

Chapter 6

Test and evaluation

The fundamental purpose of test and evaluation (T&E) in a defence system's development and acquisition program is to identify the areas of risk to be reduced or eliminated. During the early phases of development, T&E is conducted to demonstrate the feasibility of conceptual approaches, evaluate design risk, identify design alternatives, compare and analyse trade-offs, and estimate satisfaction of operational requirements. As a system undergoes design and development, the iterative process of testing moves gradually from development test and evaluation (DT&E), which is concerned chiefly with attainment of engineering design goals, to operational test and evaluation (OT&E), which focuses on questions of operational effectiveness, suitability and survivability. Although there are usually separate development and operational test events, DT&E and OT&E are not necessarily serial phases in the evolution of a weapon system development. Combined or concurrent development test and operational test is encouraged when appropriate (possible cost/time savings).

Test and evaluation ... provides information to many customers. The T&E gives information to: developers for identifying and resolving technical difficulties; decision makers responsible for procuring a new system and for the best use of limited resources; and to operational users for refining requirements and supporting development of effective tactics, doctrine and procedures.¹

6.1 From the outset of this inquiry, and when trying to come to grips with the series of acquisition failures that were the inquiry's genesis, one question has nagged persistently at the Committee: *why was it not possible for problems to have been detected in a timely manner during the development and acquisition phases of projects?* The Committee is assuming here that the members of project teams were not simply negligent, and ignored problems, or suppressed them, or glossed over them in the hope that they would go away.

6.2 The question, in the Committee's view, demands a systemic answer. Given that the capability development/acquisition process is an example of 'systems engineering', the only answer that seems tenable is that projects did not have in place effective feedback mechanisms for managing risk.

6.3 The Committee is completely in accord with the remarks quoted at the outset of this section that the 'fundamental purpose of test and evaluation ... is to identify the areas of risk to be reduced or eliminated'. The Committee therefore sought to establish

1 Defense Acquisition University, *Test and Evaluation Management Guide* (4th edition), The Defense Acquisition University Press, Virginia, 2001, p. 2-1

whether the policies, structures and processes to achieve this ‘fundamental purpose’ had received sufficient attention by Defence and have been attended to by DMO as part of the necessary reforms.

6.4 Overall, the Committee is not satisfied that this is the case. Indeed the Committee is very concerned by what appear to be systemic, and potentially serious, shortfalls in test and evaluation across the whole capability life cycle.

6.5 The *Capability Systems Life Cycle Management Manual 2002*, with characteristic thoroughness, sets out in a clear and relatively comprehensive way the function of T&E in capability procurement. It recognizes that T&E

... begins in the Requirements Phase when a concept is established for progressive evaluation of the capability to ensure that it meets the approved requirement. This must include provision for corrective action in the event that the requirement is not being met.... In that an evaluation indicates non-compliance, corrective action should be initiated as early as possible in the life cycle.²

6.6 The *Manual* goes on to specify the requirement for ‘more sophisticated’ testing as the configuration becomes ‘better defined’; for a series of ‘formal tests’ during the Acquisition Phase which include technical, environmental and human dimensions; the tests should be ‘planned... scheduled... integrated’.

6.7 Having set down the parameters for Development T&E and Operational T&E, the *Manual* arrives at the requirement for a Test and Evaluation Master Plan, and emphasizes:

Key milestones during the Acquisition Phase should be related to T&E results. Measured progress toward achievement of the required capability baseline and retirement of risk should be a determinant of a project schedule.³

6.8 All of this is entirely consistent with the kinds of advice provided by standard texts on test and evaluation, and parallels the processes described in copious detail in the *Test and Evaluation Management Guide* published by the Defense Acquisition University for the US Department of Defense. Test and evaluation is also given substantial attention in other Australian Defence guides, such as the *Technical Regulation of Army Materiel Manual (TRAMM)*.

6.9 But the Committee has a strong sense that this abundant and sound advice is very poorly translated into the actual practice of project teams and others responsible for capability. What is worrying about this situation is that the advice of the *Capability Systems Life Cycle Management Manual 2002* does not specify any truly ‘new’ processes. The new *Manual* essentially underscores the long-established need

2 *Capability Systems Life Cycle Management Manual 2002*, paras 6.16 and 6.18

3 *Capability Systems Life Cycle Management Manual 2002*, para 6.31

for T&E to be properly integrated (planned and resourced) in the capability life cycle. The *Manual*'s processes draw on established Defence systems engineering practice in terms of key T&E-related documents and processes.⁴ In short, the systemic application of T&E seems to have been mandated in Defence guides and manuals for some time.

6.10 If this is in fact the case, Defence has a serious problem on its hands. It seems to the Committee that these mandated requirements have not been effective.

6.11 A simple indication of the fundamental importance of T&E—and the seriousness with which it is to be attended to—can be found in the above-mentioned Army *TRAMM*, which requires that 'All materiel involved in T&E is to have its details clearly recorded for identification and traceability' and 'All records relating to T&E activities are to be retained in a form subject to the Commonwealth Archives Act 1983.'

6.12 The Committee's concerns about T&E in capability development and materiel procurement received its initial boost from the submission to the inquiry by the ANAO, and from the ANAO's Report no. 30 2001–02 on *Test and Evaluation of Major Defence Acquisitions*.

6.13 The ANAO's Report No. 30 concluded that:

... there was little evidence of effective corporate initiatives to implement Defence's test and evaluation (T&E) policy, which aims to promote a unified approach to T&E to guarantee effective and efficient use of all T&E resources and avoid unnecessary duplication of effort. The policy needs to be reviewed and to articulate how the 'unified approach' is to be implemented.⁵

6.14 The Committee was concerned that some of the key recommendations of that report had been rejected by Defence. Defence's disagreements related to strategic management and oversight of T&E and the training of personnel responsible for safety critical system development, maintenance and test and evaluation.⁶ The Committee asked the USDM why Defence had dismissed the ANAO's recommendations.

I think that 'dismiss' is probably not the word I would choose. We thought very carefully about what the Audit Office had to say on this... Firstly, it was a single-issue audit. They looked at the role of T&E on its own and tried to give it a status in their recommendations and in our project management approach that put it on a pedestal almost on its own, without it being integrated into the total collection of tools that we use for project management. It does not stand on its own, in our view. We have a

4 Advice to the Committee in correspondence from ANAO, dated 27 February 2003

5 *Submission 2*, p. 4 (ANAO)

6 Australian National Audit Office, Report No. 30 2001–02, *Test and Evaluation of Major Defence Acquisitions*, p. 21

fundamental disagreement with Audit on that—not a dismissal, but we disagree that it should be out there on its own. We disagree with the need to provide, for example, a specific isolated budget for test and evaluation.

Audit also proposed a role for test and evaluation that went down the American route, where the test and evaluation authority was responsible for independent reporting—I think it may have been to Congress. We did not agree with a separate role for it—again, because we believe that it should be integrated as part of our project management techniques...

The audit, in some respects, was quite extreme. The Audit Office believed—and I think I saw a suggestion to this effect in one of the early drafts—that the proper test and evaluation procedures would have, in fact, prevented the problems with Collins and would have avoided the difficulties we face with Bushranger. I disagree with that. That is putting test and evaluation on a pedestal far above where it should be. The problems that we had with those projects were not related to not knowing what the problems were. It was not a case that these platforms were not tested and that no-one knew what the problems were. We knew what the problems were, but it was a question of whether people were doing anything about them to fix them.⁷

6.15 The Committee is somewhat surprised by this response on at least two counts.

- The Committee has examined the ANAO's Report No. 30, and contrary to Mr Roche's view that ANAO 'put [T&E] on a pedestal almost on its own, without it being integrated into the total collection of tools that we use for project management', the Committee considers that the ANAO report clearly recognizes that T&E is part of the systems engineering and management processes. Indeed the relationships were explored at several points in the ANAO report.⁸
- Mr Roche says 'It was not a case that these platforms were not tested and that no-one knew what the problems were. We knew what the problems were, but it was a question of whether people were doing anything about them to fix them.' This indicates project management problems that go beyond T&E. The Committee would argue that perhaps a more strategically placed T&E function—one which is captured, say, in the governance structures of Defence—would result in closer scrutiny of, and better accountability for, the implementation of prescribed T&E procedures.

6.16 In this context, the Committee notes the view of former Defence Secretary Dr Allan Hawke, expressed in a communication with a defence industry firm in July 2001:

T&E is an important tool in our plans for the management of Defence capability to ensure successful achievement and maintenance of operational effectiveness.

7 *Committee Hansard*, pp. 347–348 (Mr Michael Roche)

8 For example, at pp. 27–30, 42–46, 54–58 and 110–112

As Defence moves to consider its governance strategies in a theme of organisational renewal it is timely for us to consider T&E as a key management tool. I am aware of the proposal to review the governance structure for T&E and the proposal is currently under consideration by members of the senior leadership team in Defence.⁹

6.17 It seems to the Committee that the ideas flagged by Dr Hawke have not come to fruition. If a unified T&E approach is what risk management requires, then it seems to the Committee that some kind of feedback loop which is wholly integrated into the governance of Defence and the management of acquisition and logistics is exactly what is needed. Giving T&E status at this level would encourage and support the iterative, integrated, ongoing T&E being conducted (supposedly) throughout all phases of the capability life cycle.

6.18 The Committee felt generally uneasy about the evidence given by Defence witnesses in response to the Committee's questions about test and evaluation. Responses seemed to be sometimes contradictory, occasionally equivocal and frequently confusing. As a consequence, test and evaluation is not something about which the Committee has full confidence in Defence's capacity or will to seriously address.

6.19 Even at the level of clarifying what was generally understood by the phrase Test and Evaluation, the Committee struggled to establish common ground with Defence witnesses.

6.20 For example, one witness declared that the process known as IV&V (Independent Verification and Validation) was: 'the process within the systems engineering arena which goes beyond test and evaluation and which is far more important when dealing with software systems, particularly software systems that are safety critical.' When challenged about this statement, the witnesses reiterated that: 'Test and evaluation is a key component of verification and validation, which is a part of the systems engineering process.'¹⁰

6.21 The Committee is puzzled by this elevation of IV&V to overarching status with T&E falling under it. Key standard texts and guides consulted by the Committee placed T&E as the generic concept, with IV&V as a sub-class of T&E, and then one which was applied almost exclusively to software development projects. This certainly is the case in Blanchard's *System Engineering Management*, and the US Defense Acquisition University's *Test and Evaluation Management Guide*. The Australian *Technical Regulation of Army Materiel Manual (TRAMM)* also has IV&V as an annex to T&E. Nor does IV&V appear as a discrete entity in Defence's *Capability Systems Life Cycle Management Guide*.

9 Private correspondence made available to the Committee by the recipient

10 *Committee Hansard*, pp. 348–349 (Ms Shireane McKinnie)

6.22 The argument here is not a semantic one. Rather the Committee is highlighting the perplexity it invariably encountered when seeking to engage Defence witnesses on T&E matters.

6.23 For example, the Committee asked Defence witnesses about reports it had received about certain vessels entering service without T&E having been fully completed, and without the documentation necessary to enable this to happen.¹¹

6.24 There followed a series¹² of convoluted exchanges and disagreements about whether the vessels were actually in service or were not in service; about whether the Chief of Navy was doing a ‘review’ of Acceptance Into Naval Service procedures or had simply ordered a new direction for that to proceed; about the different responsibilities of the DMO and the Service chief with respect to confirming that contractual and functional requirements had been met as opposed to determining whether the system delivered fulfilled current service requirements; and whether or not relevant T&E data had actually been fully documented for some major platforms.

6.25 The Committee believes it an urgent task to have these matters clarified. The distinction between DMO’s responsibilities and those of the Service customer when it comes to T&E must be made clear. The point for the Committee is not that any of the platforms going into service are provably unsafe, but rather that the platforms seem not to be as assuredly safe as they are required to be according to Defence’s own guides—guides such as DEF (AUST)5679 dealing with the procurement of computer-based safety critical systems, or Navy’s delivery and acceptance document TI 338.

6.26 In response to a question on notice about the provision of T&E documentation (see 6.22 above), Defence stated that ‘all documentation related to pre-delivery tests and trials [for the vessels specified] has been made available to Navy to support its operational test and evaluation phase, culminating in the Acceptance into Naval Service milestone’. It is not clear whether that documentation was adequate, nor its delivery timely. The Committee assumes that it was. But a close examination of the table provided raises more questions than it gives answers.

6.27 For example, for each of the specified platforms, the Overall Test and Evaluation Assessment is stated to be ‘Transient’.¹³ The Committee assumes that this means that all the platforms are yet to be found ‘Compliant’, and presumably therefore not yet suitable for formal Acceptance into Naval Service. But in any event, with respect to what are these vessels being assessed as ‘Transient’ (or ‘Compliant’ or ‘Non-compliant’)?

6.28 The Committee notes that Test & Evaluation Master Plans (TEMPs) are not a new requirement. But where they exist, do they specify testing outcomes required to achieve Provisional Acceptance, or full Acceptance into Service? The Committee also

11 *Committee Hansard*, p. 352 (Senator the Hon Peter Cook)

12 *Committee Hansard*, pp. 352–358

13 The options were Compliant, Transient, Non-compliant.

notes that the Reports of Materiel State at Delivery (TI 338) are listed for each platform. But do these TI 338's record required data in terms of safety case, contractual liabilities, work to be completed, the status of the regulatory/certification system, Integrated Logistic Support, and performance limitations? Only a much more detailed examination of T&E processes would enable an assessment of their adequacy.

6.29 The Committee's view at this stage¹⁴, and pending the outcomes of the work of DTRIALS in developing a new T&E concept document (see below), is that:

(a) Operational Test and Evaluation (OT&E) should be planned and scheduled from the outset of a project and should be conducted by the relevant Service's OT&E organization, independent of the DMO and the contractors;

(b) OT&E should progress from a well-established foundation of Development Test and Evaluation (DT&E); and

(c) VCDF (and the Owner Support Executive) should be advised of the extent to which equipment delivered by DMO to the Services has met the functional requirements which had been originally specified in the documents that secured government approval for the project.

6.30 At one stage during the rather confusing evidence referred to above, there was discussion about the Navy receiving vessels from the DMO, then Navy seeing how far they could 'push the operational envelope', with 'the humans in the loop and asking, "How does this work as a bit of operational kit?"'.¹⁵ The Committee remains uneasy with that discussion because of the Committee's concerns about the seeming inadequacy of implementation of T&E during the development and acquisition phases. Operating 'with humans in the loop' should focus the minds of Service personnel on T&E issues with particular alacrity.

6.31 The Committee may appear over cautious in these matters, but insists that integrated and effective T&E throughout the capability life cycle is the only way to ensure the delivery of a fully functioning platform with safety-critical systems operating at peak efficiency and effectiveness. The recent 'sea water pipe burst incident' of the HMAS *Dechaineux*, for example, has reinforced to the Committee the imperatives of optimal T&E (both developmental and operational)—whether or not the pipe failure in the case of *Dechaineux* can be attributed to inadequate T&E. The Committee notes that the 1998 ANAO Report on the *New Submarine Project* included several pages devoted to stainless steel pipe risks which included the conclusion that the 'SMO pipe welding case study provides an example of uneven and unsystematic risk management.'¹⁶

14 On the basis of advice from the ANAO in communications with the Committee.

15 *Committee Hansard*, pp. 352–356 (Mr Michael Roche, Vice Admiral Chris Ritchie)

16 ANAO Report No. 34 1997–98, *New Submarine Project*, p. 93

6.32 While the Committee has not engaged in a thorough examination of risk management and T&E matters, it nevertheless asserts a right to express concerns when it feels that its exploration of test and evaluation matters elicits muddled responses from Defence witnesses and produces no convincing evidence that T&E has a profile in Defence commensurate with its fundamental importance in systems engineering generally and in weapons platform development in particular.

6.33 The Committee is pleased to note that there are some resources being devoted to a review of T&E policy within Defence.

Following the 2001 ANAO report, VCDF directed DTRIALS to initiate a review and redevelopment of Defence's T&E policy... Throughout 2002, DTRIALS has been developing a T&E concept paper and working to achieve alignment between the T&E policy and the guidance provided in the Defence Capability Systems Lifecycle Management [Manual]... and to address the ANAO report's recommendations.

This proposed T&E concept paper has been circulated and it has been agreed in principle as a conceptual basis for development of defence T&E policy. Importantly, it... seeks to ensure that the Operational Acceptance is validated against the Operational Concept Document—which is a new document proposed by the manual—so we have a chance to carry out test and evaluation against what has been delivered and against what was proposed in the Operational Concept Document. In addition to that, we have also recommended that a business case closure procedure be carried out, whereby the capability is measured against the original business case. That would then identify whether the field capability has met all the requirements. I am looking to within the next six months to achieve that.¹⁷

6.34 The Committee looks forward to examining a copy of the revised T&E policy when it has been finalized. The Committee will be particularly keen to ensure that the policy provides for T&E which is fully *integrated* (planned and funded) with the capability development process, that it provides for T&E to be carried out in an *independent* fashion, and that the policy embeds a 'cradle to grave' philosophy.

Recommendation

6.35 The Committee recommends that the Senate, under Standing Order 164, order the production, upon its completion, of the report by Director of Trials (DTRIALS) of the Review of Test and Evaluation in Defence, and that the Senate refer the document to the Senate Foreign Affairs, Defence and Trade References Committee for examination and report.

17 *Committee Hansard*, pp. 352–358 (Group Captain Michael Gaspert)