

Senate Inquiry Submission

Defence Capability Assurance & Oversight Bill

Dr Keith F. Joiner, CSC (Group Captain, Ret'd)

Senior Lecturer Test and Evaluation, Capability Systems Center, University of New South Wales, Canberra

Executive Summary

I have a career in test and evaluation culminating in Defence as Director General for Test and Evaluation in 2010-14. Since 2015 I have been a Senior Lecturer in test and evaluation with the University of NSW, researching and teaching test and evaluation planning and techniques. I have graduated over 1200 students in test subjects within the coursework Master programs.

Test and evaluation are crucial for informed decision-making. Contrary to popular perception, when tests and evaluations are planned and resourced correctly, it improves the speed of capability delivery by involving representative users and stakeholders directly in practical test activities that reduce bureaucracy, build consensus and disclose risk early. However, to do that efficiently with modern technology, the complexity of interconnected and autonomous Defence capabilities and advanced persistent cyber threats requires acquirers and testers to be proficient in modelling, simulation, digital engineering, agile requirements methods, cybersecurity and modern test design. There is an alarming and widening gap in all these areas for testing and evaluation in Australian Defence compared to the U.S. Defense. If this gap is not addressed, Australia will not have the agility to match near-peer threats and will be less trusted to adapt and integrate with its allies. This gap spans competencies, policies and infrastructure, slowing down and obfuscating most acquisitions.

Further, perspectives in the U.S. and the U.K. on Defence capability assurance have fundamentally changed compared to Australia. Public accountability has led to independent oversight of capability assurance mechanisms in those countries but not here. In the U.S., it is Title 10 congressional law founded in the so-called 'Pentagon Wars' which the proposed Australian Bill seeks, inter alia, to emulate. In the U.K., the Infrastructure and Projects Authority (IPA) is the Government's recent centre of expertise for infrastructure and major projects, including oversight of Defence Projects, reporting to the Cabinet Office and His Majesty's Treasury.¹ These legislative initiatives in capability assurance shift oversight on key disciplines in capability assurance from the 'toolbox' and discretion of program managers to legislated oversight agencies. In the simplest form, they implement public expectation that *'you shouldn't buy a car without a test drive'*, or in Defence terms, by *'a fleet of cars with public monies without a test drive by representative users before contract and production.'* These agencies do an enormous amount in quality education, policy initiatives and 'best practice guidance. Australia is long overdue to follow its AUKUS leads with similar legislated initiatives and oversight, be that just Defence or more broadly.

There are two reasons for the widening gap between Australia and the U.S. Defence in capability assurance. The first reason is an Australian problem with stagnant and self-reinforcing policies and practices in acquisition, system engineering, and testing and evaluation. Globally research and practices in these fields have moved dramatically since the 20th Century dogma on which the majority of Defence capability acquisition and assurance is practised. Defence largely internally promotes an aging acquisition and assurance workforce with self-consultation about competencies that do not embrace global practices, like agile requirements development, software development and testing, test design, modelling and simulation, cybersecurity, and digital engineering. Research notes that bureaucracies with directive control can become entrenched in waterfall, lifecycle and contract approaches. Shifting these dogmatic practices cannot rely on individuals finding their own

¹ <https://www.gov.uk/government/publications/infrastructure-and-projects-authority-mandate/infrastructure-and-projects-authority-mandate>; <https://www.gov.uk/government/news/ipa-2022-23-annual-report-on-major-projects>.

training and trying innovative approaches; instead, changing this dogma requires pro-active policy and directed re-education programs, with some significant test infrastructure support.

Stagnant Australian policies are also caused by the fragmentation of the Defence Department across 13 services and groups, or 14 if you count the new Nuclear Propulsion Organisation. There is a consensus approach to central policy that can take years for even minor changes. Information gathered in one area, such as a space satellite, can pass through numerous other groups and services, like datalinks, base infrastructure, command planning, etc., each with its own technical assurance and differing testing standards. Defence needs to be able to prescribe the minimum standards of cybersecurity assessment and test across all services and groups, including the nuclear capabilities, quickly and to a higher standard than other Government departments due to Defence's generally higher threat risk.

The second reason for the widening gap is the U.S. decision to create the Directorate of Operational Test and Evaluation (DOT&E), working directly for the Office of Secretary of Defense (i.e., Minister). Despite having 'Operational' in the title, under Title 10 Congressional Law, DOT&E licenses and approves all test and evaluation planning and appointments across the whole lifecycle, reports annually to Congress, and provides funding to cover significant test processes, education and infrastructure improvements. This independence has enabled DOT&E to help its Defense respond more quickly to cybersecurity threats and use modern, efficient modelling, simulation and test designs in their acquisition. Importantly, their influence has created three passes of Government approval compared to Australia's two passes, whereby contracted development can only proceed to low-rate initial production until an operational assessment is complete. This extra milestone avoids over-commitment in costly production until proper testing of prototypes and the associated risk disclosure, generally with schedule and cost advantages.

U.S. Agencies like the Test Resource Management Centre (TRMC) and their support from the Defence Acquisition University (DAU) mean the leading edge of test practices and infrastructure are directly funded, are quickly supported by policy for Defense use, and there is education and competency to use these. In contrast, in Australia, the responsible areas for testing and evaluation are chronically underfunded, have a minimal policy and test infrastructure, with very modest overall test competency requirements. These deficiencies are despite two decades of significant related ANAO reports, Defence's Roadmap for Test and Evaluation, and a Senate Inquiry into Defence Procurement. The First Principles Review sought a strong strategic centre with the Vice-Chief, which ultimately included test and evaluation governance. The test and evaluation governance has a new strategic plan for test and evaluation based on priorities in the Government's Sovereign Industrial Capabilities, but funding and progress since show that Defence lacks the wherewithal to address the significant test policy, infrastructure and education changes needed.

One example of the paucity of Australian Defence response to ANAO auditing and the complexity of modern technology is the inadequacy in the minimum competency levels for managing tests and evaluations. Adding up the Defence training courses in test and evaluation, you get just four weeks of training, more appropriate for a Council roundabout than the programs Defence manages. Even the systems engineering master program I manage has one core introductory subject to test and evaluation at 160 hours of work or four weeks of full-time intensive. The second elective subject I run to help match U.S. DoD competencies in test design attracts less than five per cent of my students and extends our UNSW offerings to just two months full-time. A month or two of education is much less than the one year that has been the standard in overseas flight tests for the last four decades. While our Defence sends officers to do this one-year education, it is a handful and does not cover the many non-aircraft capabilities that today have equal complexity to aircraft. Many of the subjects outlined in this submission as deficient in most of our capability assurance staff would fill the remaining year of complete education in this field.

I support the formation of a more independent capability assurance agency similar to the U.S. DOT&E. The ANAO first recommended this in 2001 (Recommendation 2), and Defence argued it was unnecessary on several grounds that, with the benefit of two decades of hindsight, was probably

disingenuous. The new agency must first determine and enforce the minimum standards to assure all Defence capability and then proactively pursue best-practice improvements that address technology and threat advancements with greater efficiency. Such an agency needs the imprimatur and funding mechanisms of some legislation to prescribe these new policies, curriculum and infrastructure in design assurance, systems engineering and testing. Eleven main areas need urgent action across the pillars of assurance policy, infrastructure and education. These areas include agile testing, software testing, integrated testing, cybersecurity assessment and testing, test designs, autonomous system tracking, modelling, and simulation.

For example, urgent funding is needed for a joint network integration centre (JNIC) to be the heart of such assurance and test capability, federating Australia's test facilities into a Defence simulation network and cybersecurity-accredited cyber-ranges. These initiatives can be undertaken with the direct assistance of DOT&E under the multilateral T&E program (MTEP) MOU and FMS funding or through new AUKUS agreements. If the Government does not legislate and create some form of Capability Assurance Agency with authority and independence, our military commanders, positioned as they are in a very diverse and interdependent organisation, will continue to be self-determinant, dogmatic and mostly dated, unknowing and inefficient in the way they assure their capabilities. We will continue to lag behind our U.S. allies who took this legislative step and investment decades ago and are reaping the benefits. The AUKUS agreements represent a renewed investment in Australian Defence that warrants Australia showing it will align our capability assurance to the U.S. standards. A Defence Capability Assurance Agency offers the authority, funding and independence to close this otherwise widening gap in capability assurance.

Scope and Limitations

This submission is set out in more detail per the overview above. I first briefly outline my qualifications to submit on this topic before outlining the widening gap between the U.S. and Australia in Defence Capability assurance. I then cover two main reasons why that gap is widening before assessing how it can be addressed, concluding with 11 recommendations across assurance testing policy, infrastructure and education.

As a researcher, I am critically aware that these submissions do not have the latitude to provide background and research findings. Much of the necessary research, definitions and supporting principles to support many of the arguments herein are at the references I have co-authored [1-5] and in key overseas reports [6-13]. I am willing to testify 'in-camera' to assist the Inquiry.

This submission must generalise, and, in every criticism, I regret there are pockets of excellence that may rail at the criticism. If your experience is in one of those pockets, please reflect on what the middle half of all capability projects are doing and then the bottom quarter. For example, there are test pilots and engineers educated overseas for a year in the highest testing standards, capabilities developed and supported by World's-best modelling and simulation, and a few projects trying agile approaches. However, like cyber threats, Defence is judged by its weakest links, so please, if the criticism seems unfair, do try to consider the average practice in Defence, not your best practice.

A. Qualification to Submit

My distinguished career in test and evaluation culminated in Defence as Director General for Test and Evaluation in 2010-14, with responsibility for vetting all master test plans to the Government and for the conduct of testing across all domains (i.e., air, sea, land, joint, space) for at least some parts of their lifecycle (i.e., maritime preview & development, land preview and operational etc.). Some of my unique blend of test experience in Defence include the following over 30 years of permanent service and eight years of reserve service (so far):

- Maritime testing of air-delivered sea mines, torpedoes, anti-ship missile defence, amphibious craft and ships, future submarines, air warfare destroyers, underwater tracking range development, and uncrewed underwater vessels for environmental assessment and counter-mine warfare (1993-2022).
- Land testing of amphibious vehicles, surveillance vehicles, protected mobility vehicles, artillery, soldier combat ensemble, enhanced rifles, counter-sniper and counter-IED capabilities, battle-management systems, fire trucks and forward observer capabilities (2010-2014).

- Air testing of laser-guided weapons, anti-radiation weapons, anti-ship missiles, stand-off penetration weapons, air-to-air missiles, aircraft-based weapon management software, airborne navigation training, engine trend monitoring, aeroelastic flutter limits, digital avionics updates, new training helicopters and aircraft, identification friend and foe (IFF), and airborne surveillance radars (1993-2018).
- Space testing of new allied satellite capabilities and communication (2010-2014).
- Joint testing of operational planning software, electronic warfare capabilities, cybersecurity capabilities, datalinks, and special forces capabilities to respond to CBRNE events (2010-2023).

This diverse test exposure is unusual, as following the First Principles Review of 2015, the responsibilities for test conduct and much of the vetting of test plans reverted to each Service and Group through the acquisition lifecycle instead of passing the baton between agencies. The Vice-Chief Group took responsibility for test policy, competency and governance but not detailed test planning or conduct.

I was privileged to be Australia's lead negotiator for two years, securing the Multilateral Test and Evaluation Program (MTEP) MOU with the five-eyes partners (2013-14). I served for six months in 2009 with allied forces in Baghdad, Iraq helping plan the drawdown of forces. Identifying key differences between U.S. and Australian acquisition, testing, and evaluation has been greatly assisted by annual visits to or from the U.S. DOT&E and International T&E Association symposia every year from 2010 to 2023.

Since 2015, I have been a Senior Lecturer in Test and Evaluation with UNSW, researching and teaching test and evaluation planning and techniques. I have graduated over 1200 students in test subjects within the Master of Systems Engineering, Cybersecurity, Space Systems and Project Management programs. The opportunity to teach and research has helped highlight differences between my experience in the service and best practices in systems engineering, project management, cybersecurity, and testing and evaluation.

B. The Widening Gap with U.S. Defense in Assurance Testing

Test and evaluation are crucial for informed decision-making. Contrary to popular perception, when test and evaluations are planned and resourced correctly, it improves the speed of capability delivery by involving representative users and stakeholders directly in practical test activities that reduce bureaucracy, build consensus and disclose risk early [2, 14, 15]. However, to do that efficiently with modern technology, the complexity of interconnected and autonomous Defence capabilities and advanced persistent cyber threats requires acquirers and testers to be proficient in modelling, simulation, digital engineering, agile requirements methods, software development, cybersecurity and modern test design [7, 16-22].

Further, perspectives in the U.S. and the U.K. on Defence capability assurance have fundamentally changed compared to Australia. Public accountability has led to independent oversight of capability assurance mechanisms in those countries but not here. In the U.S., it is Title 10 congressional law founded in the so-called 'Pentagon Wars' which the proposed Australian Bill seeks, inter alia, to emulate. In the U.K., the Infrastructure and Projects Authority (IPA) is the Government's recent centre of expertise for infrastructure and major projects, including oversight of Defence Projects, reporting to the Cabinet Office and His Majesty's Treasury.² These legislative initiatives in capability assurance shift oversight on key disciplines in capability assurance from the 'toolbox' and discretion of program managers to legislated oversight agencies. In the simplest form, they implement public expectation that '*you shouldn't buy a car without a test drive*', or in Defence terms, by '*a fleet of cars with public monies without a test drive by representative users before contract and production.*' These agencies do an enormous amount in quality education, policy initiatives and 'best practice guidance. Australia is long overdue to follow its AUKUS leads with similar legislated initiatives and oversight, be that just Defence or more broadly.

A 2018 research article by Joiner and Tutty [2] explores the widening gap in capability assurance between the U.S. and Australian defence departments thoroughly. They find the following four main factors driving the need for better integration, interoperability and information (I3) assurance, expanded later in Gorod, et al. [23, Ch. 16] and supported by subsequent significant research by Ferreira, et al. [13] into the prevalence and effect of the modern system of systems phenomenon:

"First, the systems are becoming so synthesised or fused, complex and interdependent that they can, even without taking into account human agency, have emergent properties or exhibit behaviours that vary to an extent that is not easily predicted. Moreover, the number of permutations of modern software-intensive systems make classical rigorous

² <https://www.gov.uk/government/publications/infrastructure-and-projects-authority-mandate/infrastructure-and-projects-authority-mandate>; <https://www.gov.uk/government/news/ipa-2022-23-annual-report-on-major-projects>.

testing of them, all but impractical (Cofer, 2015), such that there has to be a reliance on some mission and safety-critical assurances, (Tutty, 2016).

- Second, as the software-intensive systems enable higher order human-like functions (i.e. strategies and decision-making not simply control), the difficulty in specifying what the system must do becomes harder and it is more crucial to include representative human agency and decision-making to adapt the systems during development.
- Third, the threat to weapon systems has adapted as a result of the push for information exploitation and dominance, so as to exploit the broader cyber-attack surfaces of such inter-connected systems, not just with malicious attacks but as part of multi-layered hybrid1 or hyper2 warfare – in short the threat is more complex and probably adaptive.
- Fourth, there is a requirements stasis during development and build of large complex systems, largely through an emphasis on project management cost and schedule achievement on contracts, and unfortunately including their processing and software. Such a requirements stasis soon creates an alternative reality that is too far out of alignment with the contemporary family-of-systems and strategic/operational reality into which that complex new system must go into service – more so today because of the three previous change factors.”

Joiner and Tutty [1] find and document six major U.S. Defense initiatives at the time, creating a widening gap with the Australian Defence that has largely not followed their ally:

“Initiative 1 – augmenting operational exercises with formal experimentation

The U.S. DoD has developed experimentation exercises in each key capability area where developing capabilities are deliberately networked and evaluated with legacy systems. Such experimentation exercises have annual battle-rhythms, dedicated planning staffs and evaluation scientists, such that project offices who look after updates of legacy systems and new developmental systems, only have to cover the cost and effort of sending or linking their systems and occasionally providing representative users. ...

Initiative 2 – integration system program offices and new certifications

The U.S. DoD have developed program offices for each of the services focused on integration. These integration offices form a nucleus of specialty integration staff to advise other program offices on the evolving architecture requirements and to acquire integrating capabilities. ...

Initiative 3 – enhanced T&E regime – earlier, evidence-based rigour and innovation: test smart not test often!

The most recent U.S. Director of Operational T&E (DOT&E), Dr Gilmore, was appointed in 2009 and until his departure this year, he oversaw a dramatic improvement in the rigour, timeliness and joint networking of test sites and distributed simulation. In particular, he rolled out mandatory use of probabilistic test design and test analysis techniques to all U.S. test centres, test staff and acquisition programs that are inherently and mathematically efficient.

...

Initiative 4 – T&E network infrastructure

The U.S. DoD test networks are now extensive, connecting every major design development facility and test range in the U.S., using three networks each with different levels of security and purpose; namely the Test Enabling Network Architecture, the Joint Mission Environment Test Capability network and the Joint Information Operations Range. The networks were developed by, and are run by, the Test Resource Management Center whose mission is to provide the necessary enterprise-wide architecture and the common software infrastructure ...

Initiative 5 – cybersecurity protection plans and T&E

The U.S. DoD response to cyber threats began in earnest in 2008 with a Presidential Directive and has since leveraged the DoD’s strengths in T&E and its T&E network infrastructure extremely well. Joiner (2017) outlines the three-phased approach taken by the U.S. DoD which has ended with a completely updated acquisition policy with cybersecurity integrated into all acquisition lifecycle stages.

Initiative 6 – permeating these U.S. initiatives into industry

The U.S. DoD initiatives outlined above have all permeated U.S. Defense Industry.”

Since the work above, AI-enabled systems are challenging assurance processes and competencies in both Defence departments [7, 24]. The U.S. are leveraging their investment in areas like modelling, simulation, software testing, test designs and test networks to meet the assurance challenge [25, 26], further widening the gap with Australia’s assurance competencies, policies and infrastructure. If this gap is not addressed, Australia will not have the agility to match near-peer threats and will be less trusted to adapt and integrate with its allies. This gap spans competencies, policies and infrastructure, which slows down and obfuscates most acquisitions.

C. Primary Australian Reason for the Widening Gap

There are two primary reasons for this widening gap, one Australian-based and the other U.S. based. The first reason is an Australian problem with stagnant and self-reinforcing policies and practices in acquisition, system

engineering, and testing and evaluation. Globally research and practices in these fields have moved dramatically since the 20th Century dogma on which the majority of Defence capability acquisition and assurance is practised. Defence largely internally promotes an aging acquisition and assurance workforce with self-consultation about competencies that do not embrace global practices, like agile requirements development [4, 12, 22, 27-31], software development and testing [5, 7, 32, 33], test design [1, 34], modelling and simulation [1, 6, 35, 36], cybersecurity [5, 8, 16, 37, 38], and digital engineering [11, 20, 39-47]. At this point in the submission, I will borrow from recent research into agile approaches by Kazakevich and Joiner [4], noting the Defence Strategic Review's repeated calls for Defence to seek '*minimum viable capability in the shortest possible time*' [27] which will require agile acquisition and testing through the concept of a minimum viable product. In comparing project success by methodology, Pace [48] summarise the disadvantages of waterfall approaches as follows, all of which resonate with other major Defence reviews [49, 50]:

Critics argue that this tactic is not appropriate when the specifications and requirements cannot be correctly collected at the project onset or are in a state of flux (Saynisch, 2010b). Also, due to the linear nature of the work, changes to requirements can require large amounts of rework or wasted work, which can be detrimental to the project regarding schedule and cost (Haughey, 2009). Another element of criticism is the amount of control required. The traditional approach takes the perspective that a rigorous, hierarchical control best manages complexity (Saynisch, 2010b), but critics assert that project problems stem from this framework (rather than from a lack of process or planning) (Parker et al., 2015). Finally, this traditional approach carries the perspective as bureaucratic in nature; the project completes large amounts of documentation throughout its lifecycle (Phatak, 2012).

Software engineers recognised that the development process could be improved by creating an Agile methodology that has evolved and widely adopted worldwide as a framework of choice for developing software-focused products [7]. Pace [48] describes the '*basis of agile project management methodology*' as '*a series of recurring iterations*' where '*each iteration cycle includes planning, design, coding, and testing.*' Again according to Pace (2019), '*The advantages of the agile approach are cost savings and speed of delivery (Stettina and Horz, 2014) ... a flexible method that embraces change (Stettina and Horz, 2014) ... right for any project that involves uncertainty, volatility, or risk*' and eliminating the '*bureaucratic overhead common to the traditional approach.*' Research notes that bureaucracies with directive control can become entrenched in waterfall, lifecycle, and contract approaches [48, 51]. Shifting these dogmatic practices cannot rely on individuals finding their own training and trying innovative approaches, partly due to the Defence bureaucracy of consensual committees, poor incentives and a one-team mindset [52, 53]. Instead, shifting this dogma requires a proactive assurance policy and directed re-education programs, with some significant test infrastructure support. The research by Smith, et al. [53] in counteracting harmful Defense incentives through testing and evaluation is particularly noteworthy, especially in retaining innovative staff.

Stagnant Australian policies are also caused by the fragmentation of the Defence Department across 13 services and groups, or 14 if you count the new Nuclear Propulsion Organisation. Defence has a consensus approach to central policy that can take years for even minor changes. Such fragmented groups can create innovative policies (the diversity argument), like the ICT Software Test Manual 2018 [33]. However, it is too easy for such a policy to apply only to capabilities within that group or, in the example's case, be focused only on ICT and not software in all capabilities due to hesitancy about who has authority. Another example is cybersecurity assessment, where the lead guidance has been in the same ICT management group using the Whole-of-Government Information Security Manual (ISM) as its authority for years. Information gathered in one area, such as a space satellite, can pass through numerous other groups and services, like datalinks, base infrastructure, command planning, etc., each with its own technical assurance and testing standards. Defence needs to be able to prescribe the minimum standards of cybersecurity assessment and testing best practices across all services and groups, including nuclear capabilities. Such policies must evolve quickly and to a higher standard than other Government departments due to Defence's generally higher threat risk [5, 8, 37].

One example of the paucity of Australian Defence response to ANAO auditing and the complexity of modern technology is the inadequacy in the minimum competency levels for managing tests and evaluations. Adding up the Defence training courses in test and evaluation, you get just four weeks of training, more appropriate for a Council roundabout than the programs Defence manages. Even the systems engineering master program I manage has one core introductory subject to test and evaluation at 160 hours of work or four weeks of full-time intensive. The second elective subject I run to help match U.S. DoD competencies in test design attracts less than five per cent of my students and extends our UNSW offerings to just two months full-time. A month or two of education is much less than the one year that has been the standard in overseas flight tests for the last four decades. While our Defence sends officers to do this one-year education, it is a handful and does not cover the many non-aircraft capabilities that

today have equal complexity to aircraft. Many of the subjects outlined in this submission as deficient in most of our capability assurance staff would fill the remaining year of complete education in this field.

D. Primary U.S. Reason for the Widening Gap

The second reason for the widening gap is the U.S. decision to create the Directorate of Operational Test and Evaluation (DOT&E), working directly for the Office of Secretary of Defense (i.e., Minister). Despite having 'Operational' in the title, under Title 10 Congressional Law, DOT&E licenses and approves all test and evaluation planning and appointments across the whole lifecycle, reports annually to Congress, and provides funding to cover significant test process, education and infrastructure improvements. This decision to give greater autonomy to testing and evaluation followed the 1980's issues known as the Pentagon Wars [54, 55]. This independence has enabled DOT&E to help its Defense respond more quickly to cybersecurity threats [16, 21] and use modern, efficient modelling, simulation and test designs in their acquisition [10, 35, 56, 57]. Importantly, their influence has created three passes of Government approval ('A' to 'C') compared to Australia's two (1 & 2), whereby contracted development can only proceed to low-rate initial production until an operational assessment is complete [58]. This additional milestone avoids over-commitment in costly production until proper testing of prototypes and the associated risk disclosure, generally with schedule and cost advantages [14, 59]. Australia demonstrated it could do this approach in the ANZAC Ship Missile Defence project (known as '1+7', as set by central agencies) and by the Chief of Army in the Joint Light Tactical Vehicle contracts.

U.S. agencies like the Test Resource Management Centre (TRMC) and their support from the Defence Acquisition University (DAU) mean the leading edge of test practices and infrastructure are directly funded, are quickly supported by policy for Defense use, and there is education and competency to use these. In contrast, in Australia, the responsible areas for testing and evaluation are chronically underfunded, have a minimal policy and test infrastructure, with very modest overall test competency requirements. These deficiencies are despite significant related ANAO reports in 2001-2002 [60], 2010-11 [61] and 2015 [52], Defence's Roadmap for Test and Evaluation in 2008 [62], and a Senate Inquiry into Defence Procurement in 2012 [50]. For example, at the peak of operational difficulties with the Collins in 1999, a Ministerial report expressed considerable frustration at the inadequacy of test and evaluation not to have screened for technical risks earlier, including a prescient call for preview testing as part of any future selection.³ The ANAO confirmed the inadequacy of the Collins testing in a more detailed review published in 2002, and they ascribed the difficulty to project governance not enforcing the necessary test policy [60]. Unfortunately, Defence at the time rejected the ANAO recommendations to improve along U.S. lines the independence and governance of such project testing, ironically arguing it did not do projects that warrant such rigorous oversight!⁴

The First Principles Review of 2015 [49] sought a strong strategic centre with the Vice-Chief, which ultimately included test and evaluation governance. The test and evaluation governance has a new 2021 strategic plan for test and evaluation [63] based on 2020 priorities in Government Sovereign Industrial Capability [64], but funding and progress in the two years since show it lacks the wherewithal to address the significant test policy, infrastructure and education changes needed.

Examples. Two examples illustrate the differences between Australia and the U.S. in undertaking capability assurance. First, in 2014, a Capability Gate Review Board considered the 'OneSky' project proposal in preparation for Government submission. Serious concerns were raised by me as the senior test and evaluation representative concerning the technical risk, whereby ninety per cent of systems required were commercial of the shelf, and only ten per cent of the capability would be developmental, mainly integrating software. None of the U.S. best practice methods for assessing the technological risk levels of mixed maturity systems (i.e., [9]) were used. As the test representative I argued for a low-cost pilot implementation in low-density airspace before committing to full development but was outvoted. Some of the harmful incentives noted by Smith, et al. [53] were evident in the discussion. The One-Sky project is now a '*black swan*' [65] or project of concern with over 200 per cent overrun in expenditure and schedule amounting to hundreds of millions of dollars [66, 67]. Had this proposal occurred in the U.S., the test perspective would have accompanied the submission to Government, formal caveats would have been put on the test and evaluation master plan, and independent oversight of the test progress would have been reported annually to Congress/Government.

³ <https://www.aspistrategist.org.au/wp-content/uploads/2016/04/Discussion-Paper.pdf>

⁴ RAND Corporation. (2011). Learning from Experience, Volume IV - Lessons from Australia's Collins Class Submarine Program. Santa Monica: RAND Corporation on behalf of Australian Department of Defence (available www.dtic.mil/dtic/tr/fulltext/u2/a552686.pdf), especially pp. xix, 12, 24-26, 41, 44-45, 50.

Further, the Milestone C checkpoint would have provided an off-ramp to the Government without the extent of escalating commitment. Other possible risks are likely due to the different cybersecurity assessment and test standards, test design, and software testing between the two countries. Air traffic systems are cyber-physical systems that warrant testing on cyber-ranges both in development and through life [68-70], infrastructure that is scarce in Australia. Another object lesson from this example is the benefit such an agency could have across complex capabilities needing to be assured in all Government agencies, whether directly or by positive example and guidebooks.

Another example is the development of test infrastructure to track underwater, surface and above-water military assets in the same range. Defence developed a project in 2009 to expand its aging deep-water tracking range to include a contiguous shallow-water tracking capability. Due to the physics of noise in the water, shallow-water tracking is more costly. Such capability is needed locally as underwater assets do not quickly move to U.S. ranges, and the nearest such U.S. ranges have limitations in depth. The assistance of Naval Underwater Warfare Command (NUWC), Code 70 (Ranges) was sought to scope the project in 2010-11. A decade of acquisition activities has seen requirements engineering, tendering, portable systems contemplated, transfer to the submarine project, transfer back to the Navy and many other complications. The major environmental and estate planning requirements, estimated to take a minimum of seven years to implement, obtained part of the decision a decade later (2020),⁵ with the bulk of the work currently still being tendered.⁶ I attribute this decade of delay to the competition such projects have from the funding pressures of the very capabilities they seek to support, mainly submarines and torpedoes but also maritime patrol aircraft, anti-submarine warfare ships, counter-sea-mine and smaller uncrewed underwater vessels (UUVs). I also see an unnecessary reinvention of allied expertise, whereby the simple Foreign Military Sales (FMS) agreement for NUWC Code 70 to oversee the capability is repeatedly argued away for lengthy commercial tendering.⁷ Such a capability in the U.S. is evident in the Atlantic Underwater Test and Evaluation Center (AUTECE), built from 2009-2019. While the USN was the key stakeholder, the U.S. DOT&E lobbied Congress on the importance of the funding, highlighted annually in reporting on the importance of related projects. The U.S. Test Resource Management Centre (TRMC) is also pivotal in securing and sustaining such capability (i.e., <https://www.dote.osd.mil/Portals/97/pub/reports/FY2022/FY22DOTEAnnualReport.pdf>, p. 25).

E. Addressing the Widening Gap

One more deleterious effect is caused by the widening gap between U.S. best practices and Australian Defence practices in capability assurance. Australian capability assurance staff who receive advanced education in test techniques, such as the few subjects I teach or the test pilots and engineers receiving overseas education, become dismayed when they try what they have learned in a Defence context. Initiatives launched in a Defence context without policy support and the necessary test infrastructure absorb enormous energy and burn some of our best and brightest, including those with the values and integrity to seek test roles. Again, the research by [53] is seminal in counteracting harmful incentives with testing and evaluation and would likely reduce the wastage of highly educated testers.

I fully support the formation of a more independent capability assurance agency. That agency must first determine and enforce the minimum standards to assure all Defence capability and then proactively pursue best-practice improvements that address technology and threat advancements with greater efficiency. Such an agency needs the imprimatur and funding mechanisms of some legislation to prescribe these new policies, curriculum and infrastructure in design assurance, systems engineering and testing. The record of Defence in the 20 years since the ANAO first suggested such an agency aligned to the U.S. model strongly suggests Defence's disagreement was a mistake. There has been a litany of concerns from reviews and difficult projects, highlighting the need to better oversee the majority of capability assurance at the expense of the few projects which excel. Alternatively, to put that constructively, to capture and prescribe the best practice of the few projects which excel and prescribe that across all services and groups.

⁵ <https://www.minister.defence.gov.au/media-releases/2020-11-20/new-maritime-underwater-tracking-range-be-built-wa>

⁶ Tender at <https://www.tenders.gov.au/Atm/Show/57595d52-1778-48c7-b989-7a5447b1b395> closing 17 Aug. Scope of works now being tendered "Overall, design services required will likely involve the following key activities: Planning Phase Activities: Concept Design Report (30% design); Site selection; P80 Detailed Business Case; risk workshops; and Public Works Committee Notification. Delivery Phase Activities, Schematic Design Report (50% design); Detailed Design Report (90% design); Head Contractor procurement comprising, Stage 1 – Invitation to Register Interest; and Stage 2 – Request for Tender; Final Design Report (100% design; for-construction documents); construction period; Defects Liability Period; lessons learnt workshop; and risk workshops."

⁷ <https://adbr.com.au/rft-issued-for-navy-underwater-instrumented-range/>

Eleven main areas need urgent action across the pillars of assurance policy, infrastructure and education. These areas include agile testing, software testing, integrated testing, cybersecurity assessment and testing, test designs, autonomous system tracking, modelling, and simulation. For example, urgent funding is needed for a joint network integration centre (JNIC) to be the heart of such assurance and test capability, federating Australia's test facilities into a Defence simulation network and cybersecurity-accredited cyber-ranges. These initiatives can be undertaken with the direct assistance of DOT&E under the multilateral T&E program (MTEP) MOU and FMS funding or through new AUKUS agreements. Priority would be to:

- 1) Have TRMC expand the U.S. test networks across all Australian test sites, including cybersecurity and simulation training and accreditation.
- 2) Align Australian Defence assurance competencies and education programs with the U.S. and their DAU, including re-education programs for all acquisition and test staff in agile software methods, cybersecurity assessment and test, modelling, simulation and digital engineering.
- 3) Align Australia's autonomous system tracking ranges and assurance laboratories with the U.S. Defense, particularly a primary contiguous deep and shallow underwater tracking with the direct support and oversight of Naval Underwater Warfare Centre, Code 70.

The advantage of such immediate U.S. assistance is that it will continue to close the gap and does not require the usual delay in defining and selecting a commercial support partner illustrated in my second example.

F. Recommendations

The following eleven recommendations are where Defence Capability Assurance has been deficient and widening compared to the U.S. Defense initiatives. These areas all need to be urgently improved. Progress over the past decades indicates that Defence is incapable of giving central priority to the necessary investment in policy, education and infrastructure to sustain informed decision-making at the level of technological and threat advancement. The level of investment sought is not to lead against this technological and threat advancement but to respectfully follow the U.S. Defense initiatives sufficient to remain a trusted partner. In Defence capabilities, over 150+ significant projects are spread across 13-14 groups and services without a strong strategic centre and funding impetus to consistently and proactively drive acquisition and assurance policy, education and infrastructure for these eleven areas.

- 1) **Agile Testing.** Agile acquisition practices require changes in T&E policy, procedures and education around the agile testing of minimum viable product (MVP) and agile requirements setting.
- 2) **Software Testing.** The increasing amount of software-driven functionality means all preview, developmental and acceptance T&E practitioners need competency in software test management, and the T&E policies and practices need to be updated to cover these aspects.
- 3) **Cybersecurity Assessment and Test.** Advanced persistent cyber threats require changes in T&E policy, procedures and education on assessing and testing the cyberworthiness of weapon systems, ICT and Defence infrastructure.
- 4) **Integration Testing and Simulation.** The system-of-system (SoS) phenomenon comes from the integration and interconnectivity of apparently mature, multi-proprietary, multi-generational systems or software to form new capabilities [13]. The SoS phenomenon requires changes in T&E policy, procedures and education to improve integration testing and supporting disciplines like modelling, simulation, bench-level pilots, combinatorial test design and test coverage.
- 5) **Test Design and Analysis Competence.** T&E competencies need to include test design and analysis (i.e., what runs, repetitions for what factors) so that rigour and efficiency can be sought through screening the effect of the essential factors, modelling these, and validating the adequacy of capability performance and delivered models to predict that performance.

- 6) **Joint Network Integration Centre.** A joint network integration centre (JNIC) servicing and linking all Services and Groups is an urgent and crucial organisation and infrastructure needed to enable T&E for all technical, human, and procedural integration and cybersecurity aspects.
- 7) **Cyber-physical Connected Cyber-range.** A federation of cyber-range test infrastructure and associated accreditation and training is urgently needed, leveraging extant allied capabilities to progressively roll-out in priority order across all Defence capabilities.
- 8) **Autonomous Systems Ranges.** Defence needs to urgently re-prioritise ranges to test robotic and autonomous systems in general and specifically address the decade of glacial progress on underwater, surface and above-surface tracking and signature ranges in deep and shallow water. The underwater tracking ranges should be well ahead of new submarines and autonomous underwater vessels, and to do so, these need to be managed for consistency and efficiency by U.S. Naval Underwater Warfare Command (NUWC), Code 70.
- 9) **Independent funding.** Improving T&E in Defence requires more centralised funding for infrastructure and education, including much greater direct use of allied capabilities like the U.S. Test Resource and Management Centre (TRMC) and U.S. Defense Acquisition University courses.
- 10) **Improving Contractors' T&E.** T&E governance in Defence needs to include and drive standards and practices in Defence Industry as much as it does T&E conducted inside Defence, including more direct oversight of competencies, infrastructure, contract procedures and deliverables.
- 11) **Proactive Policy Guidance.** T&E policies and practices need to be more adaptable and proactive to cater for new technology and threats, using guidebooks like the U.S. National Institute for Standards and Testing (NIST), funding for initiatives like the U.S. Test Resource and Management Centre (TRMC), and education like the U.S. Defense Acquisition University.

G. References

- [1] K. F. Joiner and M. G. Tutty, "A tale of two allied defence departments: new assurance initiatives for managing increasing system complexity, interconnectedness and vulnerability," *Australian Journal of Multi-Disciplinary Engineering*, vol. 14, no. 1, pp. 4-25, 2018.
- [2] R. Yesudas, K. Castelle, K. F. Joiner, J. Bradley, and M. Efatmaneshnik, "Solving tensions and impacts of overlapping project management and systems engineering with the elegance of the complex systems governance approach," *International Journal of System of Systems Engineering*, vol. 10, pp. 164 – 193, 2020.
- [3] P. F. Katina, A. Tolk, C. B. Keating, and K. F. Joiner, "Modeling and Simulation in Complex System Governance," *International Journal of System of Systems Engineering*, vol. 10, no. 3, pp. 262-292, 2020.
- [4] B. Kazakevich and K. F. Joiner, "Agile Approach to Accelerate Product Development using an MVP Framework " *Australian Journal of Multi-Disciplinary Engineering*, 2023: submitted.
- [5] (2023). *Australian Defence Guidebook for Test and Evaluation of Software-Intensive Systems, ICT and IS Projects (draft)*.
- [6] A. Tolk, P. Barry, and S. C. Doskey, "Using modeling and simulation and artificial intelligence to improve complex adaptive systems engineering," *International Journal of Modeling, Simulation, and Scientific Computing*, vol. 13, no. 2, pp. 2241004.1-19, 2022.
- [7] J. Weiss and D. Patt, "Software Defines Tactics: Structuring Military Software Acquisitions for Adaptability and Advantage in a Competitive Era," Hudson Institute, Online2022, Available: <https://www.hudson.org/national-security-defense/software-defines-tactics-structuring-military-software-acquisitions>.
- [8] GAO, "Weapon Systems Cybersecurity: DOD Just Beginning to Grapple with Scale of Vulnerabilities," United States Government Accountability Office, Washington2018, vol. GAO-19-128.
- [9] (2020). *GAO-20-48G, Assessment Guide Technology Readiness: Best Practices for Evaluating the Readiness of Technology for Use in Acquisition Programs and Projects*. Available: <https://www.gao.gov/assets/gao-20-48g.pdf>

- [10] E. Bjorkman, "Going Faster: Implications for Test and Evaluation," *ITEA Journal*, vol. 40, pp. 83–85, 2019.
- [11] E. B. Rogers and S. W. Mitchell, "MBSE Delivers Significant Return on Investment in Evolutionary Development of Complex SoS," *Systems Engineering*, vol. 24, no. 6, pp. 385–408, 2021.
- [12] K. M. Castelle, A. W. Dean, and C. B. Daniels, "Benefits and challenges of implementing agile development in modular shipbuilding," *Naval Engineers Journal*, vol. 131, no. 2, pp. 75-85, 2019.
- [13] F. H. Ferreira, E. Y. Nakagawa, and R. P. dos Santos, "Reliability in Software-intensive Systems: Challenges, Solutions, and Future Perspectives," in *47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, 2021, pp. 54-61: IEEE.
- [14] E. J. Copeland, T. H. Holzer, T. J. Eveleigh, and S. Sarkani, "The effects of system prototype demonstrations on weapon systems," *Defense ARJ*, vol. 22, no. 1, pp. 106–134, 2015.
- [15] K. F. Joiner, "How New Test and Evaluation Policy is Being Used to De-risk Project Approvals through Preview T&E," *ITEA Journal*, vol. 36, no. 4, pp. 288 – 296, 2015.
- [16] K. F. Joiner, "How Australia can catch up to U.S. cyber resilience by understanding that cyber survivability test and evaluation drives defense investment," *Information Security Journal: a Global Perspective*, vol. 26, pp. 74-84, 2017.
- [17] J. D. Hagar, T. L. Wissink, D. R. Kuhn, and R. N. Kacker, "Introducing combinatorial testing in a large organization," *Computer*, vol. 48, no. 4, pp. 64-72, 2015.
- [18] D. R. Kuhn, R. N. Kacker, L. Feldman, and G. White. (2016, May) Combinatorial Testing for Cybersecurity and Reliability. *Information Technology Bulletin*. Available: <https://www.nist.gov/publications/combinatorial-testing-cybersecurity-and-reliability>
- [19] E. Lanus, L. J. Freeman, D. R. Kuhn, and R. N. Kacker, "Combinatorial testing metrics for machine learning," in *IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, Online, 2021, pp. 81-84: IEEE.
- [20] A. Mukhopadhyay *et al.*, "A Perspective on the Adoption of Digital Engineering Within an Enterprise," Ansys, Online July 2023, issue Revision 1.1. Available: <https://www.ansys.com/content/dam/resource-center/white-paper/ansys-digital-engineering-in-the-enterprise.pdf>.
- [21] P. Christensen, "Cybersecurity Test and Evaluation: A Look Back, Some Lessons Learned, and a Look Forward!," *ITEA Journal*, vol. 38, no. 3, pp. 221–228, 2017.
- [22] D. Wickert, "Test in the Age of Agile: Rising to the Challenge of Agile Software Development," *ITEA Journal*, vol. 40, pp. 95-105, 2019.
- [23] A. Gorod, L. Hallo, V. Ireland, and I. Gunawan, A. Gorod, L. Hallo, V. Ireland, and I. Gunawan, Eds. *Evolving Toolbox for Complex Project Management*. Boca Raton, Florida: Auerbach Publications, 2019, p. 535.
- [24] J. T. Jacobsen and T. Liebetau, "Artificial intelligence and military superiority," in *Artificial Intelligence and International Conflict in Cyberspace*, vol. Routledge Studies in Conflict, Security and Technology, F. Cristiano, D. Broeders, and F. Delerue, Eds.: Routledge, 2023, pp. 135-156.
- [25] L. Freeman, "Best Practices for Addressing New Challenges in Testing and Evaluating Artificial Intelligence Enabled Systems," ed. <https://itea.org/professional-development/webinars/>: International Test and Evaluation Association, 2023.
- [26] T. Cody, E. Lanus, D. D. Doyle, and L. Freeman, "Systematic training and testing for machine learning using combinatorial interaction testing," in *IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, Online, 2022, pp. 102-109: IEEE.
- [27] (2023). *National Defence, Defence Strategic Review*. Available: <https://www.defence.gov.au/about/reviews-inquiries/defence-strategic-review>
- [28] S. Hooda, *Agile software development : trends, challenges and applications*. Hoboken, New Jersey: John Wiley & Sons, 2023.
- [29] M. V. Leticia, "Project Agile Management For Software Development: A Comparative Study On The Applicability Of Scrum Together With Pmbok And / Or Prince2," *Revista De Gestão e Projetos*, vol. 7, no. 3, pp. 48-60, 2016.
- [30] M. Lotfi, "Which practices are lean, agile and resilient? Literature review and practitioners' perspective," *International Journal of Advanced Operations Management*, vol. 11, pp. 142-170, 2019.
- [31] T. Thesing, C. Feldmann, and M. Burchardt, "Agile versus waterfall project management: decision model for selecting the appropriate approach to a project," *Procedia Computer Science*, vol. 181, pp. 746-756, 2021.

- [32] (2018). *Defense Innovation Board Do's and Don'ts for Software (draft)*. Available: https://media.defense.gov/2018/Oct/09/2002049593/-1/-1/0/DIB_DOS_DONTS_SOFTWARE_2018.10.05.PDF
- [33] Australian Government. (2019). *ICT Software Testing Manual*.
- [34] Department of Defense, U.S. (2017). *STAT COE-Report-01-2017 Automated Software Testing Implementation Guide*. Available: https://www.afit.edu/stat/statcoe_files/Automated%20Software%20Testing%20Implementation%20Guide.pdf
- [35] M. G. Lilienthal, "A look back at the past 44 years of live virtual and constructive (LVC) simulation and lessons for cyberspace LVC," *Journal of Defense Modeling and Simulation*, vol. 154851292211095, 2022.
- [36] G. Schweiger, H. Nilsson, J. Schoegg, W. Birk, and A. Posch, "Modeling and simulation of large-scale systems: A systematic comparison of modeling paradigms," *Applied Mathematics and Computation*, vol. 365, no. 124713, pp. 1-14, 2020.
- [37] (2020). *Cybersecurity Test and Evaluation Guidebook*. Available: <https://www.dau.edu/cop/test/DAU%20Sponsored%20Documents/Cybersecurity-Test-and-Evaluation-Guidebook-Version2-change-1.pdf>
- [38] (2018). *Defense Acquisition Guidebook, Chapter 8 on Test and Evaluation*. Available: <https://www.dau.edu/pdfviewer/Source/Guidebooks/DAG/DAG-CH-8-Test-and-Evaluation.pdf>
- [39] M. Price, "Current and Emerging trends in the Aerospace sector: How shifting priorities and developing technologies are shaping the industry today and into the future," Atkins Corporation, SNC Lavalin Group, Online July 2018, Available: <https://www.atkinsglobal.com/~media/Files/A/Atkins-Corporate/aviation-trends-white-paper-digital.pdf>.
- [40] T. Y. Pang, J. D. Pelaez Restrepo, C.-T. Cheng, A. Yasin, H. Lim, and M. Miletic, "Developing a digital twin and digital thread framework for an 'industry 4.0' shipyard," *Applied Sciences*, vol. 11, no. 3, pp. 1-23, 2021.
- [41] A. Rasheed, O. San, and T. Kvamsdal, "Digital twin: Values, challenges and enablers from a modeling perspective," *IEEE Access*, vol. 8, pp. 21980-22012, 2020.
- [42] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital twin: Enabling technologies, challenges and open research," *IEEE access*, vol. 8, pp. 108952-108971, 2020.
- [43] F. Flammini, "Digital twins as run-time predictive models for the resilience of cyber-physical systems: a conceptual framework," *Phil. Trans. R. Soc. A*, vol. 379, p. 20200369, 2021.
- [44] E. VanDerHorn and S. Mahadevan, "Digital Twin: Generalization, characterization and implementation," *Decision Support Systems*, vol. 145, p. 113524, 2021.
- [45] D. Henriques, R. F. Pereira, R. Almeida, and M. Mira da Silva, "IT governance enablers in relation to IoT implementation: a systematic literature review," *Digital Policy, Regulation and Governance*, vol. 22, pp. 32-49, 2020.
- [46] S. A. Hajkiewicz *et al.*, "Tomorrow's digitally enabled workforce: Megatrends and scenarios for jobs and employment in Australia over the coming twenty years," in "Australian Policy Online," Commonwealth Scientific and Industrial Research Organisation (CSIRO) 2016, Available: https://www.researchgate.net/profile/Claire_Mason/publication/299953345_Tomorrow's_digitally_enabled_workforce_Megatrends_and_scenarios_for_jobs_and_employment_in_Australia_over_the_coming_twenty_years/links/570728ee08aefb22b0934be9.pdf.
- [47] A. Miceli, B. Hagen, M. P. Riccardi, F. Sotti, and D. Settembre-Blundo, "Thriving, not just surviving in changing times: How sustainability, agility and digitalization intertwine with organizational resilience," *Sustainability (Basel, Switzerland)*, vol. 13, no. 4, 2021.
- [48] M. Pace, "A Correlational Study on Project Management Methodology and Project Success," *Journal of Engineering, Project, and Production Management*, vol. 9, no. 2, pp. 56-65, 2019.
- [49] D. Peever, R. Hill, P. Leahy, J. McDowell, and L. Tanner, "First principles review: Creating one defence," *Commonwealth of Australia: Canberra*, 2015.
- [50] Australian Senate, "Senate inquiry into defence procurement," *Canberra: Australian Parliament House*, 2012.
- [51] L. Mahadevan, W. J. Kettinger, and T. O. Meservy, "Running on hybrid: Control changes when introducing an agile methodology in a traditional 'waterfall' system development environment," *Communications of the Association for Information Systems*, vol. 36, pp. 77-103, 2015.
- [52] (2015). *ANAO Report No 9: Test and Evaluation of Major Defence Equipment Acquisitions*.

- [53] N. Smith, E. White, J. Ritschel, and A. Thal, "Counteracting harmful incentives in DoD acquisition through test and evaluation and oversight," *ITEA Journal*, vol. 37, pp. 218-226, 2016.
- [54] J. G. Burton, *The Pentagon wars: Reformers challenge the old guard*. Naval Institute Press, 2014.
- [55] R. Benjamin, "The Pentagon Wars," ed. United States: HBO, 1998, p. 1 h 43 min.
- [56] J. N. Elele, D. H. Hall, M. E. Davis, D. Turner, A. Faird, and J. Madry, "M&S Requirements and VV&a Requirements: What's the Relationship?," *ITEA Journal*, vol. 37, pp. 333-341, 2016.
- [57] D. K. Ahner, "Better buying power, developmental testing, and scientific test and analysis techniques," *ITEA Journal*, vol. 37, pp. 286-290, 2016.
- [58] (2019). *GAO-19-439, DOD acquisition reform: Leadership attention needed to effectively implement changes to acquisition oversight*. Available: <https://www.gao.gov/assets/gao-19-439.pdf>.
- [59] W. F. Kramer, M. Sahinoglu, and D. Ang, "Increase return on investment of software development life cycle by managing the risk - A case study," *Defense ARJ*, vol. 22, no. 2, pp. 174-191, 2015.
- [60] (2002). *ANAO Audit Report No. 30: Test and Evaluation of Major Defence Equipment Acquisitions*.
- [61] (2011). *Audit Report No.57: Acceptance into Service of Navy Capability*.
- [62] (2008). *Defence Test and Evaluation Roadmap*.
- [63] (2021). *Defence Test and Evaluation (T&E) Strategy*. Available: <https://www.defence.gov.au/about/strategic-planning/defence-test-and-evaluation-strategy>
- [64] (2020). *Sovereign Industrial Capability Priority Implementation Plan*. Available: <https://www.defence.gov.au/about/strategic-planning/defence-test-and-evaluation-strategy>
- [65] N. N. Taleb, *The black swan: The impact of the highly improbable (Vol. 2)*. Random house, 2007.
- [66] (2019). *Auditor-General Report No.4 2019-20, OneSKY: Contractual Arrangements*. Available: https://www.anao.gov.au/sites/default/files/Auditor-General_Report_2019-2020_4.pdf
- [67] J. Bajkowski, "Defence sounds new warning on \$4.1 billion national air traffic control system," in *The Mandarin*, ed. Online: Tom Burton, Private Media, 2022.
- [68] (2021). *Cyber-Physical Systems*. Available: <https://www.nist.gov/el/cyber-physical-systems>
- [69] J. Geismann and E. Bodden, "A systematic literature review of model- driven security engineering for cyber-physical systems," *The Journal of systems and software*, vol. 169, p. 110697, 2020.
- [70] B. Carter *et al.*, "A preliminary design-phase security methodology for cyber-physical systems," *Systems (Basel)*, vol. 7, no. 2, p. 21, 2019.