A DECADE OF SUCCESS AND FAILURES IN THE DOD ACQUISITION SYSTEM

A Continuing Quality Journey

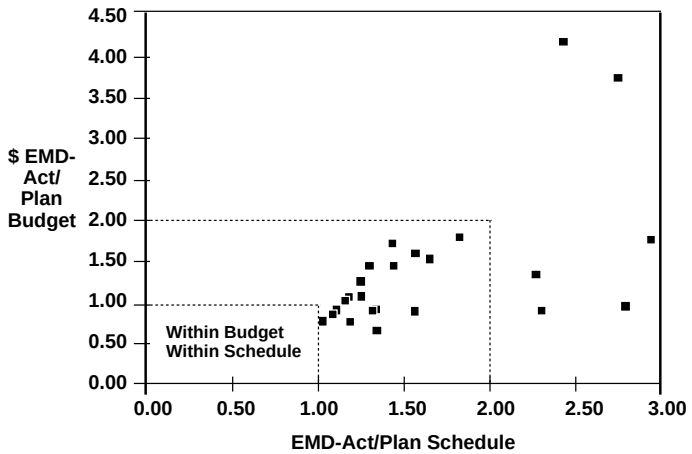
Raymond W. Reig

A recent research effort identified the current "track record" of Department of Defense weapon systems acquisitions. In the spring of 1993, the office of the Director of Operational Test and Evaluation (DOT&E), asked DSMC to undertake a research study of the current DoD acquisition system for a very specific reason.[1] Since 1991, Congressional law requires the DOT&E to specify at Milestone II, the number of test articles required for the Initial Operational Test and Evaluation in the Army and Air Force, and Operational Evaluation in the Naval Services (IOTE/OPEVAL) system test. The office desired to know if there was historical data that would help in making this important, and possibly costly decision.

(**Editor's Note**: In the July-August issue of Σ, Professor Reig commented on the acquisition of software-intensive systems. This article reviews the acquisition history of 24 recent programs of all types.)

Professor Reig is on the faculty at DSMC. He is a graduate of the U.S. Naval Academy, New York University, Auburn University, Southern Illinois University, the Air War College and PMC 90-1.

FIGURE 1. Results: Ratio of Actual vs. Planned



Engineering and Manufacturing Development Phase

Because the acquisition phase of interest was the Engineering and Manufacturing Development phase (EMD), we concentrated our efforts on this phase exclusively. We surmised that if we reviewed the data within the "Blue Books" prepared for Defense Acquisition Board (DAB) meetings at Milestone II and Milestone III, we could determine the program's success in adhering to its EMD phase. "Blue Books" are summary program plans and data compiled for the use of DAB principals and staff assistants immediately prior to milestone meetings.

Milestone II DAB information gave us the program's planned activities and expenditures. The Milestone III DAB information gave us actual activities and costs incurred during EMD. We were particularly interested in the planned and actual costs for EMD, and the planned and actual schedule. To gather this data, we also looked at the annual Selected Acquisition Report (SAR) for the year the program conducted its Milestone III DAB meeting.

We applied this approach to 24 recent DoD acquisition programs. As we reviewed the DAB "Blue Books" and program SARs, we recognized that we could extract much more information than was necessary to answer the basic question:

Is there a relationship between the number of Low Rate Initial

> Production (LRIP) articles used for testing in EMD/IOTE/OPEVAL, and the success of that program?

Also, we intuitively assumed that LRIP quantity alone could not account for the success of any program in EMD. We identified six other variables that could affect a program during EMD and gathered this data as well.

The net result of this approach is that we gathered data from the EMD phase that —

1. determined how successful these 24 programs were in meeting their planned EMD cost and schedule;
2. determined the relationship between LRIP test quantities and program success; and
3. identified how other variables affected program success.

EMD Cost and Schedule Data

Figure 1 indicates the results of planned and actual EMD cost and schedule data for the 24 programs. This is the ratio of the actual EMD results to the planned figures for cost and schedule. A program that came in essentially on-cost and on-schedule, would have a ratio of 1:1 for cost and 1:1 for schedule. Of the 24 programs, 18 came in under a 100-percent cost and schedule overrun (ratios 2:1, and 2:1). For the entire 24 programs, the average cost overrun in the EMD phase of the program was 45 percent, and the average schedule overrun was 63 percent. The data showed more programs were better able to control their cost overruns than their schedule overruns. For the remainder of this article, and for the final research report, we discuss the data obtained by reviewing all 24 programs, unless otherwise noted.

LRIP Test Quantity

Figure 2 shows the results of the data gathered to answer the question of whether the LRIP test quantity was related to program success. (Here program success was defined as a low schedule overrun.) We are dividing the number of the LRIP test articles used within a program by the total planned production quantity, expressed as a percentage. This is compared to the probability of the program having an EMD schedule slip less than 50 percent. These data (21 programs) show that programs using three percent or more LRIP test articles never exceeded a 50-percent schedule slip, and the probability steadily increased as the three-percent mark was approached.

The data also show that 28 percent of the total LRIP quantity was acquired with research, development, test and evaluation funds and presumably used for testing. The remaining 72 percent was presumably used for other than test purposes.[2] Proving out the manufacturing process and ramping up the production rate are the only other authorized purposes of LRIP systems. Since this is accomplished regardless of the end use of the LRIP systems, a question could be asked:

> Is the current 28-percent LRIP test usage/72 percent other than test usage the proper mix of LRIP systems?

In this regard, a recent Naval Postgraduate School thesis concludes that the majority of the problems that occurred during IOTE/OPEVAL are directly related to test resource issues.[3] It also recommends "sufficient test articles should be produced and available well before the operational test is supposed to start."

Evaluation Scheme

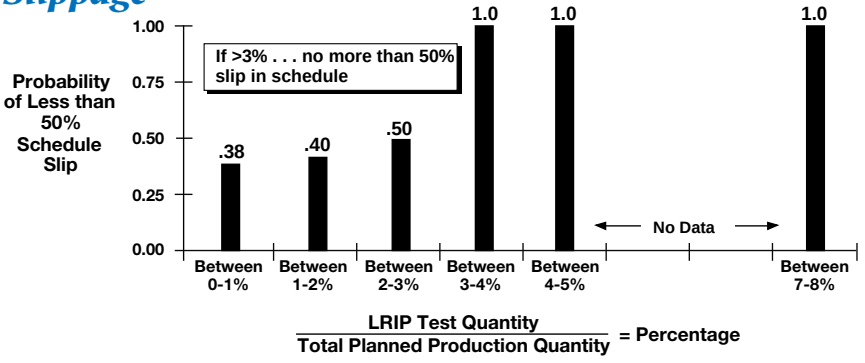
Further, in order to evaluate the impact of the other variables, we devised a standard program success rating scale of one to five. Five indicated the program was "Very Successful," essentially on planned budget and schedule. A rating of four indicated the program was "Successful," exceeding the EMD budget by less than 30 percent, and exceeding the schedule by less than 12 months. Three indicated a "Fairly Successful" program, exceeding the budget by less than 45 percent, and the schedule by less than 18 months. The two lower success ratings decrement by 15 percent for budget overruns and six months for schedule overruns.

It is important to recognize that our evaluation scheme is a measure of program management success at a point in time rather than eventual weapon system success. There was no attempt to evaluate the effectiveness of the systems in their operational roles. Also, we did not evaluate the performance (Effectiveness and Suitability) of these programs in EMD, although that would be an excellent additional study and relatively easy to do.

The six other variables selected were —

1. the degree of risk identified at Milestone II;
2. use of competition during the

FIGURE 2. LRIP Test Quantity and Schedule Slippage



Demonstration/Validation(DEM/VAL) phase;

3. use of competition during the EMD phase;

4. type of contract;

5. number of associate contractors; and

6. Joint- or single-Service program.

From the data available in the program "Blue Books" or the SAR, we assigned a success rating of one to five in accordance with the plan outlined above. We next compared this rating with the program's other variables. For example, did the program at Milestone II indicate it was low-risk or medium-risk?

Program Success Rating Results

Finally, we averaged the program success ratings of the programs that were low-risk and the programs that were medium-risk, and found in this instance, programs that were medium-risk had a higher average program success rating than programs that were low-risk. Initially we assumed a low-risk program would have a better probability of success than a higher-risk program. The data did not support this assumption. Programs that evaluated their risk as Low at Milestone II had an average program success rating of 2.4 as compared to programs with a risk evaluation of Medium, which had an average success rating of 3.3.

Ten programs had no summary risk estimate that we could find. Only one program indicated Software Lines of Code, an accepted metric of software risk. Programs that did not use competition in DEM/VAL had a higher program success rating than did those rated 3.3 vs. 2.0. Programs that did not use competition in EMD also had a higher success rating, 3.1 versus 2.2. Four EMD contract types were evaluated, and the data resulted in the following program success ratings: Cost Plus Incentive Fee, Fixed Price Incentive and Firm Fixed Price—all 3.3; Cost Plus Award Fee—1.0.

The data also showed that programs using one or no associate contractor (besides the Prime) had a success rating of 3.4 compared to a rating of 2.4 for the programs that used more than one associate contractor. For probably the same organizational complexity reasons, single-Service programs had an average success rating of 3.0 compared to Joint programs' success rating of 1.8.

Conclusions

To conclude by returning to our initial paragraph—how are we doing? Well, it depends (a classical phrase here at the college). It depends on what DoD weapon systems acquisition is compared to. Our overall averages were discussed herein, but Figure 3 shows DoD as compared to other industries.[4] Here the comparison is more favorable. And if we return to Figure 1, and consider only the 18 of the 24 programs reasonably clustered together, the average cost overrun is 20 percent, and schedule overrun is 32 percent. These are credible numbers and much better than the 45 percent and 63 percent overall overruns.

In Total Quality Management terms, our performance with 18 of the 24 programs reviewed can be thought of as our current statistical process control bounds for DoD weapon systems acquisition. "Control charts are simple devices used to establish bounds on measured characteristics, so that when the bounds are exceeded, production should be suspended and the machine in question should be adjusted or repaired.[5]" Hopefully, this data defines our current baseline on our continuing quality journey.

References

1. Gailey, C. K., Reig, R. W. and Weber, W. (1994); Ongoing, unpublished DOT&E-sponsored research of DoD acquisition, Defense Systems Management College, Fort Belvoir, Va.

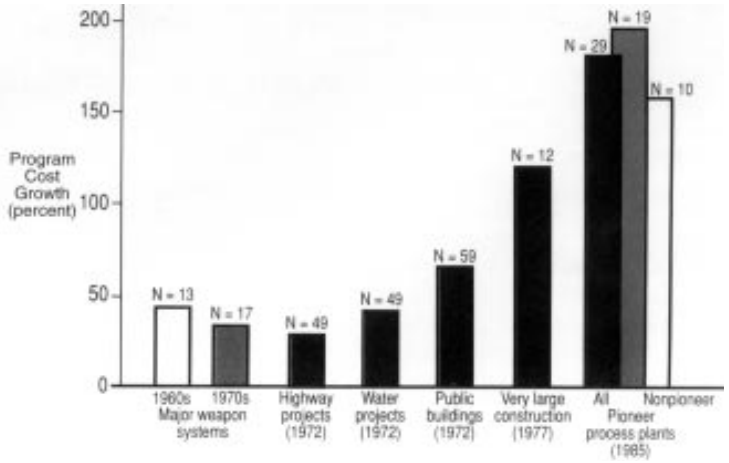
2. Reig, R. (1994 July-August), "Insufficiently Robust DT&E Means Trouble Ahead for OT&E," *Program Manager*, Vol 23, No. 4, pp. 12-15.

3. Mills, J., "An Analysis of Weapon System Readiness for Operational Testing," Masters Thesis, Naval Postgraduate School, Monterey, Calif., 1994.

4. "Final Report to the President" by the President's Blue Ribbon Commission on Defense Management, *A Quest for Excellence*, Appendix, p. 61, June 1986.

5. Bailey, M., Bowden, J. and Callahan, A.J., "Managing Ship Performance of Naval Gunfire Support Using Statistical Process Control," *Military Operations Research Journal*, Vol. 1, No. 1, Summer 1994, Alexandria, Va.

FIGURE 3. Cost Growth in Major Projects (RAND)



Source: "Improving the Military Acquisition Process — Lessons from Rand Research," (R-3373-AF/RC) The Rand Corporation, 1986.