CRM D0012661.A1 / Final
July 2005

Analysis of T&E Spending and the Cost Impact of Recommendations for Streamlining the T&E Process

Donald Birchler • Jino Choi • Viki Johnson
George Kinoshita • Pete Kusek • Julianne Nelson
Richard Sperling

CNA

4825 Mark Center Drive • Alexandria, Virginia 22311-1850

Approved for distribution: July 2005

Jino Choi, Director
Cost and Acquisition Team
Resource Analysis Division

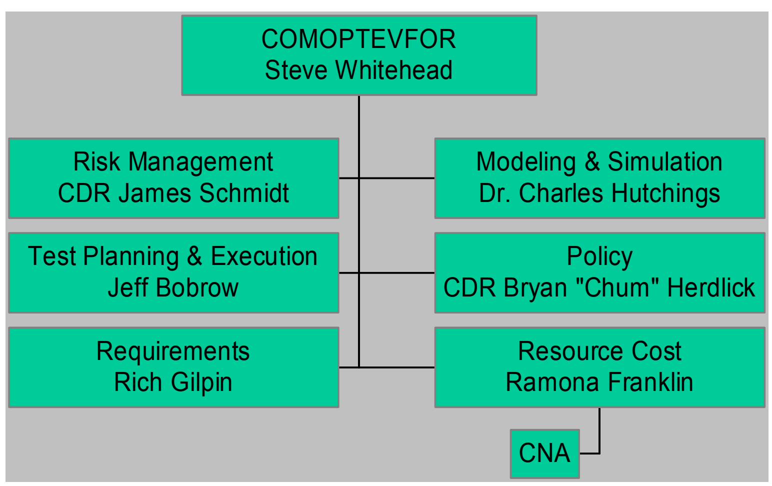
Contents

This page intentionally left blank.

Executive summary

The CNO's 2004 guidance set a goal of a 20-percent reduction in Test and Evaluation (T&E) costs. To determine ways to streamline the T&E process, Commander Operational Test and Evaluation Force (COTF) led a study to develop recommendations that ultimately would lead to cost savings. COTF, with support from the Assistant Secretary of the Navy (Research, Development & Acquisition) and the Systems Commands, established several working groups to examine different areas of T&E and offer recommendations that would reduce T&E costs. To support this effort, the Deputy Assistant Secretary of the Navy (RDT&E) asked CNA to help the Resource Cost Working Group (RCWG). Figure 1 identifies the participants in this project.

Figure 1. Working group organization chart



CNA collected financial data from 18 program offices on T&E spending from FY00 to FY04. We examined the data to determine spending trends and patterns. In addition, we estimated the cost impact of 4 of the 22 recommendations made by the other working groups. In summary, we found the following:

1. T&E spending occurs mainly in the middle phases of a program and is largely driven by systems engineering issues.
2. T&E is very labor intensive. Various forms of labor involved in T&E, such as direct labor hours, requirements development, and analysis, account for a large portion of T&E.
3. Infrastructure costs associated with T&E are also large.
4. Navy accounting systems are not adequate for tracking T&E expenditures. This problem is exacerbated by the fact that programs use various types of funds (T&E, procurement, OMN, etc.) from different program elements to pay for T&E.
5. Current T&E budget exhibits do not accurately reflect how much is really budgeted for T&E.
6. Adopting more modeling and simulation could lead to savings in T&E, especially in ship shock trials.
7. Eliminating difficulties in planning could lead to more integrated testing and "piggybacking" on fleet activities, which could yield savings.
8. Robust testing earlier in the program's development phase could save money by discovering problems early and allowing for less expensive changes.
9. Improving the process for the generation and refinement of requirements/capabilities to ensure testability could also yield some savings.

Introduction

Tasks and time lines

The CNO's 2004 guidance set a goal of reducing Test and Evaluation (T&E) costs by 20 percent. Commander Operational Test and Evaluation Force (COTF), with support from the Assistant Secretary of the Navy (Research, Development & Acquisition) and the Systems Commands, led an effort to examine what possible changes to T&E policy and practices could be made to streamline the T&E process.

Six working groups were established to develop recommendations for reducing T&E costs. To validate their recommendations, the working groups put together a qualitative survey for the 26 programs.

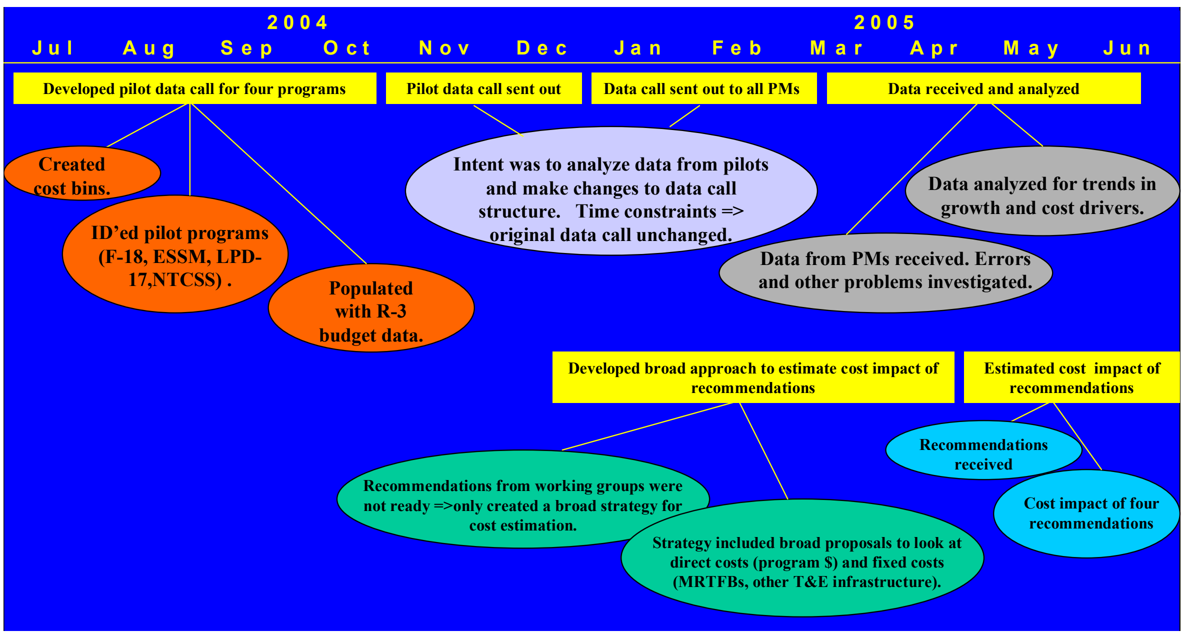
The Resource Cost Working Group had a slightly different role. In addition to developing recommendations and contributing to the creation of the survey, the RCWG was tasked with the following three assignments:

1. Develop a financial data call to collect T&E expenditure data from the 26 programs included in this study.
2. Collect and analyze financial data to identify trends and cost drivers.
3. Estimate the cost impact of the recommendations made by the working groups.

To support this effort, the Deputy Assistant Secretary of the Navy (RDT&E) asked CNA to help the RCWG complete the above tasks.

The major purpose in collecting the financial data was to determine just how much the Navy spent on T&E. These baseline numbers were used in two ways. First, we examined spending to see what trends there might have been. In particular, we were interested in seeing where most T&E spending occurs. While programs spend a great deal on actual T&E events, such as live fires and shock trials, it was thought that a great deal of money was being spent on T&E earlier in the developmental phase. In addition, as will be seen later, the working groups created a series of cost subcategories that capture, to a large extent, what programs spend their money on (targets, consumables, labor, etc.). These data were used to examine the major cost drivers for T&E. The second way the financial data were used was as a baseline from which we could calculate a percentage savings for each recommendation. Figure 2 shows the time line of the study.

Figure 2. Study time line



CNA began work on developing the data call in July 2004. The first order of business in creating the data call was to derive a set of cost subcategories that could be used by the programs to bin their expenditures. This was not a trivial task. As we have mentioned, there is no common accounting system across all the programs. Each one manages its funds and accounts for actual expenditures differently. After several meetings with various members of the working groups, a common set of generic cost bins was agreed on from which a pilot data call was created for four programs.

At the end of October, the first four data calls were sent to the pilot programs. The intent was to analyze the data received and gather suggestions from the four programs to make any changes to the data call format that they felt would facilitate the task of collecting financial data. However, the pilot programs did not respond in time for us to consider their suggestions. By January, the decision was made to continue with the format as it was, and the rest of the data calls were sent out to the remaining 22 programs.

Concurrent with this effort, CNA began to look at the methods that would be used to cost the recommendations. Although no recommendations had yet been received, a broad strategy for cost estimation was created that proposed looking at direct program costs as well as fixed infrastructure costs.

In the March-to-May time frame, we received data from the program offices and began the process of checking the data to determine if they had been broken out correctly in the cost bins and if any data elements were missing.

The first recommendations were ready for cost analysis in mid-April. From that time forward, we worked on estimating the cost impact of the recommendations and examining overall T&E spending trends.

Caveats and assumptions

We make note of several caveats and assumptions to the following analysis. First, the 26 programs chosen were assumed to be representative of typical Navy acquisition programs. While every effort was made to include programs across a variety of ACATs and SYSCOMs, it is not necessarily clear that they truly represent Navy T&E spending.

A second caveat is that the definition of T&E differs across programs. This is compounded by a lack of a common Navy accounting structure and the fact that T&E is paid for by funds other than RDT&E (for example, many programs pay for T&E with procurement funds). To some extent, this problem was mitigated by including a detailed instruction letter that defined T&E according to Navy guidance. In addition, the cost subcategories were defined in a glossary at the back of the instruction letter. Despite our efforts to make sure that all programs had a common set of definitions with which they could gather and bin their T&E expenditures, it was clear after speaking with many of them that some programs included costs that others did not.

The third caveat is that we collected data for FY00 to FY04. We felt that looking back too far would be too difficult for the programs given their high personnel turnover. Furthermore, there is no rigorous accounting archive system that allows programs to look back and examine past expenditures easily. Thus, what we collected offers only a snapshot of programs' T&E spending rather than a complete dynamic picture. This problem is exacerbated by the fact that the study programs were all in various phases of their acquisition cycles. For example, certain programs were very mature by FY00 and the T&E spending during the FY00–04 time frame was very low. By capturing the tail end of the T&E process, we may not have captured those areas where the greatest amount of money was spent.

Given the above caveats, it is logical to conclude that the actual amount of T&E spending is probably higher than the reported numbers that we collected. While we do not know just how much of an underestimation this is, discussions with the working groups members and with the various programs lead us to believe that the data collected represent the bulk of overall T&E spending for those programs that responded to our data call.

Cost approach

Overview

Each working group generated a list of recommendations tailored to its particular area (Modeling & Simulation, Risk Management, etc.). When analyzing the direct cost impact, we decided to adopt a framework that first looks at how the program offices managed T&E from FY00 to FY04. We examined how the recommendation would have changed their behavior, estimated what that change in behavior would have meant in dollar terms, and translated those savings into a percentage.

For example, suppose the program offices had adopted a policy of using modeling and simulation (M&S) more aggressively. For some programs, this might have meant substituting simulated events for live-fire events. For others, this might have meant substituting simulated events for flight hours. In both cases, we would determine the cost differential between the two types of events to determine gross savings and add in the cost of the simulator to get the net dollar benefit. This calculation would also take into account the fact that M&S may not substitute for all live events.

The first step in costing a particular recommendation is to meet with the working group members to determine which of their recommendations have cost implications that can be estimated. In some cases, a recommendation was written in broad or vague terms that are not conducive to cost analysis. We then developed the first-order savings estimates using any of the normal cost-estimating methods. In most cases, we examined a program (not necessarily one of the study programs) that successfully adopted the recommendation in question and attempted to apply the results to other study programs to achieve cost savings. The final step was to conduct a sensitivity analysis and then refine our estimates by reviewing them with the working groups and program offices (as needed).

Financial data call structure and issues

The financial data call was a spreadsheet built around the format used for the program offices' R-3 budget exhibit (see figure 3). The justification for using the R-3s is that it includes the entire RDT&E budget for a program. Given that programs often spent money on T&E that is not explicitly called T&E, we concluded that including the programs' entire RDT&E budget represented a good starting point from which the program offices could determine how much was spent on T&E.

Figure 3. Sample R-3 budget exhibit

Exhibit R-3 Cost Analysis (page 2)

APPROPRIATION/BUDGET ACTIVITY	PROGRAM ELEMENT	PROJECT NAME AND NUM
RDT&E, N/BA-7	MK48 ADCAP/0205632N	MK48 ADCAP/V0366

Cost Categories (Tailor to WBS, or System/Item Requirements)	Contract Method & Type	Performing Activity & Location	Total PYs Cost	FY 99 Cost	FY 99 Award Date	FY 00 Cost	FY 00 Award Date
Test & Evaluation	WR	NUWC Newport, RI	CONT.	2.434	11/98	2.696	10/99
Developmental	Various	Various	CONT.	0.095	10/98	0.615	10/99
Modeling & Simulation	WR	NUWC Newport, RI	CONT.	1.050	11/98	2.207	10/99
Modeling & Simulation	C,CPFF	ARL/PSU State College, PA	CONT.	0.000		0.079	10/99
GFE							
Subtotal T&E			CONT.	3.579		5.597	

Another important consideration was that we collected financial data for FY00 through FY04. The working groups decided not to ask for financial information before this time frame because they felt that institutional knowledge for T&E events that occurred before FY00 would not be present in the program offices. Without this knowledge, it was felt that any financial data before FY00 would be suspect, especially in light of the fact that there is no common accounting system across the programs.

The R-3 divides a program's RDT&E budget into four major categories (T&E, product development, support, and management), any of which may be used to fund T&E. Each category may have multiple spending lines. For example, T&E spending for a sample program (as shown in figure 3) occurred at several places, including the Naval Undersea Warfare Center (NUWC), Keyport and the Applied Research Laboratory at the Pennsylvania State University (ARL/PSU), State College. In order to cue the program office to examine all funding for T&E, we included all spending lines in all categories.

Although the R-3 is a reasonable starting point, there are some issues that make it a less than perfect document for collecting financial data. First, the R-3 does not always label funds in the T&E category as clearly developmental test and evaluation (DTE), operational test and evaluation (OTE), acceptance test and evaluation, and so forth. For example, in figure 3, the sample program had a large sum of money generically labeled "Test and Evaluation" for NUWC. In some cases, the program offices were able to separate this money into some discernible form of T&E. In other cases, they were not. Thus, as will be seen in our analysis of T&E spending, we were forced to create a bin for certain T&E spending labeled "Unidentified."

The second issue is that the Program Element (PE) line associated with the R-3 does not encapsulate all T&E spending. Money spent on T&E can also come from other budget lines, such as procurement. Fortunately, the R-3 did have a section that identified procurement PEs associated with the program. We used that information to include another set of lines in the data call spreadsheet that identified the associated procurement PEs and asked the programs to examine that money and include any funds that were used for T&E. The actual budgeted procurement dollars used to populate the data call were pulled from the P-40 exhibits.

In examining T&E spending, we are also interested in determining what drives T&E costs. To do so, the working groups created 12 cost subcategories (listed in table 1), which represented a reasonable set of cost bins that most members of the working groups felt captured the essential types of expenditures made by program offices in the T&E phase. The working groups also believed that most programs could use them to "shoehorn" their costs even though their internal accounting systems did not track T&E expenditures in quite the same manner.

Table 1. Cost subcategories and definitions

Sub-cost category	Definition
1. Test Articles	Identification and cost of aircraft, weapons, etc. that are built as part of System Development and Demonstration and used for testing.
2. Targets/Ammunition/Other Consumables	Self explanatory
3. Development of Navy model or simulator	Cost of creating a Navy (versus contractor) owned M&S capability or facility used primarily for T&E purposes. These include any new infrastructure costs as well as the development of the model or simulator and its accompanying software.
4. Operational cost of using M&S facility for T&E	Cost of using the simulator (facility charge, etc) for a T&E event.Direct labor costs for operating the M&S capability during a T&E event. An example might be the cost of engineers conducting a test on a new system in a simulator versus in an actual platform.
5. Infrastructure/ranges/services	Fee for services performed at a Navy Working Capital Fund or MRTFB.
6. Command Support Costs	These costs refer to the "people-related" costs of the command and operational units providing collateral support to the T&E effort. These are additional costs incurred because of this effort. Examples are per diem pay, travel allowances, overtime, etc.
7. Requirements Development	Labor costs incurred for writing and developing T&E requirements (e.g. TEMP development).
8. Direct labor cost to T&E	Actual manpower costs incurred by the program in performance of a T&E event. These include engineers and other personnel actually performing a test.
9. Analysis	Labor costs incurred by the program for analysis of the test data and writing of required reports.
10. Contractor Testing	These costs include any testing done by a contractor but paid for by the program. This encompasses testing performed at the contractor site or contractor-funded testing performed at NWCF or MRTFB facilities
11. Unique support test equipment	The cost of a piece of equipment designed and built specifically for supporting T&E events. An example may be a sensor uniquely created for a specific program to monitor the performance of a system.
12. Other including DT or OT not otherwise ID'ed	Catch-all bin to be used for any other T&E costs that do not fall into the other categories shown above.

Figure 4 shows a copy of a portion of the spreadsheet for the same sample program as shown earlier. Note that there is a line for "Test and Evaluation" done at NUWC with budget numbers from the R-3. Below that line are the 12 cost subcategories. The program offices' job was to take this spreadsheet as a starting point and determine how much T&E spending occurred under each spending line. They were then to break these total expenditures across the cost subcategories. Interestingly, the data in the R-3s do not always accurately reflect true RDT&E, and actual spending may have differed greatly from the budgeted numbers reported in the R-3. For example, the R-3 showed that the sample program had a budget of $2.696 million for T&E at

NUWC in Newport, RI. The program office, however, reported a higher figure of $3.752 million.

Figure 4. Sample data call spreadsheet

Program cost spreadsheet				**OPTIONAL**		APPN: RDTEN	
Cost Category	Sub-cost Category - (T&E only)	Contract Method	Performing Activity & Location	Prior Years Cost (Actual)	Prior Years T&E costs (Actual)	FY 00 Cost (Budgeted)	T&E costs (Actual)
Test & Evaluation		**WR**	**NUWC Newport, RI**	**2.434**		**2.696**	
	Test Articles						0.325
	Targets/Ammo/Other Consumables						0.000
	Development of Navy model or simulator						0.470
	Operating cost of using M&S facility for T&E						0.000
	Infrastructure/ranges/services						0.040
	Command support costs						0.156
	Requirements development						0.000
	Direct labor cost to T&E						2.150
	Analysis						0.000
	Contractor testing						0.000
	Unique support test equipment						0.000
	Other including DT or OT not otherwise ID'ed						0.611
			Total	**2.434**	**0.000**	**2.696**	**3.752**

This page intentionally left blank.

T&E spending and cost trends

Program T&E costs

As we mentioned earlier, 26 programs were selected for this study, but only 18 of the 26 responded to our data call (see table 2). One program was excused from participating in the data call because it was actually a DoD program versus a Navy one. Other programs did not reply in spite of a great deal of effort on our part to persuade them to respond. One program simply could not respond because it did not have any detailed cost data and no real accounting system. Those programs that did respond indicated that collecting the financial data for FY00–FY04 was a tremendous burden. All 18 programs took several weeks to comply with the data call, and one program estimated that it spent 320 man-hours working on it.

Table 2. Study programs

	Total Programs in Study	ACAT I	ACAT II	ACAT III	ACAT IV	Other	Number of Programs that Responded to Data Call
NAVAIR	5	4	1	0	0	0	5
NAVSEA	14	5	4	3	1	1	10
SPAWAR	6	3	0	1	2	0	3
DRPM	1	0	0	0	0	1	0
Total	26	12	5	4	3	2	18

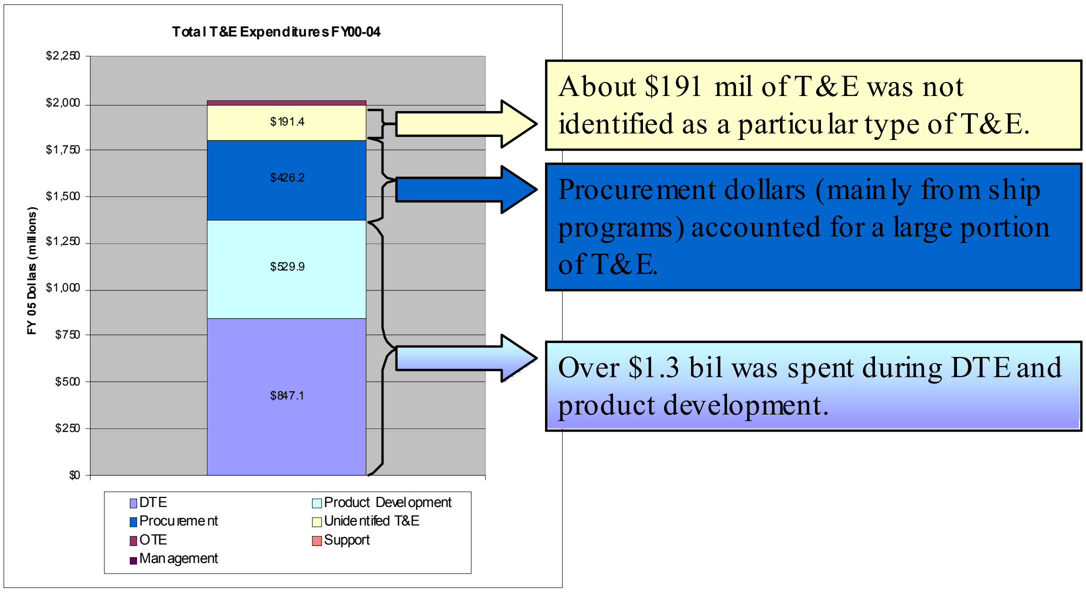
In the following analysis, we look at T&E spending in two ways. First, we examine T&E by aggregating across procurement and the four major categories outlined in the R-3 (T&E, product development, support, and management). The T&E category is further subdivided into "DTE," "OTE," and, in those cases where the programs were not able to label some T&E expenditure, "Unidentified." The purpose of looking at the data this way is to develop an understanding of where in the T&E process most programs spend their T&E dollars. The other way we analyze the financial data is by looking at spending across the cost subcategories. This analysis helped identify the various cost drivers for T&E.

It is important to note that, to avoid confusion, we did not ask the program offices to break either management (which is a major category in the R-3) or procurement into the cost subcategories. Thus, when we look at spending by major category, the total T&E expenditures will be greater than the total T&E expenditures as shown in the cost subcategories because the management and procurement dollars will not show up in the latter. To reconcile this problem, we treated both management and procurement as major categories and cost subcategories in the spending analysis.

Over the course of 5 years, the 18 programs reported that they spent just over $2.0 billion. More than $1.3 billion (or 68 percent) was spent with funds labeled DTE and product development. These two categories accounted for the lion's share of T&E. Procurement dollars were also a large source of funding for T&E. As a caveat to this finding, we point out that procurement spending was driven largely by one ship program. Finally, a fairly large portion of T&E (about 9 percent) belonged to the "Unidentified" category. These dollars represent a portion of T&E that the program offices were unable to identify. They were often a mix of spending across different types of T&E. Finally, note that T&E spending (not including procurement) was 19.5 percent of the programs' RDT&E budget. These results suggest that the majority of T&E is a large portion of a program's overall RDT&E budget and is mostly spent in the earlier phases of the program. (See figures 5 and 6.)

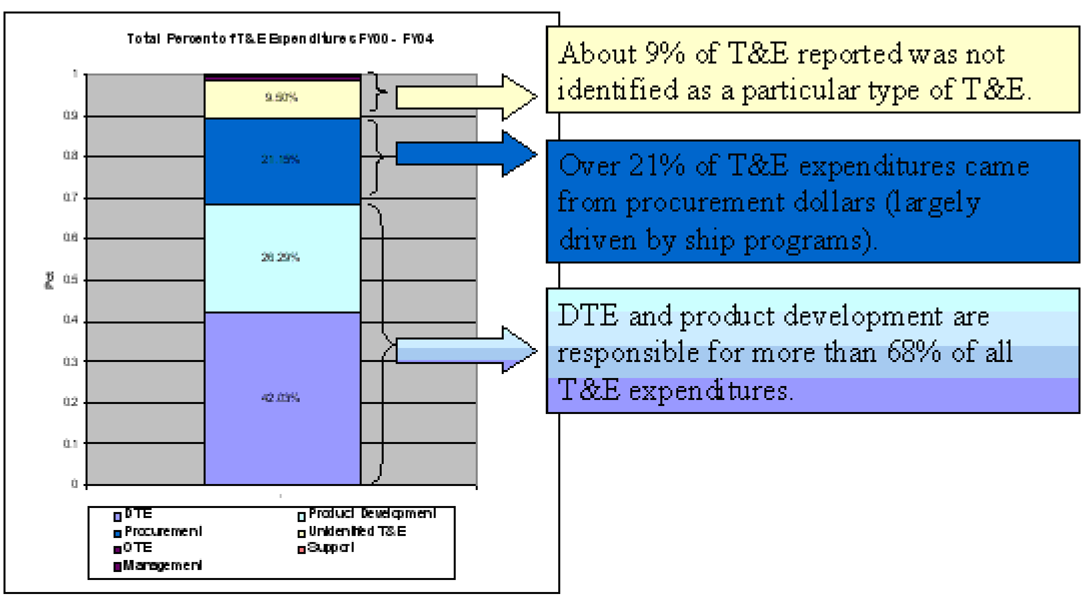
Figure 7 shows aggregate spending by the programs from FY00 to FY04. Spending by the 18 programs varied from $352 million to $440 million. Again, DTE and product development dominated spending across all years. Furthermore, the composition of program spending (defined as the percentage of T&E spent on each major category) was very stable across 5 years. This strongly suggests that, in the aggregate, the composition of Navy spending on T&E does not change much on a year-to-year basis. Newer programs whose T&E spending is far

Figure 5. Total T&E spending by major cost category (FY00–FY04)



T&E spending is driven by the development of the weapon, platform or sensor.

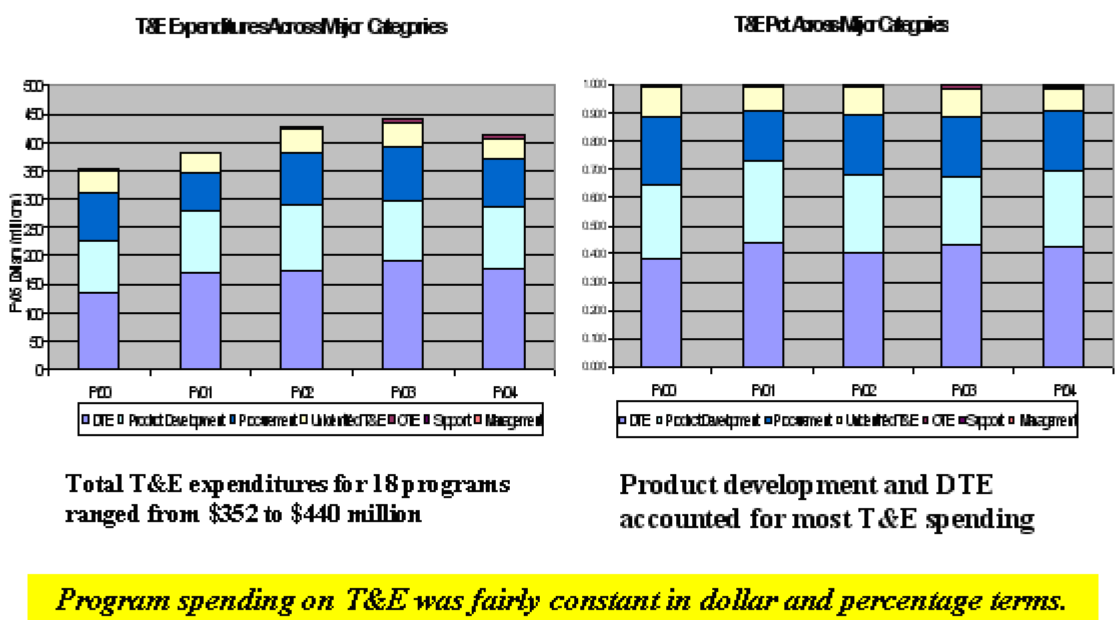
Figure 6. Total percentage of T&E expenditures by major cost category (FY00–FY04)



On average, reported T&E expenditures (not including procurement) represented about 19.5% of the programs' RDT&E budget.

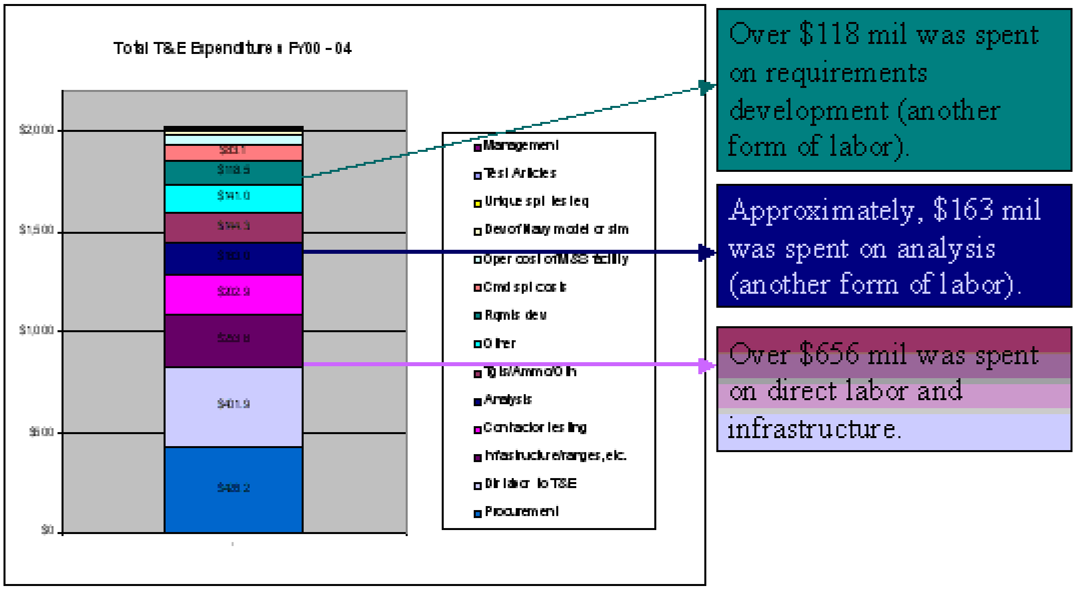
higher balance those programs that reach maturity and spend less on T&E. The implication is that any recommendation that affects a broad category of programs will have a similar cost impact in all years.

Figure 7. T&E yearly spending by major category



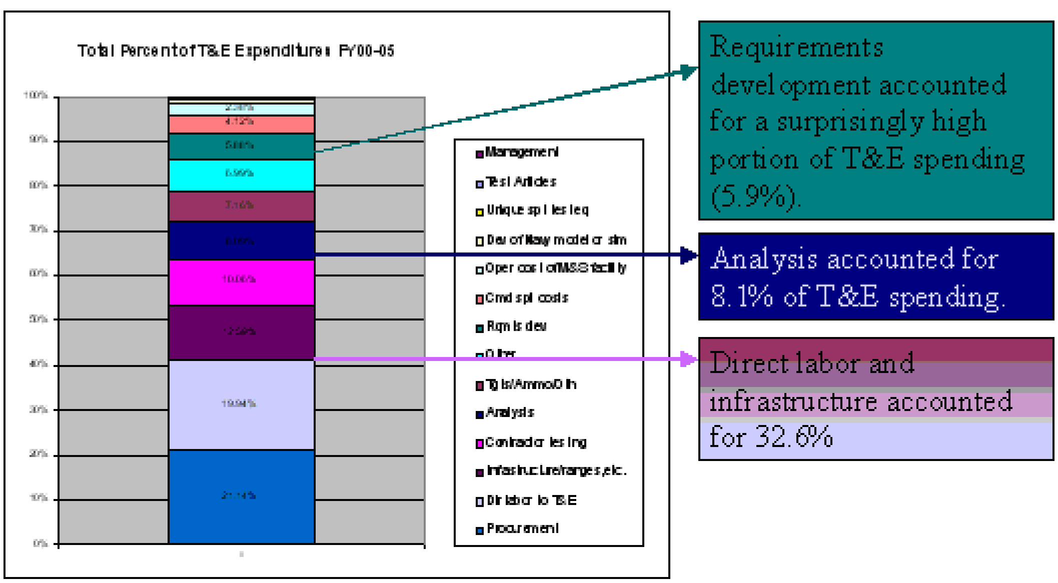
Figures 8 and 9 show that T&E is a very labor-intensive effort. About $401 million (roughly 20 percent of total T&E expenditures) went for direct labor. Actual post-event analysis and requirements development accounted for another $163 million (8.1 percent) and $118 million (5.9 percent), respectively. Added together, all forms of labor accounted for approximately $683 million, or 34 percent of the total T&E expenditures. Infrastructure costs represented the second largest cost subcategory (ignoring procurement) with programs spending about $253 million (12.6 percent of the total). These results suggest that T&E requires a great deal of labor and heavily relies on Navy infrastructure.

Figure 8. Total T&E spending by major cost category (FY00–FY04)



T&E is very labor intensive but also heavily uses Navy infrastructure.

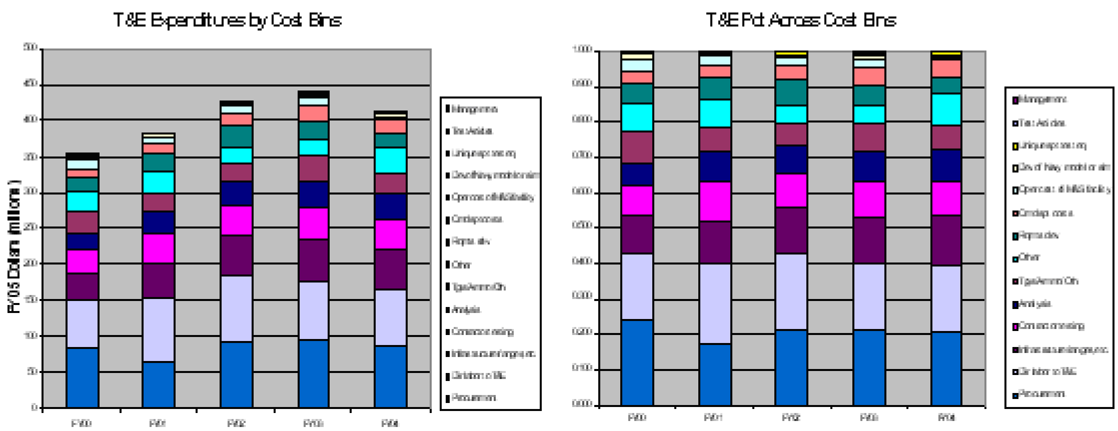
Figure 9. Percentages of total T&E spending by major cost category (FY00–FY04)



T&E is very labor intensive but also heavily uses Navy infrastructure.

We earlier noted that yearly spending by major category was very stable. This holds true for the cost subcategories as well, as figure 10 shows. Spending across cost subcategories varied slightly more than by major categories. But the analysis shows that labor and infrastructure still remained the largest pieces to overall spending. The implication from this is the same as for the major cost categories. Namely, any recommendation that affects a broad category of programs will have a similar cost impact in all years.

Figure 10. T&E yearly spending by cost subcategory



Total T&E expenditures for 18 programs ranged from $352 to $440 million.

Direct labor, analysis, and infrastructure account for 36% to 46% of total spending.

Program spending on T&E by sub cost categories was fairly constant in dollar and percentage terms.

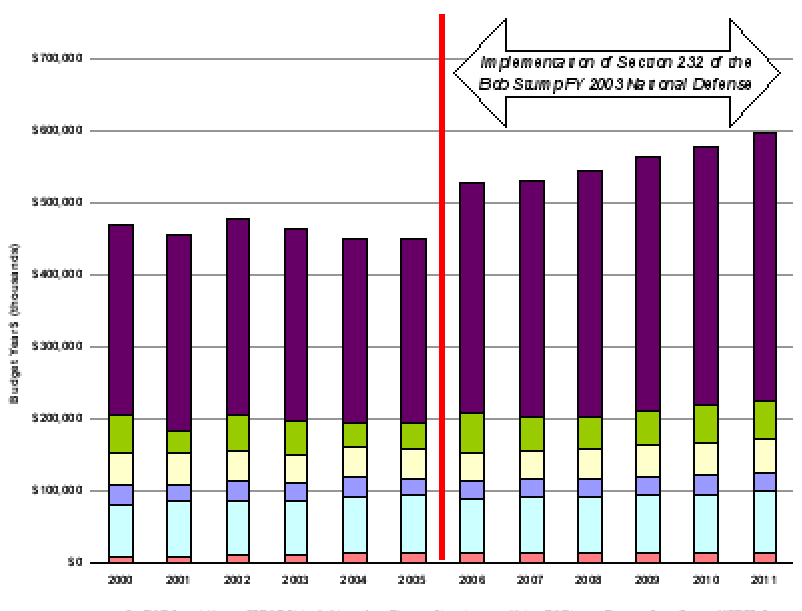
Navy infrastructure costs

While the programs control the largest portion of the Navy's overall T&E dollars, the Navy itself pays for the upkeep and capital improvements of several T&E facilities, such as the Major Range and Test Facility Bases (MRTFBs). To offer a more complete picture of Navy T&E spending, we also collected data on the following PEs that funded the Navy infrastructure costs:

- 0604256N—Threat Simulator Development
- 0604258N—Target Systems Development
- 0604759N—Major T&E Investment
- 0605863N—RDT&E Ship and Aircraft Support
- 0605864N—Test and Evaluation Support
- 0605865N—Operational Test and Evaluation Capability.

Figure 11 shows aggregate spending by year. RDT&E expenditures on infrastructure have decreased slightly in the last 6 years. In FY06, however, there is a significant spike in the budget for these PEs. This increase is largely the result of an increase in funding for the MRTFBs due to the implementation of section 232 of the Bob Stump FY03 National Defense Authorization Act (NDAA), which directed that the Secretary of Defense shall ensure that institutional and overhead costs of facilities or resources within the MRTFBs are fully funded through the major T&E investment accounts. It also directs that charges to DoD customers of these facilities or resources will be only for the direct costs of using the facility or resource. As such, for the Navy this plus-up for the MRTFBs represents net zero budget based transfer from the acquisition programs to the MRTFB account.

Figure 11. Spending on Navy T&E infrastructure



This page intentionally left blank.

Assessment of cost impact of recommendations

Due to time constraints, not all of the working groups' recommendations were examined for their potential cost impact. Instead, each working group met with CNA to discuss which recommendations should be analyzed. The first step in filtering the recommendations was to see which of them could be analyzed. In many cases, the recommendation was vague or spoke to a process or policy that simply made cost analysis impossible. For example, the Resource Cost Working Group's recommendation that the Navy adopt a more rigorous cost accounting system to track T&E expenditures should lead to cost visibility, which should, in turn, help the Navy save money on T&E. But any future savings will be an indirect consequence of this action. As such, this recommendation does not lend itself to cost analysis.

The second step was to determine whether data were readily available to do cost analysis on any of the remaining recommendations and whether there was sufficient time to do cost analysis. In some cases, a recommendation was theoretically amenable to cost analysis. However, the data collected in this study were insufficient, and it would have required another significant data call to the program offices to correctly determine the cost impact for a particular recommendation. In other cases, the cost analysis would have required only a minor request for additional data from the program offices, but there was not enough time to do the analysis. In the end, we performed cost impact analysis on four recommendations.

Recommendation 1: Reduce the need for ship shock trials

The Modeling and Simulation Working Group (M&SWG) offered one recommendation that was analyzed. This recommendation suggested that the Navy reduce the need for ship shock trials by investing in enhancements of related M&S. This recommendation specifically targeted ship shock trials versus other types of shock trials. Even though, in theory, this recommendation could be applied to other programs, we performed the cost analysis by looking at the one ship program in our study that sent in financial data.

The approach we used was to first determine how much the ship program spent on its most recent shock trial. We then gathered an estimate on how much a virtual shock trial would cost (experiment setup, software changes, etc.). The delta between these two represents the gross savings. We then included the cost of developing the appropriate model or simulator to estimate a net savings.

The gross savings calculation is shown in table 3. The ship program sent CNA a complete year-to-year breakdown of the ship shock trial costs. For example, in FY98 the program spent $318,000, mostly on environmental documentation. The actual live shock trial occurred in FY01 with the program incurring a $20.8-million bill in that year. Altogether, the program spent about $29.748 million on the shock trial over the course of 5 years. Members of the M&SWG estimated that the cost associated with a virtual event would be from $1 million to $2 million, which we divided evenly across the final 3 years since much of the expense of the first 2 years dealt with environmental compliance, which would not be necessary in a virtual test. The gross savings per year was calculated by subtracting the virtual cost from the actual. We then divided the savings by the program's reported T&E expenditures for that year to come up with a gross yearly savings to the program percentage (from 3.12 percent to 3.33 in FY00, etc.). Dividing the same dollar savings by the total T&E expenditures reported for the 18 study programs gave us the total gross savings percentage (1.56 to 1.66 percent in FY00).

PEO ships has a plan in place (currently not funded) to develop a DOTE-acceptable M&S equivalent to full-scale shock trial. The fixed cost of developing the model and simulator is about $33 million. This M&S tool would not be program specific but could be used by the Navy to do virtual ship shock trial for all future ships. Thus, in developing the net savings to the Navy, we did not apply the full cost of the M&S tool to the study program. Instead, we simply argue that the Navy would recoup the cost of this tool after two new ships, assuming that each of the future generation ships is required to do a ship shock trial and that the cost of the event is approximately $29 million (the cost of the study program's ship shock trial).

Table 3. Estimation of gross savings

	FY98	FY99	FY00	FY01	FY02	*Total*
Total actual ship shock costs FY05 $ (mil)*	.318	1.523	5.836	20.806	1.265	29.748
Cost of virtual trial **	0	0	.333 to .666	.333 to .666	.334 to .667	1.000 to 2.000
Gross Savings FY05 $ (mil)	.318	1.523	5.503 to 5.170	20.473 to 20.140	.932 to .599	28.748 to 27.748
Gross Savings/Program T&E Expenditures (Pct)	NA	NA	3.33% to 3.12%	14.29% to 14.06%	.51% to .33%	NA
Gross Savings/Total Program T&E Expenditures (Pct)	NA	NA	1.56% to 1.66%	5.36% to 5.27%	.21% to .14%	NA

* *Actual cost of program ship shock trial.*

** *Estimate of virtual ship shock trial cost as reported by NAVSEA Test and Evaluation Office = $1 to $2 million.*

Recommendation 2: Eliminate the difficulty in planning and coordinating T&E

The Test Planning and Execution Working Group (TPEWG) offered two recommendations that were analyzed. The first recommendation suggested that the Navy should eliminate the difficulty in planning and coordinating T&E on complex systems and increase the opportunity to test with training and experimentation events, thereby efficiently utilizing test resources. Adopting this recommendation would allow programs to perform more integrated weapon firings between DT, OT, CSSQT, and other fleet training events to maximize data collection. In discussions with CNA, the TPEWG identified three ways that programs could integrate testing.

1. Leverage test event off of CSSQT, training exercises, and other fleet events.
2. Combine DT and OT events.
3. Do concurrent testing during ship post-delivery test phase.

The approach we used to doing to cost analysis was to examine programs (not necessarily in the study) that have successfully used this business practice and estimate their achieved cost savings/avoidance. We then looked for study programs that, in hindsight, could have adopted or used the recommendation more aggressively and developed cost savings/avoidance estimates (possibly using the cost estimates from the successful programs). The three integrated testing strategies are sufficiently different that we treated each one separately.

Question 85F of the survey asked, "To what extent do you share resources with other events (CSSQT, Joint Force exercises, Sea Trials, exercises, etc.)?" Nine programs answered this question in the positive. We attempted to investigate to what extent these programs successfully leveraged fleet exercises but were able to receive a quick turnaround from only the Mk-48 program. As an example of a successful program, Mk-48 integrated 15 live-fire events with various fleet exercises. Unfortunately, we were not able to identify the costs that the Mk-48 program would have incurred had they needed to use dedicated assets for these events.

Four of the study programs responded "Not at all" or "Slight extent" to the same question (see table 4). To analyze the cost impact, we spoke with the four program offices to see where they might have integrated testing with fleet events more and what obstacles existed that prevented them from doing so.

The ACAT ID - NAVSEA program indicated that it already makes every effort to do this type of integrated testing. The problem is that the program cannot forecast fleet exercises (i.e., where it will occur, with what assets) very far in advance, which leads to scheduling conflicts. For example, in FY03, this program had been scheduled to do T&E during an exercise with USS *Nimitz*. However, shortly before the exercise, it was bumped from the schedule to make room for other program T&E priorities. This led to a 1-year slip in the schedule with an estimated cost impact of $12 million to $13 million (this added cost will be incurred at the end of the program due to the rightward movement of the program's schedule). Had a better scheduling tool been in place, all the programs involved with this scenario might have been able to agree to a better use of exercise assets with no program experiencing a slippage in its schedule.

Table 4. Programs that responded negatively to question 85F

Q. 85F - To what extent do you share resources with other events (CSSQT, Joint Force exercises, Sea Trials, exercises, etc.)?			
Program ACAT	**SYSCOM**	**Type**	**Response**
ACAT ID	NAVSEA	Network	Slight Extent
ACAT III	SPAWAR	Network	Not at all
ACAT IC	NAVAIR	Weapon	Not at all
ACAT IC	NAVAIR	Platform	Not at all

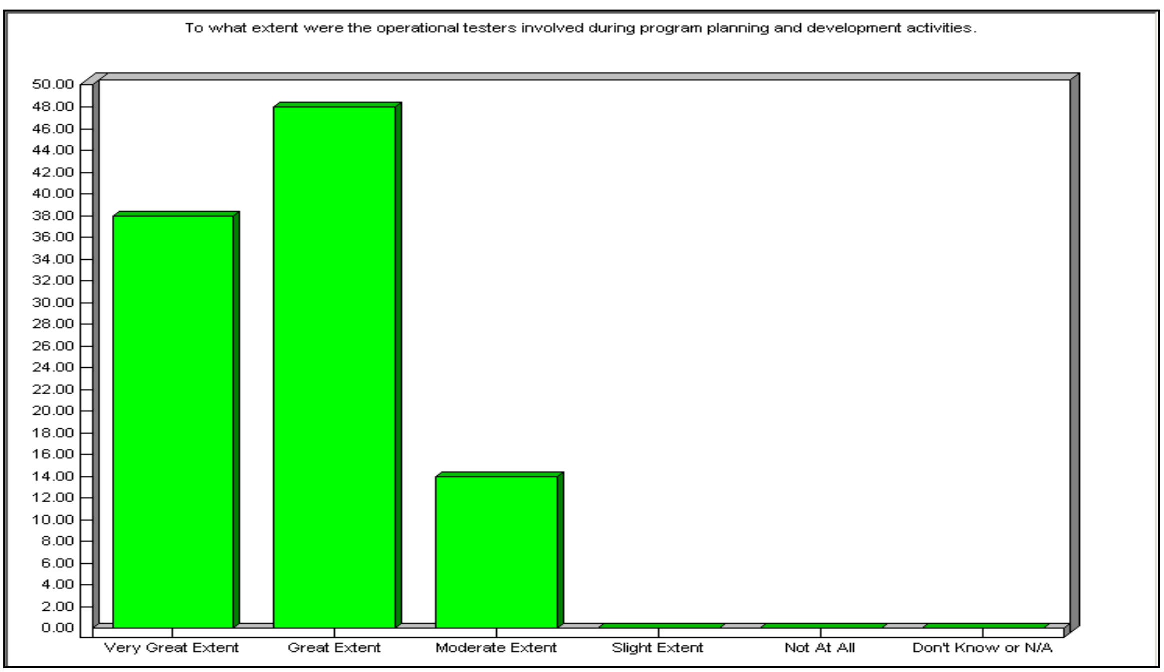
The ACAT III - SPAWAR program indicated that most of its T&E occurs on shore. Consequently, the program did not do integrated testing and found little reason to begin. In sum, programs such as these would experience no cost savings/avoidance as a result of this recommendation.

The ACAT IC - NAVAIR - Weapon program pointed out that, during FY02–FY03, it had a dedicated DDG that it used for T&E. While most of the T&E actually occurred while the ship was in port, the ship did go to sea eight times for live-fire events. We estimated that each live fire required the ship to be at sea for about 5 days. N43 provided us with a fuel burn rate of $30,000 per day for a typical DDG, leading to an added cost of $1.2 million (8 shots x 5 steaming days/shot x $30,000 fuel cost per day) incurred by the fleet (not the program).

The ACAT IC - NAVAIR - Platform program confirmed its survey response saying that it did not leverage fleet exercises for T&E. While the program was unable to provide any clarifying data to help us determine how much it might have saved had it adopted this business practice some time in the FY00 to FY05 time frame, it did acknowledge that there was probably room to do so, which would have yielded some savings.

Turning to the next form of integrated testing, integrating DT and OT, we note from the survey responses that all programs try to get the operational testers involved in the T&E process early in the program (see figure 12). This does not necessarily equate to actual DT/OT integration, but discussions with the same four programs from table 4 show that DT/OT integration is probably already being done to the greatest degree possible (at least in minds of the programs). The greatest obstacle to doing more DT/OT integration cited by the four programs was the difference between the test requirements for OT and DT.

Figure 12. Program responses to question 47



For the final form of integrated testing, concurrent testing during the ship post-delivery test phase, the TPEWG pointed to the T-AKE program as a successful case. In particular, in redesigning its post-delivery ship test schedule to do multiple tests concurrently, it successfully squeezed this phase from 18 to 10 ½ months. The program achieved this also by reducing the time needed for every test. OPEVAL, for example, originally slated to be 90 days, was condensed to a 30-day format. Projected estimates of cost savings/avoidance to this program are $24.5 million. The one ship program that responded to the data call indicated that it had already looked at its post-delivery schedule and had condensed the schedule in much the same way that T-AKE had done. Thus, while both of these programs should be used as successful templates for future ship programs, looking back at our programs, there does not seem to have been any savings that could have been achieved in the FY00 to FY04 time frame.

In sum, most programs are already doing integrated testing. The larger issue is that there are certain obstacles that prevent the programs from being more aggressive. For leveraging fleet exercises, allowing the T&E community to see fleet exercises further into the future would greatly help. To integrate OT and DT, requirements for each of the events would need to be aligned. See table 5 for a summary of cost savings.

Table 5. Summary of cost savings for integrated testing

	FY00	FY01	FY02	FY03	FY04	*Total*
Savings from Leveraging test events off of fleet events FY05 $ (mil)	-	-	$1.2	$12.0 to $13.0	-	$13.2 to $14.2
Combine DT and OT events FY05 $ (mil)	-	-	-	-	-	-
Do concurrent testing during ship post delivery test phase FY05 $ (mil)	-	-	-	-	-	-
Total savings	-	-	$1.2	$12.0 to $13.0	-	$13.2 to $14.2
Total savings/Program T&E Expenditures	-	-	N/A*	NA**	-	-
Total savings/Total Program T&E Expenditures	-	-	N/A*	NA**	-	-

* *The $1.2 million cost was paid for by the fleet versus the program office.*

** *The $12-13 million cost was due to a one year schedule slippage due to a missed T&E event in FY03. As this is an ongoing program, the actual cost of the slippage has not yet been incurred by the program.*

Recommendation 3: Robustly test earlier in the development cycle to mitigate the risk of late discovered failures

The other TPEWG recommendation suggested that Navy programs should robustly test earlier in the development cycle to mitigate the risk of late discovered failures. As part of this recommendation, TPEWG urged the effective use of modeling and simulation (M&S), such as models, hardware in the loop (HITL), and integration lab/ system integration lab (SIL), to test systems rigorously in buildup fashion to identify problems as early as possible in the development process rather than later in DT or OT. The working group argued that failure during testing would be expensive to a program. The recommendation for effective use of M&S is too broad, so we concentrated on the potential cost savings/avoidance from the use of SILs.

Our approach to doing to cost analysis was to examine programs (not necessarily in the study) that have successfully used or will use SILs to estimate their cost savings/avoidance. We then looked for study programs that, in hindsight, could have adopted or used the recommendation more aggressively and developed cost savings/avoidance estimates (possibly using the cost estimates from the successful programs).

We identified two programs that currently use SILs in their RDT&E programs: a rotary and a fixed wing aircraft program. Neither of these was a part of the 26 programs included in this study; therefore, we did not send out a data call, nor did we obtain T&E data for either of them. However, we were able to obtain some information that led to savings implications for the T&E programs.

The rotary wing aircraft test program contracted to build two SILs, consisting of a full SIL at the contractor site and a partial one at the government site in Patuxent River, MD. A full SIL consists of the following: (1) Hardware/Software Integration Testing section, (2) Master System Bench (MSB), and (3) survivability suite. The partial SIL at Patuxent River would not include an MSB, and the program office is considering adding that capability. To support the decision, it performed a cost-benefit analysis of procuring and maintaining an MSB at the Patuxent River site. The conclusion was that the net cost avoidance to include an MSB would be about $105 million. In coming up with the figure, it postulated that the program might expect a delay of almost 5 months if the MSB was not included in the Patuxent River SIL facility. The expected delay was based on an engineering assessment by the NAVAIR T&E team that 80 percent of the total SIL hours will be performed at Patuxent River, with 60 percent of those test hours performed on MSB (3,428 total SIL hours x 80% x 60% = 822 hours, or about 4.9 months). Such a delay would translate to an added cost of $123 million at the estimated loaded monthly expenditure rate of $25 million. The cost of MSB was estimated to be $18 million, including the costs to procure it and to support the integration and testing activities. The net cost avoidance would be the difference between avoiding the cost associated with the program delay and the total cost of SIL with MSB.

To understand the full effects of two full SILs, we took the total SIL hours of 3,428 and converted them to the expected schedule delay and the additional costs using the same factors as those above. The total cost avoidance from the use of two full SILs was about $515 million. Because one full and one partial SIL were already under contract, it would have been difficult to break down their costs precisely. We made a rough cost estimate by assuming that each of the three functions would cost about the same (although one could surmise that an MSB would cost more than the other functions since 60 percent of SIL test hours would be performed on MSB). Subtracting the costs (2 SILs x 3 functions x $18 million per function = $108 million) from the total cost avoidance of $515 million, we arrive at a ballpark estimate of the net cost avoidance of a little more than $400 million.

To put the cost avoidance in perspective, we examined the total RDT&E cost for this program. The net cost avoidance of $400 million represents about 10 percent of the total RDT&E costs, which is $3,511 million (from the FY06 President's Budget Exhibit R-2).

The fixed wing aircraft test program uses three test aircraft: an air vehicle/weapons carriage test aircraft (no mission systems), a dedicated mission systems test aircraft, and a mission systems/weapons delivery aircraft. The development flight test would begin in January 2009 and end in October 2011, and the total number of flight hours for integrated testing would be 2,280. The test program developed during the early stages of the RDT&E program would continue to be refined during the System Development and Demonstration (SDD) phase.

The test program is based on the maximum utilization of the SILs throughout the SDD program. The SIL program consists of four laboratories—three at the contractor site and one at the government site. The labs and their functions follow:

- Software Development Lab (SDL)—Kent, WA
  - Develops software on processor/operating system to be used
  - Reduces "port to other operating systems" problems
  - Connected to distributed network
- Mission Systems Integration Lab (MSIL)—Kent, WA
  - Integrates "box level" hardware/software
  - Limited hardware/software-in-the-loop testing
  - Connected to distributed network
- Weapons System Integration Lab (WSIL)—Kent, WA
  - Fully integrated hardware/software-in-the-loop capability
  - Representative racks, wiring, layout, etc., including stores management
  - Connected to distributed network
- Patuxent River Integration Lab (PAXSIL)—Patuxent River, MD
  - Fully integrated hardware/software-in-the-loop capability
  - Representative racks, wiring, layout, etc.
  - Connected to distributed network
  - Support test planning, training, conduct, and integration troubleshooting.

The program budget reflects an estimated T&E savings from the use of SILs for the flight test program of 300 to 500 hours. This equates to 6 to 11 aircraft-months saved at an average flight rate of 45 hours per month per aircraft, which would result in a schedule reduction of 3 to 6 months based on two mission systems/weapons aircraft. There would be an additional savings of 3 months of on-aircraft ground test for a total schedule reduction of 6 to 9 months.

The RDT&E budget for this fixed wing program is about $1 billion per year from FY09 to FY11 when DT/OT flight-testing is scheduled (from the FY06 President's Budget Exhibit R-2). This implies that a schedule reduction of 6 to 9 months would result in a cost savings of roughly $500 million to $750 million. The program office could not break down the costs of SILs because they were a part of the overall contract. We made a rough estimate that they would range from $110 million to $220 million based on the SILs' costs of the rotary wing program (about $110 million for two full SILs). The lower number would reflect an assumption that the SILs' functionality for the fixed wing program would be about the same as for the rotary wing program, and the higher number would reflect an assumption that the costs of SILs for the fixed wing program would be about twice as much given that there are four SILs for that program vice only two for the rotary wing program. Based on the foregoing assumptions/estimates, the net cost avoidance would be about $300 million to $600 million, which would represent 4 to 9 percent of the total RDT&E budget.

For the two aviation programs we examined, the estimated net cost avoidance from the maximum utilization of SILs ranged from 4 to 12 percent of the total RDT&E expenditures. We expect that the savings percentage for the T&E program would be higher than the savings percentages for the whole RDT&E program because the use of SILs would affect the T&E program more than the rest of the RDT&E program.

The program office representatives we talked to implied that all future aviation programs (both fixed and rotary wing aircraft) could achieve similar magnitudes of savings. They cautioned, however, that for the programs that currently use SILs, the budget already reflects the savings (hence the term cost avoidance to highlight the point).

We attempted to apply the savings analysis to some of the study program to see if any savings might have been achieved had some of them adopted the use of SILs. In particular, we looked at the platform programs in NAVSEA and NAVAIR that shared traits with the rotary wing program described earlier. We concluded that, although there might have been room in some of these programs to adopt SILs, there was probably no cost avoidance to be had in the FY00 to FY04 time frame since these programs were still in the middle of the program phase. Recall that in the rotary wing case, the savings were to be achieved by avoiding delays that push the schedule to the right. Given that our study programs were still in phase B or in the early stages of phase C, any cost avoidance applicable to these programs would happen in the future years.

Recommendation 4: Improve the process for the generation and refinement of requirements/capabilities to ensure testability

The Requirements Working Group (RWG) offered one recommendation that was analyzed: Navy programs should improve the process for the generation and refinement of requirements/capabilities to ensure testability. This recommendation is very broad and speaks to various issues, such as redundant testing and better T&E scheduling. We refine our cost analysis to look at one issue in particular—testing that is either unnecessary or excessive.

Our approach to cost analysis was to look for study programs that did excessive testing and estimate the cost savings/avoidance had the programs been able to skip these tests. We identified these programs by look at the responses to question 25 in the survey, which asks, "In the execution of this program, to what extent were there any testing requirements that were imposed to ensure compliance with KPP/thresholds that, in your opinion, were entirely unnecessary, inappropriate, or excessive?" From table 6, we see that four programs reported that there had been some degree of unnecessary T&E. As part of the cost analysis, we spoke with members of all four programs to gather information on the type and scope of unnecessary testing.

Table 6. Programs that responded positively to question 25

Q. 25 - In the execution of this program, to what extent were there any testing requirements that were imposed to ensure compliance with KPP/thresholds that, in your opinion, were entirely unnecessary?			
Program ACAT	**SYSCOM**	**Type**	**Response**
ACAT IC	NAVAIR	Platform	Moderate Extent
ACAT IC	SPAWAR	Network	Great Extent
ACAT II	NAVSEA	Network	Great Extent
ACAT IC	NAVSEA	Platform	Slight Extent

The ACAT IC - NAVAIR program office acknowledged that it did incur a cost associated with testing that was excessive. According to the program office, various adjustments were necessary to accommodate changes in the interpretation of requirements as the personnel at various test organizations changed. For example, the entire electronic warfare (EW) section of the ORD was rewritten after more than 2 years of testing based on "new" interpretation(s). These changes resulted in the program having to repeat some previous tests. A much greater impact, however, was the degree of churn and programmatic ripples caused by these new interpretations. The program office estimated the cost impact to have been about $400,000 incurred over the period of FY00 through FY04.

The ACAT IC - SPAWAR program also acknowledged that it incurred a significant added cost for executing unnecessary T&E. In this case, there was a duplication of Navy testing and JITC testing required for a system to receive Joint Interoperability Certification.

The program office was able to provide only rough cost estimates because actual cost data were scattered across a variety of budget documents. Using a very rough cost of about $65,000 per week per combat system test bed and $15,000 per week for the test command, a 2-week Navy-only test cost about $160,000. Note that this amount does not include test procedure development, post-test analysis, dry run of test procedures, and so on. The cost of a corresponding joint test was approximately $810,000. Additional costs included about $350,000 to COMOPTEVFOR and $200,000 to Navy DT to support the event. Aggregating the numbers yields a total cost estimate of $1.4 million per platform certification. Multiplying this figure by the 27 platforms results in a total cost estimate of $33.8 million for unnecessary testing. This amount does not include all the pre- and post-test work, which could be as much as four times the testing costs.

The ACAT II - NAVSEA program also reported having to do excessive T&E—in this case, due to the Combat System Key Performance Parameters (KPPs) being assigned to the Element TEMP. The additional labor for completing the added tests was estimated to be about 6 hours of prep work and 6 hours of test conduct for each of the 15 test events conducted. In addition, there were 20 hours per event for data analysis. This adds up to 480 additional hours of test schedule over the test period.

The costs associated with the additional requirements are estimated to be:

- 6 hours of prep work per test for 15 tests at $75/staff hour for 5 staff = $33,750
- 6 hours of conduct per test for 15 tests at $75/staff hour for a 10-man test team = $67,500
- 6 hours per test event of Lear A/C at $3,500/hour for 15 tests = $315,000
- 20 hours of data analysis for 15 tests at $75/staff hour for 3 staff = $67,500.

The total estimated cost over FY00-FY04 is $483,750.

The ACAT IC - NAVSEA program experienced a slightly increased level of cost and risk attributable to differences in the definition of operational availability (AO) at various Operational Testing (OT) agencies. These differences in key performance parameters led to requests for a greater degree of retesting during FOT&E than would have otherwise occurred. The final extent of this cost increase has not yet been determined; the OPEVAL tests occurred during the period covered by our data call, but the schedule of follow-on tests has not yet been finalized. The program cost for an added FOT&E event is likely to be in excess of $500,000.

Table 7. Summary of cost savings for eliminating excessive testing

	FY00	FY01	FY02	FY03	FY04	*Total*
ACAT IC – NAVAIR – Platform	0.080	0.080	0.080	0.080	0.080	$0.400
ACAT IC – SPAWAR – Network*	NA	NA	NA	NA	NA	NA
ACAT II – NAVSEA – Network	0.097	0.097	0.097	0.097	0.097	$0.484
Total savings	$0.177	$0.177	$0.177	$0.177	$0.177	$0.884
Total savings/Program T&E Expenditures	1.03%	0.54%	0.51%	0.54%	0.51%	0.58%
Total savings/Total Program T&E Expenditures	0.05%	0.05%	0.04%	0.04%	0.04%	0.04%

** The cost savings for the SPAWAR program was not included in the summary due to the fact that we are uncertain when the program incurred this cost. However, the savings appear large and would greatly increase the percentage savings.*

In summary, table 7 shows that eliminating excessive testing could lead to some programs saving anywhere from 0.5 percent to 1.0 percent. For all programs taken together, this equates to a relatively small 0.04 percent to 0.05 percent. However, the final calculations do not include the cost savings for the SPAWAR program because we were not sure when these costs were incurred. As a result, these savings percentages should be viewed as a floor with potential savings likely to be much higher.

This page intentionally left blank.

List of figures

This page intentionally left blank.

List of tables

This page intentionally left blank.

CRM D0012661.A1 / Final