E.

In considering how best to achieve capability assurance, it is advisable to look back but also look forward. In looking to past lessons, there are recurring themes that constitute persistent challenges in capability programs. These arise from factors such as poor requirement setting, production inefficiencies, workforce challenges, misaligned assumptions, poor risk management, platform-centric approaches to capability, cash flow adjustments to the investment program, insufficient contingency to align with risk, optimism, lack of institutional memory and lack of penalties for poor performance.[12] While each capability program will have relatively unique requirements and challenges, there are lessons from historic programs that should be accounted for. In this respect, a key priority to overcome many of these downfalls is to lift institutional capability through an emphasis on professionalisation of project managers.

Notwithstanding the importance of past lessons, there are strategic changes that pose new challenges for future capability assurance. These include evolving missions, a changing defence industrial base, the need for interoperability, cybersecurity, the shift to longer service lives with more upgrades during capability life, the need for investment in workforce after a period of underinvestment, cost growth in weapon systems, and the importance of alignment with strategic goals.[13] While these challenges were identified by RAND in a U.S. DoD context, they apply to Australia as well.

A significant risk to capability assurance will arise from competition for suitable qualified, experienced and cleared workforce. Competition for such workforce will be in the context of an expected shortage of engineering skills in Australia.[14] While skilled migration is proposed to address this gap[15], security clearance requirements will constrain their utility within the defence sector. This dilemma will occur as departments are reducing use of 'inappropriate outsourcing',[16] thus putting more pressure onto the retention of the professional workforce overseeing capability acquisition and sustainment activities. A high turnover of workforce within the defence sector (both within Defence and within industry) would impact capability assurance, noting experience is relevant to productivity, and quality of advice and decision-making. While this will vary with context, RAND analysis indicates a benchmark within shipbuilding studies of at least 5 years before employees can be presumed to be fully productive.[17] Even higher levels of competency through suitable training and experience is critical for personnel involved in training and mentoring the next generation of workforce, as well as those undertaking T&E functions.[18] The prospect of inexperienced project and capability management should give greater priority to the conduct of early risk reduction, including independent assessments, and with scrutiny maintained throughout the program lifecycle.

Capability assurance is critical throughout the service life of a system. This puts an onus onto capability managers and delivery agencies to consider objectives, system performance, costs and risks on an ongoing basis, and applies to upgrades and sustainment activities. Such activities should be

[11] Defence Strategic Review, loc cit.
[12] Lucia Retter et al., Persistent Challenges in UK Defence Equipment Acquisition, RAND Report A1174-1, 2021, https://www.rand.org/pubs/research_reports/RRA1174-1.html
[13] Jonathan Wong et al., Improving Defense Acquisition, RAND Report A1670-1, 2022, https://www.rand.org/pubs/research_reports/RRA1670-1.html p v-vi.
[14] Professionals Australia, Engineering a Better Future: Australia's Growing Crisis in Engineering Skills, March 2023, https://apesma.informz.net/apesma/pages/EABF_2023
[15] Tech Council of Australia, Getting to 1.2 million: Our roadmap to create a thriving Australian tech workforce, August 2022, https://techcouncil.com.au/wp-content/uploads/2022/08/2022-Getting-to-1.2-million-report.pdf
[16] Australian Government, APS Strategic Commissioning Framework: Strengthening APS capability through reduced reliance on contractors and consultants, APS Commission, October 2023, https://www.apsc.gov.au/sites/default/files/2023-10/Strategic%20Commissioning%20Framework_2.pdf
[17] John Birkler et al., Australia's Naval Shipbuilding Enterprise, RAND Report 1093, 2015, https://www.rand.org/content/dam/rand/pubs/research_reports/RR1000/RR1093/RAND_RR1093.pdf, p 185.
[18] Australian National Audit Office, Test and Evaluation of Major Defence Equipment Acquisitions, 24 November 2015, https://www.anao.gov.au/work/performance-audit/test-and-evaluation-major-defence-equipment-acquisitions-0

planned early, should leverage common strategies, and should balance the extent of such activities with the risks that they mitigate.[19]

As noted above, the key to effective capability assurance is early (pre-gate 0) and ongoing risk reduction through analysis of objectives, alternatives and constraints. RAND has developed a framework that provides robust evidence to inform Defence decision-making in relation to the achievability, affordability, capability outcomes and risks of capability investments, including FIC considerations. This process utilises Assumption-Based Planning to model critical vulnerabilities, to assess how different assumptions will affect outcomes, and to allow decisions to be reconsidered as circumstances evolve and throughout the capability lifecycle. The utility of such analysis provides clear accountability for decisions and the basis by which the performance of defence capability processes can be enhanced.

Artificial Intelligence and Autonomous Weapons

With great power competition and a general increase in global instability, nations such as Australia are considering employment of emerging technologies such as Artificial Intelligence (AI) and autonomous systems. This includes our U.S. allies, who consider exploitation of such technologies as the 'third offset'.[20] Australia's involvement in allied development of technologies and operational concepts for AI and autonomous systems represents a key source of asymmetric advantage.[21]

Any discussion of military applications of Artificial Intelligence, including but not limited to Autonomous Weapon Systems (AWS) should begin with an acknowledgement of three challenging aspects of the technology. First, AI is a broadly applicable term for what is essentially an enabling technology. Thus, it is more helpful to conceptualise AI as electricity or steam than as a car or a train. While current generation systems could be classified as task-based or narrow AI that generate predictable outputs from its interpretation of a human-given objective, the technology continues to advance rapidly. AI enables 'autonomy', which in turn refers to the relative (not binary) capacity of a system to 'execute a task, or tasks, without human input, using interactions of computer programming with the environment'.[22]

Second, there is no universally agreed definition, standard or benchmark that an agency could use to definitively state whether a given system is truly autonomous. There are a number of published works that discuss the definitional challenges in greater detail, however, for the purposes of this submission, it is sufficient to utilise the autonomy spectrum adopted by the Australian Army in its Robotics and Autonomous Systems Strategy. This framework is built around four levels of autonomy: remotely operated (which are often incorrectly conflated with autonomous systems), automatic (where a human remains in the loop to monitor and potentially intervene), autonomic (where the human supervises or tasks a system, thus remaining in the decision loop), and autonomous (where the human starts the decision loop but the system can then act independently).[23] While similar to definitions adopted by the U.S. military, China, the UK, and civil society groups,[24] each differs to some legally significant degree. Where the Australian Government chooses to draw this line, and the technical standards required to assess against that line, will have a significant impact on the feasibility and practicality of future policy levers.

[19] John Drew et al., Enabling Early Sustainment Decisions, RAND Report 397, 2013, https://www.rand.org/pubs/research_reports/RR397.html
[20] Gian Gentile et al., A History of the Third Offset, 2014-2018, RAND Report A454-1, 2021, https://www.rand.org/pubs/research_reports/RRA454-1.html
[21] Andrew Dowse et al., Allied Asymmetric Advantage in the Air, limited distribution concept paper, January 2024.
[22] V Boulanin, 'Mapping the development of autonomy in weapon systems: A primer on autonomy', *Stockholm International Peace Research Institute*, 2016.
[23] RICO, 'Robotic and Autonomous Systems Strategy v2.0', Australian Army Future Land Warfare Branch, 2022.
[24] For an overview: Austin Wyatt, "So Just What Is a Killer Robot?: Detailing the Ongoing Debate around Defining Lethal Autonomous Weapon Systems." *Wild Blue Yonder*, 2020.

Third, the technology has continued to advance at lightning pace, but in the absence of effective international regulation or united policy response. The release of Chat GPT-3 in November 2022 clearly captured the attention of both policymakers and the general public, with both groups still grappling with the implications of the wave of proliferation in other generative AI tools (across writing, audio and video domains) through 2023. Similarly, the Russian invasion of Ukraine has seen the widespread use not only of remote operated first-person UAVs, but also the use of AI and semi-autonomous weapons on the battlefield. The best known of these concerning developments has been the widespread use of loitering munitions (e.g., Shahed-136),[25] which can independently select and engage targets based on a pre-established database.[26] Although loitering munitions have been described in the media as lethal autonomous weapons systems,[27] debate in the legal[28] and scholarly[29] communities is ongoing. The spread of these technologies is even more pronounced and rapid in the commercial space. This trajectory will further increase the complexity and political sensitivity of imposing effective and enforceable regulatory guardrails around actors developing such technologies. It would, therefore, be valuable for policy makers and the ADF to devote resources to understanding how to best undertake risk reduction and management for both our own use of autonomous systems as well as that of other international actors.

The ADF has a clear interest in a range of emergent applications of AI technologies in the military domain, including increasingly autonomous weapon systems. Each of the service branches has established units focused on integrating such emerging technologies. RAND has successfully supported a number of the resulting efforts, including providing support to the Royal Australian Navy's Campaign Plan for Robotics and Autonomous Systems,[30] and the Australian Army Robotics Integration and Coordination Office's consideration of how to overcome barriers to adopting such technologies at scale.[31] RAND has similarly supported efforts by the U.S. military,[32] and recently completed an exploration of how to most effectively approach co-design of military AI under the AUKUS framework.[33]

Crucially, underpinning the ADF's approach to AWS is a commitment to ensuring that such systems are only ever employed in a manner that is ethical and compliant with Australia's obligations under international humanitarian law. It is easy to make the argument that such commitment places the ADF at a disadvantage compared to less scrupulous strategic threat actors, but it is important that Defence leadership continues to disregard such an argument. Quite aside from Australia's obligation

[25] S Russell, 'AI weapons: Russia's war in Ukraine shows why the world must enact a ban', *Nature*, vol. 614(7949), 2023, pp. 620–623.
[26] Boulanin, 2016.
[27] J Kahn, 'A.I. drones used in the Ukraine war raise fears of killer robots wreaking havoc across future battlefields', *Fortune*, 29 March 2022, https://fortune.com/2022/03/29/artificial-intelligence-drones-autonomous-weapons-loitering-munitions-slaughterbots-ukraine-war
[28] I Bode, H Huelss, A Nadibaidze, G Qiao-Franco and TFA Watts, 'Prospects for the global governance of autonomous weapons: Comparing Chinese, Russian, and US practices', *Ethics and Information Technology*, vol. 25, no. 1, 2023, p. 5.
[29] Boulanin, 2016.
[30] Peter Dortmans, Joanne Nicholson, James Black, Marigold Black, Carl Rhodes, Scott Savitz, Linda Slapakova, and Victoria M. Smith, *Supporting the Royal Australian Navy's Strategy for Robotics and Autonomous Systems: Building an Evidence Base*, RAND Corporation, 2021; Slapakova, Linda, Paola Fusaro, James Black, and Peter Dortmans, *Supporting the Royal Australian Navy's Campaign Plan for Robotics and Autonomous Systems: Emerging Missions and Technology Trends*, RAND Corporation, 2022; Black, Marigold, Carl Rhodes, James Black, Rebecca Lucas, and Linda Slapakova, *Supporting the Royal Australian Navy's Campaign Plan for Robotics and Autonomous Systems: Enhancing Innovation*, RAND Corporation, 2022; Black, Marigold, Linda Slapakova, Paola Fusaro, and James Black, *Supporting the Royal Australian Navy's Campaign Plan for Robotics and Autonomous Systems: Human-Machine Teaming and the Future Workforce*, RAND Corporation, 2022.
[31] Austin Wyatt, Joanne Nicholson, Marigold Black and Andrew Dowse, *Understanding how to scale and accelerate the adoption of RAS into deployable capability. Phase 1—Identifying Barriers*, RAND Corporation, Forthcoming.
[32] Recent examples that are publicly available include: Ryseff, James, Eric Landree, Noah Johnson, Bonnie Ghosh-Dastidar, Max Izenberg, Sydne J. Newberry, Christopher Ferris, and Melissa A. Bradley, *Exploring the Civil-Military Divide over Artificial Intelligence*, RAND Corporation, 2022; Wilson, Bradley, Ellen M. Pint, Elizabeth Hastings Roer, Emily Ellinger, Fabian Villalobos, Mark Stalczynski, Jonathan L. Brosmer, Annie Brothers, and Elliott Grant, *Characterizing the Uncrewed Systems Industrial Base*, RAND Corporation, 2023.
[33] Austin Wyatt, James Ryseff, Elisa Yoshiara, Benjamin Boudreaux, Marigold Black and James Black, *Towards AUKUS Collaboration on Responsible Military Artificial Intelligence: Co-Design and Co-Development of AI Among the United States, the UK and Australia*, RAND Corporation, forthcoming.

to ensure that weapon use is consistent with the laws of war, the ADF owes an obligation to Australian soldiers, sailors and aircrew that it will not expose them to the kind of moral injury that would result from a co-deployed AWS engaging an illegitimate target. That being said, this position does present Australia with a wicked problem: how should the ADF secure a competitive advantage whilst ensuring that its use of AWS, and other military applications of AI, remains ethical and legal?

The use of AI in military applications has raised a number of ethical, moral, and legal challenges that continue to dominate the discourse around these technologies. While prominent in the case of AWS, these issues range from the protection of people, society and the environment; algorithmic bias and discrimination; explainability, reliability and accessibility; accountability and responsibility mechanisms; and privacy and transparency.[34] There are related concerns that the use of AWS would lead to the dehumanisation of the decision to apply lethal force,[35] and that policy makers would be less deterred from resorting to warfare without the political risk of placing human soldiers at risk of harm. Similarly, there is a risk that the deployment of autonomous systems and AI into sensitive environments, even if they are not armed, without norms governing their deployments raises the risk of unintended escalation or mistaken engagements.[36] Miscommunication or unexpected responses could occur if, for example, one of our regional neighbours were to breach the airspace of another (intentionally or not), with an uncrewed aircraft, or if an uninhabited surface vessel were to collide with a crewed fishing vessel in contested waters. These risks are driving calls for retaining human control over such systems, even if that reduces their comparative effectiveness.[37] How to actually achieve such control remains an open question, and one that is complicated by the dual-use nature of AI, and its growing ubiquity in civilian systems,[38] raising the risk that the pursuit of competitive advantage leads to a 'race to the bottom' when it comes to ethical and legal compliance of military AI.[39]

Arguably the most publicly prominent of the issues surrounding specifically military applications of AI is the ongoing discussions occurring at the United Nations surrounding how to integrate Lethal Autonomous Weapon Systems (LAWS) into international humanitarian law. Australia has been an active participant in these discussions, for example presenting the ADF's system of control to a 2019 Meeting of the Group of Governmental Experts on Emerging Technologies in the Area of LAWS. Whilst this process has generated valuable discussion and some steps have been taken - for example, the affirmation of eleven guiding principles in 2019 - this process has not yet resulted in new international law. Proponents of the position that LAWS should be subjected to a binding prohibition recently celebrated the passage (with Australian support) of UN General Assembly Resolution 78/241, which acknowledged "serious concerns" stemming from LAWS and requested that the Secretary General seek the views of member states toward a report (due September 2024) on how to address ethical, legal and humanitarian risks stemming from LAWS.[40] Although a strong signal in moving the issue to the agenda of the General Assembly, this resolution still simply calls for more discussion and review. As at the time of writing, therefore, we remain far from specific international humanitarian law prohibitions on any military application of AI beyond those that Australia is already obligated to follow.

Among the few principles that is almost universally agreed to amongst the international community in relation to AWS (and military applications of AI more broadly) is that the decision to end human life should remain under 'Meaningful Human Control'. This was presented within

[34] Also concerns that use of LAWS may create anti-Western sentiment in areas affected by strikes, due to the perceived indignity and unfairness of being injured or killed by an unmanned system. Interestingly, a recent article demonstrated whether racial biases shape support for drone strikes. See P Lushenko, K Carter and S Bose, 'Do racial biases shape Americans' support for drone strikes? We asked them', *Modern War Institute*, 17 April 2023.
[35] YH Wong et al., *Deterrence in the age of thinking machines*, RAND Corporation, 2020.
[36] I Bode and H Huelss, 'Autonomous weapons systems and changing norms in international relations', *Review of International Studies*, vol. 44, no. 3, 2018, pp. 393–413.
[37] FE Morgan et al., *Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World*, RAND Corporation, 2020, pp. xiv–xv.
[38] A Wyatt and J Galliott, 'The revolution of autonomous systems and its implications for the arms trade', in ATH Tan, ed., *Research Handbook on the Arms Trade*, Edward Elgar Publishing, 2020.
[39] Morgan et al., 2020, p. xiv.
[40] United Nations, Lethal Autonomous Weapon Systems, General Assembly Resolution 78/241, 12 October 2023.

quotations deliberately, because there is no universal agreement of what this term actually means, or how it could be reflected in an objective measurable standard. For example, the UK military recognises that it must maintain meaningful human control, as does the Australian Army,[41] but neither offers criteria for determining if a system does so.[42] In 2022, RAND Australia undertook an analysis of Human-Machine Teaming and the future workforce for the Royal Australian Navy, in which common characteristics of Meaningful Human Control were identified including that an operator should have sufficient information to make a "conscious and informed decision", that the system should be predictable, transparent, and reliable – which requires effective testing and evaluation, and that the system should allow for sufficient time for the human to make an informed decision whether or not to allow that engagement.[43] This reflects the position advocated by the U.S. delegation to the UN Group of Governmental Experts, which focused on the legal review process and design factors without direct human control over the release of force. The UK delegation went further, arguing that a human must remain in the decision loop. The Australian delegation presented a process-based system of control, focusing on mechanisms for ensuring human decision makers retain effective, if not temporally direct control over autonomous weapons.[44]

Central to establishing, and demonstrating, that meaningful human control (however one defines the specific requirements) is maintained of a military application of AI is an effective Test, Evaluation, Verification and Validation (TEV&V) model.[45] However, this is complicated by the fact that AI enabled systems can act in undesirable or unpredicted ways[46] due to the complex, and often opaque, interactions between system elements, behaviour and performance.[47] Our existing conceptions of TEV&V do not serve us particularly well when the system must be tested for the reliability and effectiveness of its decision-making, rather than the binary effectiveness of if the weapon successfully deploys; nor do current testing approaches fully capture the effect of post-deployment learning, emergent non-deterministic behaviour of AI-enabled systems,[48] or the convergence effect (where integrating AI has unexpected effects at the system level).[49] As a result of the potential that such systems would act in unexpected ways during testing, militaries require significant safety features and larger than traditional exclusion zones. Complex AI-enabled systems have a well-known tendency to fail spectacularly and destructively, usually with little obvious warning,[50] although this can be somewhat mitigated by design stage decisions that encourage the system to return to a safe state upon detection of unexpected input or error. Australia could leverage its superior access to physical space by facilitating the testing of such systems by the ADF and partner militaries; for example, this is a known challenge for the South Korean and Singaporean militaries.

In the absence of meaningful progress toward international regulation of LAWS, or other military applications of AI, Australia has an opportunity to leverage its position as a trusted middle power to promote the generation of confidence building and de-escalatory norms of behaviour. Equally, Australia should invest in identifying mechanisms and potential friction points in how the ADF and

[41] Australian Delegation, 'The Australian Article 36 Review Process', Second Session, Group of Governmental Experts of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to be Excessively Injurious, 27–31 August 2018.
[42] Development, Concepts and Doctrine Centre, *Joint Concept Note 1/18 Human Machine Teaming,* Joint Concept Note, U.K. Ministry of Defence.
[43] Marigold Black, Linda Slapakova, Paola Fusaro, and James Black, Supporting the Royal Australian Navy's Campaign Plan for Robotics and Autonomous Systems: Human-Machine Teaming and the Future Workforce. Santa Monica, CA: RAND Corporation, 2022.
[44] Australian Delegation, 'The Australian Article 36 Review Process', Second Session, Group of Governmental Experts of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to be Excessively Injurious, 27–31 August 2018.
[45] M Verbruggen, 'No, not that verification: Challenges posed by testing, evaluation, validation and verification of artificial intelligence in weapon systems', in T Reinhold and N Schörnig, eds, *Armament, Arms Control and Artificial Intelligence: The Janus-faced Nature of Machine Learning in the Military Realm*, Springer, 2022.
[46] S Ferreira et al., 'Unmanned and autonomous systems of systems test and evaluation: challenges and opportunities', 2010.
[47] Verbruggen, 2022.
[48] Verbruggen, 2022.
[49] Verbruggen, 2022.
[50] Wyatt, 2021.

Australian industry could leverage our improved access to the U.S. and UK defence ecosystems to build its influence and access to such technologies.

Another short term policy response that the Australian government could take in this space would be to provide training for APS and ADF members on the potential, limitations, and risks associated with emergent AI and robotics tools. Humans have a tendency to overly trust in technology once it diffuses and matures; however this automation bias can lead trained and experienced humans to trust in the technology rather than their own judgement, amplifying the risk of unanticipated consequences, for example via malicious actor interference or training data bias. Early familiarisation with the current state of technology (both its opportunities and its limitation) reduces these risks. While the decision to trial Microsoft Co-Pilot is one example of this, more could be achieved through investing in a concentrated effort to build AI literacy more broadly both within Defence and in the wider public service. Such training need not be expensive, with serious policy wargaming, class-room based education, and the use of training proxies being just some alternatives to speculative procurement.

Finally, Australia should continue to invest in the education and training pipeline for relevant STEM skill sets, from higher-order mathematics to software engineering. Whilst Australia lacks the specialised industrial base and resource capacity to compete with the United States and China to become a leading source of compute power or semi-conductors, it can generate a leading advantage as a source of top-level talent. Talent remains an important bottleneck in the development of AI technologies in both the commercial and military spheres. Although recruitment and retention is already a known challenge for the ADF and a focus of efforts to develop Australia's AUKUS Pillar One workforce, developing the capacity to build a knowledge edge would give Australia an outsized influence in the development of AI that would support Pillar Two efforts.

Reflecting the importance of understanding the potential risks of future Artificial Intelligence developments, RAND recently established the Center for Technology and Security Policy (TASP). This centre focuses on answering the challenges posed by potentially existential risks emerging from Artificial Intelligence and its intersection with our world. RAND's Artificial Intelligence research continues to be largely made available to the public and policy makers on our website.

Armaments Manufacture, Procurement and Inventory

RAND Australia has undertaken two analyses in support of Defence's Guided Weapons and Explosive Ordnance enterprise. The first was a comparative evaluation of the lessons of case studies relevant to such an enterprise. A published report[51] examined the weapons industries of five nations, with a separate non-public report examining two domestic case studies of relevance. This first report found that:

- While case studies revealed lessons for Australia, our needs for a weapons enterprise are unique and require definition and prioritisation of our desired outcomes;
- Similarly, our need for sovereign capability should be defined – operational sovereignty is concerned with removing the risks of loss of access to or control over capability elements required to achieve an effect;
- Complex weapons enterprises require decades to develop, requiring investment in supporting innovation and education systems;
- There is a cost premium to reducing sovereign risk;
- Joint development with allies and partner nations is attractive from an interoperability and commonality perspective, as well as to benefit from shared development investments, but is not without risks;
- A continually adapting offsets program can enable growth of the defence sector; and
- Deliberate analysis is needed to right-size industrial capacity to achieve a capability that is sufficient for Australia's needs and adaptable to compete internationally.

[51] Christopher Mouton et al., Establishing a Sovereign Guided Weapons Enterprise for Australia: International and Domestic Lessons Learned, RAND Report A1710-1, 2022, https://www.rand.org/pubs/research_reports/RRA1710-1.html.

In 2022, RAND began a subsequent analysis to inform the establishment of the weapons enterprise. It examined:

- The outputs of a guided weapons and explosive ordnance enterprise;
- How the enterprise might serve as a catalyst for innovation;
- The intent for sovereign capability and what trade-offs might be involved in achieving it;
- The implications for different options for the enterprise for weapons stockpiling and production;
- The requirements for weapons in protracted conflict; and
- The costs associated with establishing a weapons enterprise.

The report of this second tranche of analysis was not openly published, due to classification, but an unclassified derivative short perspective[52] was published in mid 2023. The perspective noted that a significant source of sovereign risk is the prospect of conflict in which demand exceeds supply from overseas sources. A shift towards greater domestic manufacturing would need to consider five key factors:

- Capability: the greater the variety of weapons, the more challenging and costly it will be to rely upon domestic production.
- Supply chains: more sophisticated weapons involve complex supply chains, which makes it more difficult to mitigate sovereign risks.
- Time: there needs to be a balance between reducing short term risks (through stockpiling) and investing in longer term development of a weapons industry. Such decisions need to be made on the basis of *strategically relevant timeframes*, which are associated with preparation of capabilities before conflict. The associated risk that such timeframes may be short may demand an initial focus by the enterprise on enhancing the force-in-being.
- Cost: planning of a weapons industry should account for the whole of life cost of a domestically produced weapon and the demand for that weapon. Analysis might also account for the associated cost that might be imposed on an adversary through its application – which may lead the enterprise to consider affordable mass options.
- Surge: the predominant risk to the enterprise is the gap between weapons requirements in peacetime and in times of conflict, which may be further exacerbated by protraction of conflict. Stockpiling may be an option, albeit one with significant practical constraints. This risk is characterised in terms of *operationally relevant timeframes*, which is associated with the ability to replenish capabilities within conflict. The potential utility of weapons should inform the need for stockpiling of weapons as well as to invest in industry to enable a subsequent capacity to surge production.

The Ukraine conflict has been used to support various ‘lessons’ for Australia, including the relative utility and effectiveness of various weapons, although care needs to be exercised in view that Ukraine is a quite different environment compared to potential threats that Australia may face. One clear lesson however is that the gap between peacetime production and the demand of protracted conflict represents a significant strategic risk. The need to enhance the force-in-being demands that the balance between buy and build initially should favour the former to increase stock in the short term. Whereas developing domestic production is an option that should be pursued, it should be done so (1) in a targeted approach, based upon likely demand and sovereign risks, noting that attempting to produce multiple weapons will come at an opportunity cost, and (2) viewing schedule optimism of domestic industry with caution, especially noting the need to invest time and money in extensive T&E in order to have confidence in new sources of weapons.

The perspective also highlighted four non-mutually exclusive pathways to develop a domestic production capability. These included improving whole-of-system expertise through increased maintenance roles, development of technological expertise through supply chain participation, involvement in co-production of existing systems in which demand is exceeding supply, and

[52] Andrew Dowse et al., Australia’s Sovereign Capability in Military Weapons, RAND Perspective A2131-1, 2023, https://www.rand.org/pubs/perspectives/PEA2131-1.html

development of new weapons technologies. A key activity however is for the GWEO enterprise to engage with Defence's Force Design organisation to appreciate weapon priorities, gaps in inventory and industry options to deliver on those needs.

CONCLUSION

The team at RAND Australia appreciates the opportunity to contribute a submission to the Joint Standing Committee's review of the Defence Annual Report 2022-23. We have confined the comments in our submission to three areas of interest in which we have undertaken recent studies.

One common theme across the three areas is the vital importance of evidence-based analysis to inform decision making about future defence capabilities. RAND has demonstrated value in undertaking such studies, in order to appreciate alignment of investments with strategic objectives, analyses of alternatives, evaluation of cost, schedule and other risks as well as FIC implications, and use of frameworks for ongoing management of programs and uncertainty.

While the terms of reference of this inquiry identified test and evaluation as relevant to capability assurance in Defence, it represents a second theme across the three areas. T&E is a key contributor to capability assurance, although we note that broader analysis is needed to identify all sources of capability risk; and the requirement for T&E needs to be balanced with the need for speed to capability and greater use of off-the shelf acquisitions. Systems that feature artificial intelligence technologies represent a challenge to T&E, in that evaluating trust in a system's predictability is difficult to achieve in systems that do not act in a deterministic manner.[53] Finally, we note the criticality of comprehensive T&E activities as an integral element of a domestic weapons industry, to give confidence in weapon safety and fitness for purpose.

As an organisation that is committed to only undertaking high quality objective analysis to address public policy challenges, RAND Australia is well placed to support Defence into the future.

[53] Andrew Dowse, The need for trusted autonomy in military cyber security, in H.A. Abbass et al (eds) Foundations of Trusted Autonomy, 2018, https://doi.org/10.1007/978-3-319-64816-3_11, p 207.