Image designed by Schatz Publishing Group

LESSONS FROM THE DEVELOPMENT OF ARMY SYSTEMS

William A. Lucas and Richard G. Rhoades

Several of the statistically significant relationships found in this study of 13 Army systems involve factors that are related to the stability of the program. For example, uncertainty of a project's future and funding cutbacks are found to have a strong predictive influence on development program effectiveness, which may be explained in part by their impact on program staffing turnover and the disruption of testing regimes. A central conclusion from this study is that shorter development cycle times favorably correlate with lower levels of these sources of program instability, and with substantially better project outcomes.

This article is based on the results of Army Materiel Command-sponsored research of several years' duration (Lucas & Rhoades, 2004). A structured case study approach was employed to examine the history and processes that had resulted in the introduction of a number of technology-based Army systems in time to make a positive contribution to the outcome of DESERT STORM. The 13 case studies that resulted were developed on systems ranging from the M829A1 "silver bullet" to the GUARDRAIL Common Sensor and the APACHE attack helicopter.

METHODOLOGY

The case studies were prepared largely based on interviews with key participants from the government/contractor team that developed each system, and using a research questionnaire to structure the discussions. The authors designed the questionnaire to provide coverage of a number of development process, organizational relationship, critical technology maturity, and other factors that either the authors' prior experience or the literature (GAO, 1999; GAO, 2000) suggested might be relevant to determining the relative success of projects. A portion of the questionnaire consisted of questions that were in common with a research instrument successfully used by one of the authors in a prior study (Air Force Research Laboratory, 2002) of aerospace research projects. This process resulted in collection of a common set of data for the systems studied, which could then be analyzed to identify factors contributing to the relative degree of success in system development.

The heart of any systematic analysis is the definition of a common outcome measure that allows comparison.

The heart of any systematic analysis is the definition of a common outcome measure that allows comparison. In this study, the projects (cases) were compared based on their performance relative to their agreed-upon goals and requirements. Each project had a budget, a systems procurement cost goal, a set of technical requirements, and completion dates. In addition, questions of performance are immediately observable and easily remembered by project managers: Once production was started, were problems found that required that further engineering changes be made? And did the system perform well during its use in DESERT STORM? Using structured questions, the researchers asked the key government and industry interviewees how well their projects performed in these areas, with a range of answers that characterized how badly the projects had missed meeting their objectives if they had not been completely successful. Note that researchers developing each case study had the independent views of at least two senior managers, as well as their own detailed study of their project, to enable them to make summary judgments on project success meeting these largely observable outcomes.

Six of these outcome measures were used to create a scale that scores the projects from zero to six according to the number of key outcomes a project achieved. If a project was (a) transitioned to production on time, (b) developed within budget, (c) had no late engineering changes, (d) met the goals for system unit costs, (e) met the goals for technical requirements, and (f) encountered no difficulties when deployed in the field, it was awarded six points on this scale. These results appear in the third column of Table 1.

FACTORS INFLUENCING STABILITY

Previous reports on systems development have noted the importance of various factors that influence the stability of system acquisition programs. As part of this research, program funding uncertainty and cutbacks, changes to the system requirements (e.g., changes to the threat the system was being designed to defeat), and changes in key military personnel representing the user community (e.g., Training and Doctrine Command) were all examined to see to what extent any or all of these

TABLE 1. SUMMARY CASE INFORMATION

SYSTEM/CASE	DEVELOPMENT DURATION (MONTHS)	KEY OUTCOMES ACHIEVED (0-6)	COMPLEXITY (LOW, MEDIUM, HIGH)
APACHE attack helicopter	108	1	High
TADS/PNVS (target acquisition and designation/pilot's night vision systems)	~36	3	Medium
MLRS rocket system	33	6	Low
ATACMS missile system	37	6	Medium
M40 chemical protective mask	~48	2	Low
Mounted microclimate cooler	~24	5	Low
M829-A1 armor–piercing kinetic energy tank ammunition	~36	6	Low
TOW-2A (Tube-launched missile)	48	3	Low
AN/TAS 4 infrared night sight	~24	4	Low
Joint Stars Ground Station	105	1	Medium
Guardrail common sensor	~24	3	Low
PAC-2 (PATRIOT anti-missile system)	~52	2	Medium
HELLFIRE missile system	~84	3	High

TABLE 2. PROGRAM STABILITY

Type of Instability	Question/Response
Funding uncertainty	6 of 12 disagreed strongly that there was uncertainty about the future of project funding.
Project slow-down	8 of 13 projects were not stopped and restarted or slowed down.
Funding cut-backs	5 of 13 escaped changes or compromises forced by cut-backs in project resources.
Turn-over in Army user representatives	All projects experienced changes in key TRADOC personnel during development. This occurred only once or twice for 7 of 13.
Change in systems requirements	4 of 13 had no changes in systems requirements during development.
Change in system requirements	4 projects experienced systems requirement changes* in the middle of development; in 3 these changes occurred late in development

**Responses selected as many periods as applicable from the stages of planning; early, mid- and late development; and transition.*

"instabilities" impacted program outcomes. Table 2 contains the aggregate results from questions that were used in the research, categorized by type of instability.

A somewhat surprising result is that there was widespread occurrence of all these forms of instability in the pre-DESERT STORM development programs studied here. Looking back at the successful performance of new Army systems in DESERT STORM, one might think that their development benefited from strong and stable environments, but the evidence shows that the external environment then was not unlike what one might find today.

FINANCIAL UNCERTAINTY AND PROJECT CUTBACKS

Where potential or actual funding changes were encountered ("financial uncertainty"), it appears to have had significant consequences. When one looks at the projects that are reported to have been slowed (i.e., experienced a lengthening of the planned development phase), all five also experienced problems due to financial cutbacks. By comparison, only three of eight that were not slowed experienced problems due to cutbacks (Table 3). While program slow-down may be caused by a variety of factors besides or in addition to budget cuts, once slowed, programs seem to have continuing financial problems. Given that cutbacks are often the first signal that a

TABLE 3. FUNDING INSTABILITY AND FORCED CHANGES AND COMPROMISES

A. Often uncertainty about future of project funding?		
Cut-backs forced changes	Other responses	Strongly disagree
Changes were forced by cut-backs	8	0
No changes forced by cut-backs	2	3
Total	10	3

Tau B = 0.683, significant at .001.

B. Was project slowed down?		
Cut-backs forced changes	Slowed down	Kept on schedule
Changes were forced by cut-backs	5	3
No changes forced by cut-backs	0	5
Total	5	8

Tau B = 0.688, significant at .001.

program's future is at risk, these results are expected and provide reassurance that the survey respondents' judgments of the projects are consistent.

The earlier LeanTEC research that has influenced the design of this investigation found that there were some ties between funding stability and project performance. Veteran professionals in the aerospace industry recalled a number of projects that were weakened by perceptions that project funds were limited or at risk. They suggested that when funding seemed threatened, development team engineers had a

Financial uncertainty and cutbacks are found to relate strongly to turnover.

tendency to migrate to other, more stable projects, causing turnover. Being able to bill to multiple engineering charge numbers gives the individual substantial security and control over his or her work if the primary project encounters financial cutbacks or is cancelled. Other interviewees suggested that worry about continued funding led management and team leaders to cut back on staffing or otherwise reduce costs to stretch the project out. Whatever the reasons, respondents were confident that they had seen a substantial number of projects where funding uncertainties had directly contributed to poor team performance due to team turnover and inadequate staffing.

TABLE 4. FUNDING INSTABILITY AND STAFF TURN-OVER

A. There was uncertainly about future funding		
Members turned over?	Other responses	Strongly disagree
Other responses	6	0
Disagree, no turnover	1	6
Total	7	6

Tau B = 0.667, significant at .001. Missing data for one case.

B. Cut-backs forced changes?		
Members turned over?	Forced changes	No changes forced
Other responses	6	0
Strongly disagree, no turnover	2	5
Total	8	5

Tau B = 0.635, significant at .001.

The results from this present study of Army systems support that view. When one looks in turn at how these questions about financial stability relate to other key factors, strong, negative relationships are found with staffing stability and effective testing. In particular, both financial uncertainty and cutbacks are found to relate strongly to turnover. Of six projects where respondents reported little or no funding uncertainty, none are reported to have had turnover. For the remaining projects where funding was more uncertain, six of seven experienced turnover (top, Table 4). When one compares the projects that didn't have compromises or changes forced by cutbacks with those that did, the results show that all five projects with no cutbacks also had no turnover. By contrast, six of the eight projects that experienced cutbacks also experienced turnover (bottom, Table 4).

A NOTE ON STATISTICS

Despite the limitation that the analysis only includes 13 cases, the Tau B statistical test appropriate for these variables (Blalock, 1960) shows that the strength of the relationship between the two variables is sufficiently strong that it could have happened by chance less than one time in a thousand ($p < .001$). One can thus have confidence that the reported relationships are statistically meaningful despite the small number of programs being studied.

Project cutbacks and financial stability were also found to all relate to the effectiveness and the appropriate timing of the testing used in the program. One might readily understand that slowing projects could disrupt testing schedules, and that

TABLE 5. FUNDING INSTABILITY AND THE TIMING OF TESTING

A. Was project slowed down?		
Appropriate timing of testing?	Slowed down	Kept on schedule
Other responses	4	2
Strongly agree	1	6
Total	5	8

Tau B = 0.620, significant at .004.

B. Cut-backs forced changes		
Appropriate timing of testing?	Occurred	None
Other responses	5	1
Strongly agree	3	4
Total	8	5

Tau B = 0.610, significant at .002.

investment in testing might be a casualty of budget cuts. For projects slowed during the life of the program, four of five cases did not conduct timely testing activities. For projects that stayed on schedule, six of eight were reported as having conducted appropriate testing (top, Table 5). Cutbacks could cause testing changes and compromises, which could lead to the conclusion that testing was negatively affected. Four out of five cases that did not have any changes forced by cutbacks also had appropriate testing; only three of eight programs that were slowed are reported confidently as having appropriately timed their testing activities (bottom, Table 5).

Experience suggests that stretching projects disrupts schedules, and that cutbacks and changes often lead to the need to repeat old test procedures or design new ones.

Experience suggests that stretching projects disrupts schedules, and that cutbacks and changes often lead to the need to repeat old test procedures or design new ones. That and the presence of turnover mean that revised testing programs are sometimes prepared by different individuals from those who designed the system and supported its integration. Whatever the mechanisms, the general conclusion from these results is that funding instability seriously affects staffing and the quality of testing, which were in turn shown in the research to be key predictors of program performance.

TABLE 6. FINANCIAL INSTABILITY AND PROJECT PERFORMANCE

Average number of successful outcomes and number			
	Other responses	Positive response*	Signif. at
Stability and funding:			
Uncertainty about project funding?	2.80 (10)	5.67 (3)	.001
Project ever slowed down?	1.80 (5)	4.50 (8)	.001
Turn-over on development team?	2.00 (7)	4.71 (6)	.001

**The positive responses are, in order: to disagree strongly that funding was uncertain, to say the project was never slowed or interrupted, and disagree that there was any team turnover.*

The overall effect of financial-related problems on project performance is summarized by looking at the average number of successful outcomes (Table 6). The six cases that had little or no problem with uncertain funding had an average of 5.67 successes out of a possible 6.0; those that did face this uncertainty averaged 2.80. Cases that were never slowed averaged 4.50 successes compared with 1.80 for those that were slowed. The impact on outcomes of the relationship between the funding environment and staffing is also seen; projects that experienced turnover averaged 2.0 successes, while those that avoided turnover averaged 4.71. It might be argued that financial uncertainty and cutbacks follow when projects encounter other difficulties, in which case one could expect that these differences in averages would be higher than those found for other factors. Nevertheless, when funding problems are present, there is little doubt that they are strongly associated with turnover and poor development performance for these DESERT STORM cases.

SIGNIFICANT CHANGES IN SYSTEMS REQUIREMENTS

As with other forms of instability, considerable anecdotal evidence suggests that significant changes in systems requirements will adversely impact program outcomes, particularly schedule and/or cost. Consequently, the existence of this evidence makes experienced project managers extremely wary of permitting any changes in system requirements to occur. Sometimes, however, actions on the part of potential adversaries, referred to as “changes in the threat,” can force the issue.

In the 13 cases studied, only 3 reported no change to system requirements once a system concept had evolved, and only 4 reported no change during the development phase of the project. In those cases that experienced change during development, three were judged to have required “significant” or “major” effort to make the change, while the remaining six only required “minor” or “very minor” effort. Moreover, some cases experienced multiple instances of requirements change dur-

TABLE 7. CHANGING SYSTEMS REQUIREMENTS AND PROJECT OUTCOMES

A. Frequency of changes in systems requirements		
System met cost goals?	Several, or Many times	None, or One or two times
Fell far short of cost goals	1	0
Came close to cost goals	3	3
Met or exceeded cost goals	0	6
Total	4	9

Tau B = 0.620, significant at .003.

B. Frequency of changes in system requirements		
Late engineering changes after production had started	Several, or Many times	None, or One or Two times
Significant changes	2	0
Minor changes	2	6
None, almost none	0	3
Total	4	9

Tau B = 0.537, significant at .004.

ing development, with four of the nine describing encountering "several" or "many" changes. The remainder reported "no change," or only "one or two instances" of change. The frequency of these changes seems to be at variance with the relatively stable doctrinal and operational environment prior to DESERT STORM.

When changing requirements are related to other variables, the results support the conventional wisdom that such changes are costly.

When changing requirements are related to other variables, the results support the conventional wisdom that such changes are costly. Significant correlations were found between three requirements change variables and several of the outcome metrics. None of the four projects that had several (three cases) or many (one case) requirements changes met their cost goals (top, Table 7), and none of the four avoided

TABLE 8. STABILITY OF SYSTEMS REQUIREMENTS AND PROJECT PERFORMANCE

Average number of successful outcomes			
Mid-development requirements change?	Changes occurred	No changes	Signif. at
	2.00 (4)	4.11 (9)	.005
Frequency of systems requirements change?	Several or Many times	None, or One or two times	Signif. at
	1.50 (4)	4.33 (9)	.006

late engineering changes (bottom, Table 7). For those that had none or only one or two systems requirements changes, six of nine met their cost goals, and three of nine avoided (even minor) late engineering changes. Weaker but similar negative differences on other outcomes (not shown) are found among those projects that experienced more requirements changes.

One can see the overall effects of the frequency and timing of systems requirements changes by looking at how many successful outcomes these projects had on average. The respondents were asked in what stage of development the requirements had changed, and it was found that the negative impact on the average rates of success was greatest when changes had occurred in mid-development (typically after the Critical Design Review). As shown in Table 8, the four projects that had systems requirement changes in mid-development had only an average of 2.0 positive performance outcomes, compared to 4.11 average positive outcomes of those that did not change in mid-development.

The impact is even greater when one looks at the frequency of requirements changes and the average number of successful outcomes. The four projects said to have seen systems requirements changes several or many times during development averaged only 1.50 successful outcomes, compared to 4.33 successes among those projects that had no or only one or two systems requirements changes.

TURNOVER IN KEY USER PERSONNEL

The U.S. Army Training and Doctrine Command (TRADOC) is responsible for determining the requirements that Army materiel must meet in order to have utility on the battlefield. A senior TRADOC staff member (typically a colonel) serves as the alter ego of the Project Manager in interpreting these requirements as they are translated into system technical requirements during the acquisition process. This key individual may also play a critical role in preserving the planned funding for the system development by persuading more senior TRADOC leaders to strongly reaffirm the need for the system when budget cuts are threatened or problems are encountered in the system development that increase cost or stretch schedule. The frequency and timing of turnover in key TRADOC personnel were examined to determine influence on project outcomes.

TABLE 9. CONTINUITY OF TRADOC STAFFING

Did TRADOC change during early development?		
Operational problems in the field?	Yes	No
Field problems limited effectiveness	4	1
Deployed at no loss of effectiveness	1	4
Exceeded expectations	0	7
Total	5	7

Tau B = 0.630, significant at .001. Data are available on 12 cases for this TRADOC question.

All of the respondents reported that their projects had experienced some turnover in key TRADOC personnel. Only two reported no key TRADOC personnel changes during the development phase of the project. Such regular change is consistent with the military reassignment cycle. The timing of when TRADOC key personnel turnover occurred correlated with several of the outcome metrics, most notably with the extent to which the system met expectations when used on the battlefield during DESERT STORM. Table 9 shows the negative impact of staff change early in development on system performance on the battlefield. Key TRADOC personnel changed for five cases during early development, and four of those projects subsequently encountered operational field problems. Where there was no early TRADOC change, only one of seven projects was not as effective as expected.

Cases that experienced no TRADOC changes in early development, however, are seen to be substantially more successful at an average of 4.29 successful outcomes.

The only two cases—Night Sight and the M829A1 sabot round—that did not experience TRADOC changes during the development phase are the same two cases that the respondents felt exceeded performance expectations in the field.

To further examine the possible impact of TRADOC changes at different stages of the project, one can again look at the average number of goals met for the different projects. No relationships of consequence are found between TRADOC change in mid- and late development and the average number of successful outcomes. Cases that experienced no TRADOC changes in early development, however, are seen to be substantially more successful at an average of 4.29 successful outcomes, compared to an average of 2.40 for those that did have TRADOC changes at that time.

TABLE 10. TRADOC CHANGES AND PROJECT PERFORMANCE

Average number of successful outcomes			
TRADOC change during early development?	Staff change	No change	Signif. at
	2.40 (5)	4.29 (7)	.011

TABLE 11. CHANGING SYSTEMS REQUIREMENTS AND KEY TRADOC PERSONNEL

TRADOC change in early stages* of project?			
Did system requirements change during mid development?	No change	In one stage	In both stages
No	5	3	0
Yes	0	2	2
Total	5	5	2

Tau B = 0.0685, significant at 0.001 **pre-development or early development*

One would expect that the doctrinal and operational underpinning for the systems' requirements should have been relatively constant (during this late Cold War period), but the authors found a substantial number of changes in systems requirements for these cases. The question, therefore, lingers: Did the changes in TRADOC personnel found here somehow play a role? New personnel may not feel that their predecessors had correctly defined the threat's implications on requirements, or having not been a party to earlier discussions, been more willing to pursue changes suggested by new

The most damaging requirements changes are those that occur in mid-development.

knowledge and events. This suggestion raises the possibility that TRADOC personnel change could have adverse, indirect effects by somehow permitting changes in systems requirements that have, in turn, a negative impact on project performance.

This study analyzed the relationship between early TRADOC staff changes and shifts in systems requirements and found some support for that view. As noted previously, the results in Table 10 suggest that the most damaging requirements changes

TABLE 12. PROJECT DURATION, INSTABILITY AND FIELD PERFORMANCE

How often did systems requirements change?	Three years or less	Four years or more
Never	4	0
Once or twice	3	2
Several or many times	0	4
Total	7	6

Tau B = .723, significant at .001.

Did system have operational problems in the field?		
Did not meet expectations	1	4
Met expectations	4	2
Exceeded expectations	2	0
Total	7	6

Tau B = -.556, significant at .003.

are those that occur in mid-development. To look for a relevant relationship, one can aggregate TRADOC changes by asking whether there were key staff changes during the technology selection/planning period or during early development (the two earliest time slices in the project). One can then compare the projects that did or did not have these early key TRADOC personnel changes. Table 11 shows that five projects avoided early TRADOC changes, and none of the five had mid-development changes in systems requirements. Of the five projects that had TRADOC changes in one of these early stages, two experienced requirements changes. The two projects that had TRADOC changes in both planning and early development saw later requirements changes in mid-development. It would appear that TRADOC turnover in the early stages of projects is in some way related to mid-development changes in systems requirements.

CONCLUSIONS

Taken together, these several relationships strongly suggest that stability of program resources, staffing, and objectives is a very powerful influence on the relative success of projects. Certainly, as has been noted, a wealth of anecdotal evidence suggests that this should be the case. In reflecting on this array of instabilities that could impact a system development, it became clear that they had at least one thing in common. That is, the longer a system stayed in development, the greater chance it had to experience one or more of these program destabilizing events.

Support for this view is found when the occurrence of cutbacks and systems requirements changes are related to the duration of the 13 development projects studied

TABLE 13. LENGTH OF PROJECT DEVELOPMENT AND PROJECT PERFORMANCE

Average number of successful outcomes			
Project duration?	Three years or less	Four years or more	Signif. at
	4.71	2.00	.001

here. The projects divided cleanly between 7 that took 37 months or less, and 6 that took 48 months or more. When the frequency of systems requirements changes is considered, longer duration clearly allows more time for changing external conditions and priorities to lead to changes in systems requirements. Four of seven of the shorter projects never experienced any changes in requirements, and the remaining three only experienced minimal changes. None of the six projects that ran four years or longer avoided requirements changes, two saw changes once or twice, and four experienced either several or many changes. The pattern is quite strong and statistically significant despite the limited number of cases (top, Table 12).

Given that longer schedules increase the risk of encountering instabilities, it is then not surprising that project duration is also negatively related to achieving desirable project outcomes. For example (bottom, Table 12), one of the strongest relationships between development duration and project outcomes is found for how the system performed in the field. Six of the seven systems developed in (about) three years or less met or exceeded expectations when they were deployed in DESERT STORM, compared to only two of six of the longer projects. The effect on all six outcome questions is seen in Table 13 with the projects requiring longer development time averaging success on only two outcomes, where the shorter development projects had an average 4.7 successful outcomes.

The sensitivity of this central conclusion to project complexity was examined using a measure of relative complexity (Table 1) developed for this purpose. More complex projects often require longer development cycles and are more likely to experience funding difficulties. However, complexity is much more weakly related with staffing turnover than is project duration, and differences in complexity are not found to be at all related with changes of TRADOC personnel or changes in systems requirements. Complexity is also not related to testing quality and timeliness; both of these factors are strongly (and positively) correlated with outcomes. Importantly, duration alone is more strongly related than complexity to the number of successful outcomes. Whether or not complexity plays a role, project duration has a strong effect on outcomes independent of the influence of complexity.

The evidence here is that time is not an ally of systems development. The passage of time and the inevitable intrusion of new knowledge open the door to new pressures. Financial uncertainty is created in part by the need for resources for newer projects, often leading to staffing doubts about the current program and/or the appearance of new opportunities where key personnel are also needed. Or, simple career progression incentives lead people after time to move on, taking knowledge and experience with them. An obvious conclusion is that programs undertaken for around three years or less involve less risk of destabilizing factors outside the control of program managers. These data then suggest that it is better to organize and budget projects for shorter time frames whenever possible. If projects must be planned to take four years or longer in development, this research suggests that one should recognize that longer projects open the way for substantial pressures that may be difficult to resist, and then plan to deal with instability. Of the several factors considered in this study, the areas that should receive particular attention in longer projects are the need not to take advantage of the expanded opportunity to change systems requirements, to keep key user representatives in their positions during the critical early program phases, to provide incentives and career planning to avoid turnover in key development personnel, and to ensure that cutbacks and rescheduling do not compromise testing regimes.

Dr. William A. "Bill" Lucas is executive director of the Cambridge-MIT Institute (CMI), an MIT and Cambridge University partnership. His research includes the study of knowledge exchange in cross-functional development teams in industry and in university-industry collaborations, and evaluation research on programs for young professionals who will become leaders of innovation. His work before MIT included university teaching, The Rand Corporation, government service, and time in the telecommunications industry.

(E-mail address: walucas@mit.edu)

Dr. Richard G. "Dick" Rhoades is director of the University of Alabama in Huntsville's (UAH) Research Institute and professor of Engineering Management. His current research focuses on weapon system technical risk assessment and avoidance, propulsion system analysis, strategic planning, and organizational design. Prior to joining UAH, Dr. Rhoades held numerous positions in the Missile Research, Development and Engineering Center at Redstone Arsenal, including three Senior Executive Service positions.

(E-mail address: rhoadesr@email.uah.edu)

AUTHOR **BIOGRAPHY**

REFERENCES

Air Force Research Laboratory. (2002). *Lean transition of emerging industrial capability (LeanTEC)*. ARFL-ML-WP-TR-2002-4191. Wright-Patterson Air Force Base, OH: Author.

Blalock, H. M. (1960). *Social statistics.* NY: McGraw-Hill Books.

General Accounting Office. (1999, July). *Better management of technology development can improve weapon system outcomes.* GAO/NSIAD-99-162. Washington, DC: Author.

General Accounting Office. (2000, July). *A more constructive test approach is key to better weapon systems.* GAO/NSIAD-00-199. Washington, DC: Author.

Lucas, W. A., & Rhoades, R. G. (2004, June). *Lessons from Army system developments.* UAH RI-2004-1. Huntsville: The University of Alabama.