2014 NDIA GROUND VEHICLE SYSTEMS ENGINEERING AND TECHNOLOGY SYMPOSIUM
SYSTEMS ENGINEERING (SE) TECHNICAL SESSION
AUGUST 12-14, 2014 - NOVI, MICHIGAN

RISK LEADING INDICATORS FOR DOD ACQUISITION PROGRAMS

Gary Witus, PhD
Walter Bryzik, PhD
Mechanical Engineering
Wayne State University
Detroit, MI

Edward Umpfenbach, PhD
Rebecca Addis
Jerome Tzau
Kadry Rizk
US Army TARDEC
Warren, MI

ABSTRACT

What does "exposure to risk" mean? How can acquisition programs get early warning of risk exposure? How is risk exposure related to the root causes and causal mechanisms of adverse program outcomes? How does risk early warning inform risk management? How is risk exposure related to the tradeoffs made between risk versus potential rewards? What technical and management contract data reporting requirements provide evidence of risk exposure, and how can risk leading indicators be computed? How can standard technical and management contract data reporting requirements be used to improve visibility into risk exposure? How can the magnitude of risk exposure be estimated? How does risk early warning complement traditional technical, cost and schedule risk assessment? How do risk early warning methods relate to typical proposal requirements and evaluation criteria? How are risk leading indicators related to system development leading indicators? How can risk early warning methods be verified and validated?

INTRODUCTION

Risk exposure refers to program conditions that amplify the likelihood and/or consequences of unforeseen future events and interactions, and of normal variances in activity time, cost and performance. Adverse consequences include development time and cost overrun, technical performance and reliability shortfall, and excessive production, operation, and sustainment costs. Risk exposure is the result of unrecognized or unacknowledged past events, decisions and actions, and current uncertainty, inconsistency, incompleteness, and interdependency in the system development.

Risk exposure early warning refers to detecting evidence of the root causes and effects before there are significant adverse consequences. Its purpose is to alert program management to areas of elevated risk exposure in the program. Risk exposure early warning combines cost, schedule and system development data in an integrated view.

The methods and tools described in this report are focused on Technology Development (TD) and Engineering and Materiel Development (EMD) acquisition phases, and specifically TD and EMD contractor activity. A parallel activity is underway to identify sources and indicators of risk injected into the acquisition program prior to contract award, but is outside the scope of this paper.

Risk exposure early warning complements the risk identification practices and procedures in the DoD Risk Management Guide [1]. The intent of the Guide is "*to help ensure program cost, schedule, and performance objectives are achieved at every stage in the life cycle"* and to present processes "*for uncovering, determining the scope of, and managing program uncertainties.*"

The Risk Management Guide defines a risk as a potential future event which, if it occurs, would have adverse consequences, for which the probability that it will occur and the consequences, if it occurs, can be assessed. Risks with high combined likelihood and magnitude are priorities for tracking and mitigation. The practices and procedures in the guide start with identifying risk events. Risk exposure early warning does not itself identify causes or mitigations. It detects conditions of elevated exposure to risk and it points to evidence to help orient risk and issue management investigation.

The procedures and practices in the Guide rely on Subject Matter Experts (SMEs) to identify and quantify risk events. Risk exposure early warning provides an evidence-based analytic approach that complements SME insight. The Air Force Cost Risk Handbook [2] identifies common factors leading to bias and dispersion in SME estimates of time, cost, technical performance, and of identification and estimation of likelihood and consequences of risk events. It lists eight motivational factors (management pressure, social pressure, group think, wishful thinking, career goals, misunderstanding, project advocacy, and competitive pressures) and five cognitive bias factors (inconsistency over time, anchoring, irrelevant analogies, underestimation, and human nature).

Risk exposure early warning is based on understanding the causal chains from root causes to adverse outcomes, and how the effects manifest in standard program and system development data and reports.

Risk exposure early warning has three major components: Risk Leading Indicators (RLI), outlier and cluster analysis to detect program areas with elevated relative risk exposure, and Risk Estimating Relationships (RER).

RLI are computed from standard program management and systems engineering reports. RLI compare data across different reporting requirements and over time to measure incompleteness and inconsistency, instability, uneven or inadequate progress, and interdependency of the program and system organization. The RLI are evidence of both (1) root cause problems with potential for persistent future effects, and (2) conditions that increase the potential for, and/or sensitivity to, unforeseen future events and execution-versus-plan variances. The RLI were developed through examination of risk considerations in proposal evaluation and program execution criteria and reporting, published root cause analyses, best practices metrics for schedule risk analysis, and published research on system development leading indicators.

Risk Estimating Relationships (RER) are statistical models that use the RLI to estimate future bias and uncertainty in program activities time, cost and technical performance. The RER are calibrated to data from completed Integrated Master Schedule (IMS) activities of the current program, and are updated over time. Each completed activity in the IMS provides a data point.

TECHNICAL APPROACH

We reviewed prior analyses of root causes of adverse program outcome. We identified Program Management Office (PMO) risk issues and perspectives expressed and implied in the evaluation criteria and reporting requirements in Request for Proposal (RFP) packages over several ground vehicle programs. We identified significant causal mechanisms leading to adverse acquisition outcome, based on our experience over many ground vehicle programs, and program documentation. We identified potential sources of data and evidence in program management and systems engineering baseline and update reports. We reviewed reports on system development leading indicators, and reviewed metrics in recommended procedures and "best practices" for program, cost and schedule management. We adapted applicable indicators, in the light of sources of data and evidence, to define an initial set of risk leading indicators. The risk leading indicators are computable from data in typical program management and systems engineering reports and updates. We developed recommendations regarding the timing, content and resolution of standard baselines and contract data reporting requirements that would enhance their value for timely and effective risk early warning and diagnosis. We outlined a procedure to calibrate RERs for individual development programs.

BACKGROUND

Root Cause Analyses

The Government Accountability Office (GAO) [3] reported that in fiscal year 2012, of the 85 major defense acquisition programs under review, 39 percent had unit cost growth of 25 percent or more, the average delay in initial operating capability was 27 months, the average change in development cost from the initial estimate was 49 percent, and the average change in total acquisition cost from the initial estimate was 38 percent. The report finds that programs experiencing cost and schedule growth "*share a common dynamic: moving forward with programs before the knowledge needed to make decisions is sufficient*." The report identifies seven common root causes: (1) concurrent testing and production, (2) optimistic assumptions, (3) delayed testing, (4) insufficient tradeoffs among cost, schedule, and technical performance requirements during early planning, (5) unrealistic cost and schedule estimates, (6) Insufficient testing during development, (7) insufficient attention to reliability.

The GAO [4] identified 12 root causes: unstable program requirements, funding and quantities; complex systems; diminishing industrial base; new processes; immature or cutting edge technologies; first time integration; unrealistic assumptions and projections; overoptimistic program baselines; inexperienced staff; lack of relevant historical data; unreliable Earned Value Management (EVM) data; and inadequate contingency schedule slack and management reserve funds.

The Office of Performance Assessments and Root Cause Analyses (PARCA) [5] found five root causes of adverse acquisition outcomes attributable to planning and execution: (1) unrealistic cost or schedule estimates, (2) immature technology with excessive manufacturing and integration risk, (3) unrealistic performance expectations, (4) unanticipated design, engineering, manufacturing or technology issues, and (5) poor management performance. Poor management performance included ambiguities combining requirements and requirements documents, interface management, tradeoffs within holistic performance attributes (size, weight, power, cooling, etc.), and risk assessment.

In 2008, the NDIA Systems Engineering Division in conjunction with DASD-ATL-SE produced a report on the systemic root causes of program failures [6] concluding that "*the most significant causes were directly related to poor or inadequate activities early in acquisition strategizing and planning efforts and in conducting management gate reviews during the early stages of execution. Lastly, the analysis also concluded that there was a significant root cause related to staff size, training and experience*."

The Defense-Industrial Initiatives Group at the Center for Strategic & International Studies [7] found that of 85 major programs, overoptimistic estimates were the primary driver for cost growth and changes in quantity were the second leading cause. A RAND study [8] found that in their analysis of 35 programs, changes in quantity accounted for more than half of cost growth, and that "*decisions to change the schedule, additional requirements, and cost-estimating errors account for almost all of the remaining procurement cost growth*." The Institute for Defense Analysis [9] found that "*Virtually every program we surveyed experienced cost problems that could have been avoided or ameliorated through better front-end analysis of overall design issues and risks. Serious attention to system-level risk seems to have been lacking on the part of senior decision makers*."

In summary, root causes include: unstable requirements; incomplete or unstable planning; overoptimistic or inaccurate estimating; underappreciated or misunderstood technical and engineering challenges; deferred or insufficient verification and testing during development; unforeseen interactions and interdependencies in the system requirements, architecture and design; deferred tradeoff decisions; unforeseen interactions and interdependencies in the execution task schedule or organization; lack of "margin for error" in budget, schedule, performance and design; data quality and availability of relevant historical data.

Proposal Evaluation and Contract Management

Fair, equal and open competition guidelines require that RFP packages clearly state (1) what information to provide in the proposal (section L), (2) how the proposals will be evaluated (section M), (3) the Scope of Work (section C), and the Contract Deliverable Requirements List (CDRL; typically Appendix A). The RFP package typically includes attachments that specify documentation frameworks and criteria. The CDRL section specifies requirements for reports and updates on periodic or event-based timelines, and the required format and content. Baselines are required either with the proposal, or at a specific subsequent technical review event. The RFP package typically includes guidance for duration of TD and EMD phases, funds available for each phase, number of EMD prototypes, number of Low Rate Initial Production units, etc. The RFP package also includes targets for performance (requirements) and constraints.

Section M specifies how the Government will evaluate the risk-vs-reward to reject unsuitable bids, and to select from among suitable bids. Risks are the risks of failing to deliver prototypes that will pass Operational Testing (per specified failure definitions and scoring criteria) for performance and reliability on schedule and within development budget, of failing to be ready for Low Rate Initial Production (LRIP) at the end of the EMD, of failing to meet LRIP unit cost targets, and/or failing to be able to meet fuel economy, reliability and logistics support goals. Rewards are the proposed time, costs, and system performance. Risk-reward tradeoffs are considered both in proposal evaluation and in program management decisions to trade lower performance for lower cost-schedule risk, for increased design tradespace, and/or for increased reliability.

The level of risk in the proposed program is judged based on the claimed level of design, manufacturing, production cost, and RAM maturity, considering the thoroughness and credibility of the supporting data. Maturity levels are clearly defined with specific completion criteria based on systems engineering, design, integration, analysis, manufacturing, and testing artifacts. The specific criteria for maturity levels combine essential technical performance measures, systems engineering and design technical review knowledge points, and verification results. The highest levels of maturity correspond to completion of the contract requirements with tests completed on manufactured systems, with substantiating cost data and correct action plans for any deficiencies.

Maturity level advancements are progress accomplishments on the path to successful program completion. Using maturity levels to track technical progress is a well-defined and is consistent with the PMO framework for program risk assessment. Maturity level advancement vice the program plan is a potential indicator of risk. Maturity level advancement, over the entire program and by WBS element, can be integrated into the scheduling and reporting framework simply by including maturity advancement in the IMP, so that IMP events, accomplishments and criteria include maturity advancement of the design, manufacturing, and RAM. Making maturity advancement events part of the IMP forces cost and schedule reporting to be reported relative to demonstrated maturity. Maturity levels are organized by stages of system acquisition: development, production, and operation & sustainment. Maturity level assessments take a life cycle cost and capability view. Technology, integration, and manufacturing readiness views and consistent perspectives, focused on TD and EMD acquisition stages.

The RFP package includes additional requirements for risk assessment in proposal evaluation and contract execution. Executing a risk management plan, per the DoD Risk Management Guide, is required to identify and address potential future events with significant likelihood of occurrence and significant consequences if they do. Technology Readiness Assessment (TRA) [10] is required for those technologies designated as Critical Technology Elements (CTE) and Other Technologies of Interest (OTI). TRA involves detailed, engineering-level analysis by Subject Matter Experts. It is focused on the development of selected technologies and subsystems, not the entire development program. Integration Readiness Assessment (IRL) and Manufacturing Readiness Assessment (MRL) are not always required.

Technical review checklists contain over 800 specific questions in 13 categories to gauge progress at each of the major program reviews. Progress for each question is scored red, amber, green, unknown, or not applicable. A team at RAND has proposed using the technical review checklists as the basis for risk assessment [11]. The categories are review entry readiness, planning, schedule, management, staffing, process, product support, requirements management, system design, system verification, program risk assessment, certification and legal, and review completion.

Schedule risk assessment, e.g., using the Defense Contract Management Agency (DCMA) 14-point schedule assessment, GAO schedule assessment or similar approach, is typically required. Methods for schedule risk assessment, and the understanding of critical paths, near-critical paths and high schedule risk activities in non-deterministic programs, are evolving.

The RFP package specifies the program management and system engineering baselines and update reports, including content, format, and frequency or event timing. These provide the basis to evaluate risk leading indicators and to calibrate risk estimating relationships.

System Development Leading Indicators

The NDIA System Engineering Division [12] found that "*Technical decision makers do not have the right information & insight at the right time to support informed & proactive decision making or may not act on all the technical information available to ensure effective & efficient program planning, management & execution*". The NDIA formed a working group to develop a set of system development leading indicators to provide insight into technical performance at major decision points for managing programs quantitatively across their life cycle, with emphasis on TD and EMD phases, and objective measures of commonly and readily available data.

The NDIA project used surveys to identify high-value areas for system development leading indicators, building on prior work on systems engineering leading indicators [13]. The system development leading indicator categories [14, 15] were:

1. Requirements Stability
2. Proportion of Stakeholder Needs Met and Verified
3. Interface Completion Trends
4. Staffing Skills and Trend
5. Risk Burndown
6. Technical Performance Measure (TPM) Trends
7. Technology Readiness Level
8. Manufacturing Readiness Level
9. Architecture
10. Affordability
11. Requirements Verification
12. Defects and Errors

Specific leading indicators were defined in some of the categories. The indicators were not directly tied to specific contract data requirements. Typical contract data requirements relate to categories #1, 2, 5, 7, 8, 9, 10 & 11. Proposal evaluation and contract management criteria suggest additional indicators, supported by specific technical status and progress reporting, and directly related to program risk considerations in program decisions.

The NDIA survey referenced Kohl and Carson [16] reporting on a Practical Software & System Measurement workshop on architecture measurement concepts. Although the workshop focused on software and system architecture, the concept of architecture also applies to the program architecture as expressed in the IMS, to the requirements architecture, and other knowledge structures in program management and systems engineering. The workshop consensus identified six dimensions of architecture development for measurement: size, interconnectedness, completeness, compliance, consistency, and cost. "Stability" was not included. Sources of input data and methods to calculate measures from the data were not specified.

The Defense Contract Management Agency (DCMA) approach to technical performance risk assessment evaluates variances of technical progress versus cost and schedule to indicate the level of risk and detect new risks before their effects on cost/schedule are irrevocable [17, 18]. Technical progress is measured by the Technical Performance Measures (TPM) for the Technical Performance Parameters (TPP). Cost and schedule are measured with the Earned Value Management (EVM) system. The TPP and TPM are formulated for the particular program, and are derived from the major system performance parameters. In this model, TPM progress against the plan is the basis for leading indicators of risk.

Best Practices for Schedule Risk Assessment

The GAO schedule risk assessment guide [19] addresses measurement of schedule health with quantitative approaches to analyze the completeness, consistency, interdependency, safety margins in a program schedule. The metrics include: the number and proportion of "long" activities; ratio of activities to dependency links in total and by each of the four types of logical dependency; ratio of detailed activities to milestones; number and proportion of activities no mapped to a milestone or Integrated Master Plan (IMP) event; number and proportion of activities with many predecessor links, with many successor links, with no successor, and with no predecessor; critical path float to each milestone; and number of activities with negative or low float relative to their planned duration. The guide also includes of schedule execution, e.g., number of activities that were started or finished before their logical dependency condition, number of activities that started or finished late, mean and standard deviation of the difference between actual and planned duration, start date, and end date.

The GAO Guide recommends using probabilistic schedule risk analysis to complement deterministic critical path analysis. Probabilistic risk analysis requires data on the distribution of activity duration, e.g. the minimum, most likely, and maximum, and data on the correlations between activity durations. Simulation is used to compute the probability distribution of total time. The probability the program will be late or milestone missed, and the expected amount late given it is late, are computed from the distribution. However, no published data has been found showing that people can reliably estimate the distribution of activity durations or the correlations between activities. A detailed IMS can have upwards of 5,000 activities, and providing probability distribution and correlation estimates can be onerous. Probabilistic risk analysis does not directly identify which activities are putting the schedule most at risk. Further development and verification of practical methods for schedule risk analysis are needed.

Statistical Estimating Relationships

The concept of RERs was inspired by the statistical approach to Cost Estimating Relationships (CERs). CERs estimate cost by calibrating a model of historical cost data as a function of a set of explanatory factors [2, 20, 21]. The proportion of cost variance explained by a CER is a measure of the accuracy of the model. Open questions in developing CERs include: choosing the population of "similar" cases to pool together; choosing the explanatory factors; choosing the general analytic model type. These choices are interrelated. Larger pools of more diverse programs may provide more or less consistent evidence of different interactions and dependencies. When the ratio of data points to explanatory variables and model parameters is low, there is a risk of spurious correlation. Some explanatory variables may be highly correlated, or correlated in some programs but not others. Data from the program of interest may be sparse, but is highly relevant.

The significant issues in formulating CERs are (1) choice of the systems to pool as a population of similar cases, balancing population size and diversity, (2) choice of the explanatory factors, i.e., independent variables, (3) choice of the underlying regression model framework. Of these three issues, the first two are the most important. A diverse population risks attenuating and obscuring the effects of individual factors and interaction effects. A small population risks biases from small sample size. There are many different effective regression model frameworks including parametric multi-linear and non-linear regression, Artificial Neural Networks, Bayes network models, and Aggregate One Dependence Estimators (a powerful and efficient extension of naïve Bayes models to a family of weighted one-dependence models). These are all useful statistical methods to quantify the relationship and uncertainty between observable input factors and outcomes, calibrated to historical evidence.

Adverse acquisition outcomes can be on multiple dimensions, e.g., development time and cost, production cost, unit performance and reliability, etc. Both causes and outcomes may be positively correlated or anti-correlated. Program management needs to know "what and how" not just "how bad".

Adverse outcomes result from combined bias and uncertainty between program plan estimates and activity outcomes. RERs explain the accuracy of time/cost/performance planned versus actual outcomes. Accuracy has two components: (a) bias or offset, and (b) random dispersion or uncertainty.

CERs assume some degree of similarity between different programs to estimate cost. Aggregating over dissimilar programs can obscure and bias cost estimates for any specific program. Pooling multiple programs is needed to reduce statistical uncertainty, but overbroad pooling increases uncertainty [22].

Risk sources may differ from program to program, from acquisition stage to stage, and between WBS elements. Sources of risk exposure can be highly varied between different individual programs of the same type due to a wide variety of factors including marketing strategy, experience of the engineering and engineering management team, aggressiveness of the cost, schedule and performance goals, technology, engineering, and integration challenges, etc. To obtain relevant statistical evidence RERs should calibrated to the individual program. To obtain statistically valid samples, calibration data points should be at the lowest level of the IMS with time and cost reporting – each completed IMS activity is data point, conditioned on being part of the same program.

There are subtle statistical issues in aggregation over disparate programs and different parts of the same program, and statistical analysis of multiple-input-X-multiple-output relationships. Statistical issues should be considered in the light of qualitative understanding of root causes of adverse/unpredictable outcomes, the mechanism of effects, and evidence from program artifacts along the causal chain.

FINDINGS AND RESULTS

Causal Mechanisms

This section presents a synthesis of the published analyses on root causes, acquisition, program and contract management, and system development, augmented by first-hand experience on acquisition program execution and Independent Review Teams. The observations in this section are the rationale for the choices of risk leading indicators and the integrated risk early warning approach.

Optimistic Estimates Have Real Consequences. "Success oriented" program plans create skewed distributions with long tails: appealing planned results, but with greater adverse consequences when problems occur. In an effort to "sell" a program and/or to win a bid, senior management can be tempted to make optimistic claims for time, cost, technical performance, reliability, and potential risks. Since adverse acquisition outcomes are deficiencies relative to how the program was sold, optimistic claims increase risk by reducing the margin for error and uncertainty.

Beyond this "statistical" effect, optimistic estimates have real effects on the program plans that lead to elevated risk exposure. An optimistic timeline leads to aggressive schedule structures. Characteristics of an aggressive schedule structure include: concurrent (parallel but independent) paths that come together towards the end; limited incremental integration analysis, verification and testing; low schedule slack margin for error; and time estimates that are more optimistic for activities later in the schedule. Programs concurrent paths and limited intermediate integration and verification ("it all comes together at the end") do not produce the information to detect and correct problems until it is too late. Aggressive cost goals lead to reduced and deferred developmental analysis and testing, and to eliminating parallel execution of alternative backup approaches. Aggressive performance goals lead to adopting less mature advanced technologies that are more likely to have unforeseen integration issues. Aggressive schedules lead mid-level managers and engineers to take shortcuts. Over optimistic goals have real effects that elevate exposure to risk.

Allowing Margin For Error Reduces Risk Exposure. Limited margin for error creates exposure to risk. When there is little or no safety margin, small unforeseen events and "natural" variances due to uncertainty can have amplified effects. When there is more safety margin, larger impacts are needed to produce adverse consequences. Schedule margin (also called slack or float) covers schedule slip are rework time. Cost margin (management reserve) covers unforeseen costs and gives management flexibility. Performance margin provides tolerance for unforeseen interactions that could degrade overall capability. Design margins for holistic system properties such as size, power, weight, cost, reliability, etc. provide tolerance for unforeseen growth in burdens. The GAO recognized the need for management reserve and schedule slack. Tradeoffs between performance levels and holistic burdens are recognized in program management and systems engineering. Contract award practices to recognize margins and uncertainty in cost, schedule and performance are evolving.

Past Performance Predicts Future Performance. The root causes of time and cost overruns and technical accomplishment shortfalls in IMS activities do not go away by themselves. If the planning was overoptimistic, if technical challenges were not well-understood, if the executing organization has internal problems, etc., and if there has been no underlying change, then the root causes and dynamics will still be at work, and similar patterns of bias and dispersion in the outcomes of completed activities relative to the plan are likely to show up in the future. In this situation lagging indicators become leading indicators.

Potential Risks Are Everywhere And Entangled. Every requirement that has not been verified, every schedule activity that has not been completed, every architecture element that has not been designed, integrated, and tested expose the program to risk. Decisions affecting uncertainty and accomplishment in one acquisition phase have impacts on other phases. Decisions that affect uncertainty and accomplishment in cost, schedule and technical performance are interrelated. Decisions and tradeoffs in one WBS or IMP element can have impacts on others. Lacking a model and data on these interactions creates ignorance that exposes the program to risk.

Evidence Reveals Risk Exposure. Program reports - baselines and updates of system development data, linked to program execution data – are useful to detect and diagnose risk exposure when they are timely, with sufficiently complete, consistent, and accurate content. Reporting and analysis of evidence has costs that must be considered relative to benefits, just as incremental integration and verification has costs and benefits. Risk early warning can leverage standard reports and work within the framework of standard contract data requirements. The standard reports and reporting schedules have been developed to inform particular program stage gates and PMO decisions, but were not specifically designed for integrated risk early warning. Integrate risk early warning could benefit from (1) update scheduling for proactive choices vice reactive assessments; and (2) standardized content, format, and terms across different artifacts to help ensure consistency and completeness. Standardized language for RFP packages is the mechanism for implementation. Contractors will benefit from standardized contract language by knowing what data content, information, and terms to provide, and how the data will be assessed. Increasing transparency to the offerors improves the quality of responses, and the ability to compare competitor responses. Objective data reporting and clear evaluation criteria are essential for open and equal competition among vendors

The Program Manager's Office (PMO) Establishes the Risk Tradeoff Terms and Conditions. The PMO is the definitive source for understanding priorities and tradeoffs among EMD cost, schedule, technical performance & reliability, production and O&S cost and risks of not meeting claims. The PMO determines the evaluation basis and criteria to compare the risk and reward of alternative proposals. The PMO specifies the priority levels for different performance parameters and constraints. The PMO decides how to trade off development time and cost versus initial performance and reliability versus production cost versus reliability growth and performance upgrade though continuous modernization versus operation and sustainment cost and logistics footprint. Risk exposure assessment must be consistent with the PMO value proposition to be relevant. The same risk priorities and considerations apply after contract award as during proposal evaluation, albeit with additional data. Risk exposure early warning must be informed by and consistent with the proposal evaluation criteria and the contract reporting/management criteria in the RFP package produced by the PMO.

Buried Tradeoffs & Constraints Indicate and Cause Risk Exposure. Sometimes programs have constraints and/or tradeoff relationships between one factor and another that have not been included in the product specification, e.g., size, power, weight, production cost, operational reliability and maintainability, continuous modernization capacity, etc. Sometimes increasing performance on one parameter can compensate for, or allow margin for, another. Decreasing performance goals in one area can increase the design constraint tradespace over the entire system. Increasing clarity of the constraints & tradeoffs reduces risk of offerors going "off in the wrong direction."

Instability Indicates and Causes Risk Exposure. Unstable requirements, plans, and architectures create re-work and wasted effort, and uncertain outcome. Instability can be evidence of inadequate planning and understanding the technical content, implying continued future instability. Changes create re-work, and indicate past planning deficiencies that will, if not corrected, lead to future incompatibilities and re-work.

Interdependency and Incompletely Resolved Elements Indicate and Cause Risk Exposure. Requirements, system architectures, and program schedules are all networks of interconnected nodes. Larger and more highly interdependent networks have more opportunities for unforeseen interactions, frictions, and ripple effects. There are more opportunities for adverse interactions if a new node or link is added. When nodes that are not fully resolved (e.g., a planning package in the schedule versus defined task, a requirement that has not been decomposed and linked, or an architecture element does not have completed boundary diagrams) are eventually resolved, there are more opportunities for adverse interactions in more complex networks. Unresolved elements are evidence of incomplete planning and/or understanding the system and/or program. Unresolved elements add uncertainty into estimates, and may have hidden dependencies. Unresolved elements can hide development obstacles and difficulties, and thus leading to overoptimistic estimates. The GAO recommends that no detailed activity be longer than 44 days, recognizing that long activities are evidence of incomplete resolution and that longer activities tend to have greater uncertainty. The GAO also recommends analyzing IMS network characteristics to assess schedule risk. Similar, though not identical, methods can apply to the requirements network, the system architecture, and even to the interconnected network of the requirements, IMP/IMS, and system architecture.

Compliance Verification and Developmental Testing Buy Down Risk. Less verification, incremental integration and testing during development creates more opportunity for unhappy surprises at the end of the program when time and funds for corrective actions are constrained. Earlier verification provides more time margin for corrective action. Verification and developmental testing require commitment of time and funding. Different means of verification include design inspection and engineering judgment, modeling and simulation, isolated bench testing, partially integrated testing, and fully integrated testing.

Technical Progress Needs Visibility. Reporting to reveal real technical progress is needed to assess risk. Without evidence, early warning is unfounded. Objective measures, e.g., development, production and RAM maturity by WBS element, provide diagnostics. Slower than expected technical progress indicates initial planning bias. Uneven technical progress indicates planning inaccuracy or underappreciated difficulty. Incomplete, inconsistent, missing or "to be determined" fields in program management and system engineering baselines and reports are is evidence of uncertainty, which could lead to new tasks, longer times, more coordination, etc.

Different Programs are Different. Different programs have different engineering and manufacturing challenges, different contractors, different engineering management practices, and different levels of skill and experience. Different programs will have different sources of risk. These differences make it difficult to extrapolate from one program to another. Risk Leading Indicators that were relevant to one program may not be relevant to another. Quantitative relationships between leading indicators and outcomes on one program may not be accurate for another program. Aggregating across disparate programs would dilute and obscure the significant relationships, as seen in empirical studies of Cost Estimating Relationships (CER) across different domains [22].

Proposal and Contract Data For RLI

Risk early warning employs baselines and updates for seven standard data reporting elements. Three are program management products: the Work Breakdown Structure (WBS), the Integrated Master Plan (IMP), and the Integrated Master Schedule (IMS). Five are systems engineering artifacts: the System Segment Specification (SSS), the Specification Tree (ST), the System Architecture (SA), Technical Review Checklists (TRC), and the Manufacturing Cost Estimate (MCE). The RFP package specifies the content, format, preparation instructions, initial delivery and update schedules.

The Government provides the initial WBS in the RFP package per MIL-STD-881C [23]. The contract WBS has detail added by the contractor, and is updated during the program. The WBS is the primary framework for program organization and reporting.

In risk exposure early warning, the WBS and the IMP are the key frameworks to identify areas of elevated risk. The initial IMP is prepared by the Government as part of the RFP package. The IMP is a 3-level indented list of events, accomplishments, and criteria. The IMP is the basis for IMS and Earned Value Management (EVM) reporting [24, 25]. EVM time, cost and progress reporting is at the "Work Package" level of the IMS.

The high-level IMS is provided by the Government in the RFP package. It contains the major program milestones dated from start of contract award, major program activities, and their dependencies. The contractor details the IMS activities to accomplish each event/accomplishment/criteria entry in the IMP. The IMS is updated monthly. At a minimum, the IMS has Work Packages for the next 12 months and Planning Packages thereafter. A Planning Package consists of 1 or more Work Packages. A Work Package consists of one or more detail tasks. Ideally a detail task has a singularized product, and defined completion criteria, although this is not always true in practice.

The "Work Package" and "Planning Package" identifiers link the schedule to EVM reporting. Reliability, validity, bias and accuracy of EVM reporting are concerns, as is the resolution of activity decomposition, EVM reporting, and completion criteria.

For each activity, the IMS contains the following data fields: parent WBS element, IMP entry, Planning or Work Package; activity scheduling logical dependency relationships (predecessor-to-successor Finish-to-Start, Start-to-Start, Finish-to-Finish, and Start-to-Finish dependencies); budgeted time, budgeted cost, actual time expended, slack time, actual cost expended, fraction of work performed. Some of this information comes from the EVM system. Level-of-effort tasks are not included in the IMS. The IMS is the basis for schedule and cost risk analysis. The analysis can be for the overall program, by WBS element, and/or by IMP entry. By including maturity advancement steps in the IMP, cost and schedule of maturity advancement is tracked.

The SSS begins with the performance specification (P-Spec), a part of the RFP package. The SSS contains derived requirements for system segments of the contract WBS. The SSS contains the following fields: the WBS element (level 3 or below); the parent SSS elements (there may be more than one); the singular property or characteristic; the threshold and objective criteria (performance levels and conditions of performance), priority level (e.g. "tiers"), pointers to the verification tasks in the IMS (null if no task accomplishes verification); verification results (null if the verification task has not been completed, else the performance measure results and test conditions); method of verification (e.g., design inspection, analysis, bench testing, integrated testing), compliance (compliant, partially compliant, not compliant); estimate of achievable performance; tradespace dimension and tradeoffs (holistic system attribute tradespace gained or lost if the requirements are changed to the achievable level; e.g., size, weight, power, cooling, production cost, and impact on other performance requirements); time and cost of verification if there is not a verification task. Ideally, the SSS addresses the costs of compliance verification.

Compliance verification can be at different levels, e.g., design acceptance, manufacturing and design acceptance, modeling and simulation of design and manufacturing, system integration laboratory testing, field testing etc.

The information in the SSS is the essential information for tracking compliance progress, and for performance requirements trades against time, cost, and design tradespace. It provides information for time and cost tradeoffs versus verification to reduce risk. It tracks back to the IMS to correlate compliance with cost and schedule.

The SSS would benefit from a prior definition of the holistic attribute tradespace dimensions, e.g., extracted from the Ground System Architecture Framework [26], and linking the dimensions to the subsytems that supply and consume each resource. It would also benefit from a formalized definition of the levels of verification matched to the levels of the maturity level definition, e.g., verification by design review, subsystem modeling and simulation, subsystem testing in a laboratory environment, integrated system modeling and simulation, integrated system testing in representative context, full system integrated testing in a realistic context and environment.

The ST contains the technical baseline as it is developed during the program. It specifies the functional baseline, allocated baseline and product baselines [27]. Data structures, file formats, data fields and formats have not been standardized. The functional baseline describes the functional and interface characteristics of the overall system, and the verification required to demonstrate their achievement. The allocated baseline defines the lower-level configuration items making up a system, and how system function and performance requirements are allocated across lower level configuration items, including design constraints and the verification required to demonstrate the traceability and achievement of specified functional, performance, and interface characteristics. The product baseline describes the functional and physical characteristics of a configuration item; the selected functional and physical characteristics designated for production acceptance testing; and tests necessary for deployment/installation, operation, support, training, and disposal of the configuration item. The initial product baseline includes "build-to" specifications for hardware (product, process, material specifications, engineering drawings, and other related data), and "code to" specifications for software. Verification of completeness of the technical baselines is normally reviewed at technical reviews, as specified in the IMP.

The SA defines the physical and logical system entities with boundary diagrams and behaviors. Boundary diagrams specify entity relationships with each other and the external environment, and are supplemented with formatted data describing the characteristics of each interface. Behaviors are derived by tracing operational scenarios, vignettes, mission threads, and/or use cases through system boundaries to derive internal system behavior culminating with a linked and traceable allocation of behavior for product design and development. Specific guidelines for the system architecture and design documentation have distribution restrictions.

System architecture assessment can include technology, integration, and manufacturing readiness assessments. These assessments are for selected subsystems and technologies for CTE and OTI – not the entire system. Reporting TRL/IRL/MRL for all system architecture elements and levels of decomposition would support risk early warning across the entire system to reveal evidence of previously unforeseen risks. Restricting TRL/IRL/MRL assessment to CTEs and OTIs elevates the risk of being blindsided by unforeseen challenges and events. However detailed investigation of TRL/IRL/MRL and potential risks requires commitment of time, cost, and decision authority.

The TRC are filled out at each of the major reviews, based evidence presented at the review and judgment of the reviewers. The TRC probe the status of the technical program in thirteen areas with specific, predetermined questions [28]. Status is rated red, amber, green, unknown, or not applicable. The TRC cover many different factors, from "was a responsible person appointed to conduct and approve the review" to "what fraction of the critical requirements have been shown to have been met."

The MCE contains estimates of the recurring and non-recurring costs, for each variant in the Family of Vehicles, against a detailed standardized Ground System Architecture. The Government provides the reporting framework. The contractor provides an update to the estimate at each major program review. All entries are initially rated "to be determined".

Maturity Level Advancement IMP Entries

In the proposal evaluation, the maturity levels, and the substantiating data, are inputs to assess the relative risks of alternative proposals. As the program advances in system design, manufacturing and RAM maturity, risk is retired. Maturity advancement ends with delivery the completed product. The maturity level definitions can be used integrate time, cost, and technical advancement reporting and analysis by inserting maturity advancement steps into the IMP.

The IMP framework is ideally suited to insert entries corresponding to demonstrated advancement in the maturity levels. Progress in maturity advancement, relative to time and budget consumed and remaining, is a natural indicator of risk. Since the IMP is the basis for the Integrated Master Schedule (IMS), and the Work Packages of the IMS are the EVM reporting elements, including maturity advancement events in the IMP forms a traceable link between objective technical progress, task organization, time, cost.

Maturity advancement can be inserted into the IMP by adding three events: design maturity advancement, manufacturing maturity advancement, and RAM maturity advancement. The following notional example illustrates how the maturity advancement could be resolved into IMP events, accomplishments and criteria, thus forcing traceable linkage between technical progress, time, and cost. The accomplishments are stages of maturity advancement towards completion for design, manufacturing, and RAM.

The design maturity accomplishments and criteria are design and development artifacts and test results, corresponding to major technical reviews [29] and test & evaluation events [30]. The manufacturing maturity accomplishments and criteria are drawn from the manufacturing readiness handbook [31]. The RAM maturity accomplishments and criteria are drawn from the design for reliability handbook [32].

Event: System Design Maturity Advancement

- Accomplishment 1: Requirements. Criteria: Completion of (a) requirements decomposition, (b) functional baseline, (c) derived requirements development, and (d) interface identification
- Accomplishment 2: Preliminary Design. Criteria: Completion of (a) allocated baseline, (b) component analysis and selection, (c) CAD models, (d) mass properties models, (e) load plan models, (f) hull/structure/frame finite element models, (g) powertrain and mobility models, (h) survivability models, and (i) interface models
- Accomplishment 3: Critical Design. Criteria: Completed sub-system architecture, design, integration and testing in the an integration lab for product WBS Configuration Items
- Accomplishment 4: System Integration & Test Readiness. Criteria: Completed sub-system integration, testing, analysis of deficiencies, and corrective measures in an operationally relevant environment for product WBS Configuration Items
- Accomplishment 5: Prototype Delivery and Operational Testing. Criteria: (a) delivery of prototypes, (b) completion of Operation Testing, (c) corrective action plans for residual deficiencies
- Accomplishment 6: Production Readiness and System Verification. Criteria: Completion of (a) final design, (b) Technical Data Package with CAD models, mass properties models, loading plan/model, and bill of materials, (c) live fire test

Event: Manufacturing Maturity Advancement

- Accomplishment 1: Manufacturing Concept. Criteria: Completed (a) initial manufacturing and process models, (b) materials acquisition approach
- Accomplishment 2: Manufacturing Requirements Analysis. Criteria: Completed identification of (a) manufacturing concepts, (b) producibility needs, (c) new manufacturing processes, (d) new manufacturing skills, (c) special facility requirements, (f) supply chain requirements
- Accomplishment 3: Preliminary Manufacturing Analysis. Criteria: Completed identification of (a) manufacturing modeling and simulation approaches, (b) lead times for materials, (c) exotic materials (hazardous, difficult to obtain and/or process), and (d) supply chain model and potential sources
- Accomplishment 4: Detailed Manufacturing Analysis. Criteria: Completed (a) modeling and simulation analysis at the component and subsystem levels to determine constraints, (b) assessment of issues, performance and reliability of similar full production processes, (c) identification of skill sets, special skills training and certification, (d) selection of supply chain sources
- Accomplishment 5: Manufacturing Feasibility Verification. Criteria: Completed verification of (a) the manufacturing processes in a production relevant environment, (b) availability of workforce skills, (c) adequate facilities and/or facility development plans, (d) long-lead items identified, (e) obsolescence/disposal issues identified, (f) supply chain and supplier agreements in place
- Accomplishment 6: Full Manufacturing Verification. Criteria: Completed (a) demonstration of manufacturing processing in a production relevant environment, (b) specification of manufacturing workforce resource requirements (staffing and floor managers), (c) specification of facility capabilities, (d) long-lead item procurement plan, (e) obsolescence plan

Event: RAM Maturity Advancement

- Accomplishment 1: Requirements Analysis. Criteria: Completed (a) reliability continuous improvement plan to meet reliability and maintainability requirements, (b) reliability (mean miles between system abort) and maintainability (maintenance ratio, mean time to repair, max-time to repair) allocated down to the Line Replaceable Unit (LRU) level
- Accomplishment 2: Preliminary Design Analysis. Criteria: Completed (a) design failure modes and effects analysis (DFMEA), (b) Fault Tree Analysis (FTA) for all essential functions listed in the Failure Definition and Scoring Criteria, (c) Critical Items List – items whose failure would cause a mission failure or Category III or higher Hazard Severity Rating as defined in MIL-STD-882D, (d) reliability and maintainability estimates made at the LRU level, (e) initial reliability growth plan and curve (per AMSAA Projection Maturity Model)
- Accomplishment 3: Integrated Subsystem Reliability. Criteria: Completed modeling and simulation and/or sub-system testing to estimate the reliability of integrated sub-system operation
- Accomplishment 4: Detailed Design Analysis. Criteria: Using integrated subsystem reliability date from Accomplishment 3, completed (a) updated DFMEA, (b) updated FTA, (c) updated reliability and maintainability estimates, (d) reliability growth plan and curve, and (e)

Failure Reporting and Corrective Action System (FRACAS) report

- Accomplishment 5: System Level Testing. Criteria: Completed (a) reliability, maintainability, and durability test plan, (b) testing the specified number of miles on the operational terrain profile, (c) test report with details of the driving profile, terrain profile, times and types of failures
- Accomplishment 6: Post Testing Detailed Design Analysis. Criteria: Using system level testing data from Accomplishment 5, completed (a) updated FRACAS report identifying failure modes, root causes, corrective actions, and validation, (b) updated reliability and maintainability predictions, and (c) updated reliability growth plan and curve

Risk Leading Indicators

Risk Leading Indicators are computed from baseline and update program management and systems engineering data. RLI assess the current status, as well as trends and instability. They assess (1) the change from the initial baseline, and (2) the change from the last update. This requires that the Systems Engineering system retain initial baselines and previous state. This is the minimum information to assess current state, short- and long-term trends. The reporting process samples the state of the system in periodic updates (e.g., monthly) and event-based updates (e.g., technical reviews). Early warning of risk exposure and emerging risk exposure requires early visibility into the state of the system and its progress. Waiting until late-stage technical reviews to identify issues exposes the program to risk.

The initial list of candidate Risk Leading Indicators follows. The RLI address requirements, maturity, design (baselines, architecture, and design margins), technical reviews, plan and schedule.

- Requirements Interdependency. Number of requirements, number of dependencies between requirements, ratio
- Requirements Stability. Proportion of requirements and dependencies that were (a) added, (b) deleted, and (c) changed
- Requirements Verification. Proportion of requirements with compliance verification activities in the IMS, number that have been verified, number that were fully/partially/not compliant, number of requirements linked to requirements that were fully/partially/not compliant
- Requirements Verification Schedule. Minimum and average schedule slack for remaining verification activities, minimum and average project time remaining after scheduled verification activities
- Maturity. Design, manufacturing and RAM maturity levels, difference between scheduled level and achieved level (at the criteria, accomplishment, and event levels)
- Manufacturing Cost Maturity. Proportion of applicable entries in the MCE that are blank or "to be provided", the proportion of entries with different values than the previous submittal; percentage change in the estimated unit production cost
- Baselines. Number of entries in the functional, allocated, and product baselines, number of links to requirements, number of links to architecture elements, ratios of changes to number in each baseline
- Architecture. Numbers of architecture elements and links between elements, proportion without completed boundary diagrams, proportion with changed boundary diagrams
- Design Margins. Design margin remaining as a proportion of the base, for each holistic system parameter (ground vehicle holistic system parameters are found in the Ground System Architecture Framework [26], e.g., mass properties, volume, dimensions, surface areas, main and auxiliary power, cooling capacity, fuel consumption, etc.)
- Technical Reviews. Total and by major sub-heading the number of applicable fields, for the applicable fields, the proportion rated red, amber, green, and unknown, number of changes from unknown to known, proportions of increases and decreases in level and from unknown to known and known to unknown
- Integrated Master Plan (IMP) Stability. Number of IMP entries, proportion added, deleted, or changed
- Integrated Master Schedule (IMS). Number of IMS activities (by Planning Package, Work Package, and detail task), number of dependencies (by logical dependency), ratio of links to nodes, proportion of activities completed, begun, started or finished out of sequence with their dependencies, mean and variance of the ratio of actual to planned duration for completed activities, float on the critical path relative to the critical path to each milestone, number of "high risk" activities (i.e., with ratio of float to scheduled duration is less than X percent)
- Schedule Stability. Numbers and proportions of activities and dependencies that were (a) added, (b) deleted, and (c) changed

Areas of High Relative Risk

Risk Leading Indicators can be computed for the entire program and system to assess overall risk exposure, and by segment to diagnose areas of high risk exposure. Different RLI are decomposed in different ways that give different insight into the areas at risk. Requirements RLI are linked to priority tier, and branch of the specification tree. Maturity RLI are linked to acquisition phase (development, manufacturing, and sustainment) and major subsystem. Design RLI are organized by product WBS. Schedule RLI are linked to WBS element, and by IMP entry. Technical review checklists have their own organization.

Risk Leading Indicators that are correlated can be combined using Principal Components Analysis to reduce the number of indicators and variability. This requires an accumulation of data points.

Predicting future time, cost, performance differences between actual and planned results (assuming that past relationships between RLI and subsequent outcomes will persist) requires an accumulation of data points by IMS activity.

Some decisions need to be made before there has been enough history to compute trend and stability metrics, and before there has been time to accumulate data to calibrate statistical RERs. “Shortcut” methods are needed to assess relative risk (a) in proposal evaluation without “trend and stability” data, (b) during contract execution but without sufficient history to quantify time, cost, and performance bias and dispersion, and (c) when there is insufficient data to project bias and dispersion of outcome.

The underlying concept is to find outliers, e.g., WBS elements with high levels of RLI relative to other WBS elements. Even without data to quantify the relationship between RLI and time, cost, performance outcome, it is still possible to identify WBS elements with elevated RLI – e.g. if one proposal or WBS element has RLI that are three standard deviations above the norm, it is higher risk than one whose RLI are one standard deviation above the norm.

Risk Estimating Relationships (RERs)

RERs are equations or models. The RLI are the inputs. Time, cost and performance bias and uncertainty are outputs. RERs are calibrated to program data on time, cost, and performance by IMP entry, WBS element, and the overall program. Calibration uses historical data on the program to compute model coefficients. Differences between IMP entry, WBS element, and entire program provide insight into the sources of risk exposure.

Risk Estimating Relationships are essentially regression models that explain future IMS activity cost and schedule bias and dispersion in terms of current risk leading indicators. RERs are similar to Cost Estimating Relationships (CERs) used in parametric, analogy cost models. RERs are statistically significant relationships that explain program performance as functions of earlier RLI.

The RER are calibrated to previous period data for the current program. Program-to-program differences make it unlikely that quantitative evidence from one program will be relevant to another program with different design challenges, performance objectives, contractor management, engineering team, etc. The RER contain autoregressive components (past trends predict future trends) and logical components (incompleteness, instability, inconsistency, lack of safety margins, interdependency in the current state predict deficiencies in future outcomes).

Program performance data are accumulated over time. At proposal evaluation, only uncertainties inherent in the RFP package and the proposals can be assessed. After contract award, data can be accumulated regarding input conditions and output results, by IMP entry, WBS, and overall program.

Each completed activity in the IMS constitutes a data point with a cost and schedule variance. Completed requirements verification tasks also provide compliance variance. Each IMP event is a data point.

Regression models include all computational methods that use historical data to fit a type of model between input and output variables. Candidate approaches include continuous and discrete models. Continuous models include linear regression, multi-linear regression, logistic regression, polynomial regression, and artificial neural networks. Discrete models include naïve Bayes models (that assume independence among the causes and effects of inputs on outputs), and two-stage models that aggregate over families of simple models (e.g. Aggregate One Dependence Estimators that pool one dependence naïve Bayes models), and calibrating gains to explain the historical data.

The choice of underlying RER models is pragmatic – the formalism that works best is best to explain progress variances. Previous programs can suggest high value models and parameters, but differences among program may reduce relevance.

The question is which methods work – in practice, for this application – not which are best in theory. The characteristics of the application are complex. Different programs will have different challenges. This makes extrapolatinf from on program to another problematic. There are a large number of potential risk exposure indicators. In any given program, at any stage of the program (punctuated by the major technical reviews), and divided by the WBS elements, different indicators can have different significance. Outcome states can be positively or negatively correlated. EMD time and cost tend to be positively correlated, and negatively correlated with system performance, production cost and RAM.

The large number of IMS activities provide data to correlate WBS/IMP/IMS time/cost expenditure and technical accomplishment to the RLI (e.g., IMS limited to 6,000 detail tasks). The semi-hierarchical lattice structure of requirements, program activities, system segments, and system architecture pose challenges to statistical analysis, as does potential differences in the program between technical review milestones and level 3 WBS elements.

SUMMARY

This paper develops the concept of risk exposure. Risk exposure amplifies the likelihood and/or consequences of unanticipated complications, technical difficulties and delays. Exposure to risk increases the risk of adverse acquisition outcomes.

Exposure to risk is created by overly optimistic goals that lead to inadequate margins for error, aggressive concurrent schedules counting on "things to come together at the end," deferred or limited testing, coordination shortcuts, adopting novel integration processes and promising but immature technologies. Unstable, inconsistent, incompletely resolved, and highly interdependent program plans and system engineering documents both indicate and create risk exposure. Lagging and uneven technical progress relative to the plan is a further indicator of risk exposure.

Risk Leading Indicators capture objective evidence from program management and system development reports and data to assess exposure to risk, and diagnose areas and type of risk exposure. Risk Estimating Relationships estimate the bias and uncertainty between actual and planned time, cost and performance of program activities. Risk Estimating Relationships are calibrated to past performance on the program of interest, as data are collected.

The approach is practical and relevant. It is tightly linked to standard deliverable data. It builds on the program risk evaluation frameworks and criteria set out, in formal documentation, for proposal and contact execution evaluation as determined by the PMO. It builds on prior empirical analyses of system development leading indicators, program risk leading indicators, and root causes and causal mechanisms of adverse acquisition outcomes.

STATUS AND PLANS

Software tools to compute a subset of the initial RLI are in progress. Selection of outlier detection and cluster analysis methods to detect areas of high relative risk exposure is pending pilot study data collection. Selection of the underlying modeling framework for the RER is also pending pilot study data collection. Extensive research and development has been published in both of these areas, but the choice of methods depends on the characteristics of the data set [33].

Verification plans involve a pilot study that will exercise the RLI in coordination with and in support of ground vehicle acquisition program. The pilot study will provide risk exposure early warning feedback to the PMO. This will demonstrate the practicality, relevance, and value of the tools and methods. The pilot study will also provide data for selection of technical methods as noted.

The plan to transition the risk exposure early warning tools into formalized Systems Engineering practice involves continuing to work with the TARDEC's Integrated Systems Engineering Framework (IESF) team. ISEF [34] is an Army Research, Development, and Engineering Command (RDECOM) solution to integrate previously stove-piped systems engineering information and processes, disparate tools, one-off integrations, and a lack of accepted, common standards that too often occur.

ISEF is providing standardized reporting requirements and evaluation MPT. It is a working, relevant and practical toolset/system in use on multiple programs. The system engineering and analysis tools are prioritized to support the PMO needs. ISEF defines data/evidence content, connects federated databases, and preserves knowledge patterns. It is a collection of systems engineering tools united around a common information architecture to address these issues in today's Army and other DoD agencies.

PARTING THOUGHTS

"*The most serious mistakes are not made as a result of wrong answers. The truly dangerous thing is asking the wrong question*." Peter Drucker

"*There are known knowns. These are things we know that we know. There are known unknowns. That is to say, there are things that we know we don't know. But there are also unknown unknowns. There are things we don't know we don't know*." Donald Rumsfeld

"*It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so*." Mark Twain

"*In God we trust; all others bring data." "You can't manage what you don't measure*." W. Edwards Demming

"Short cuts make long delays." Peregrin Took (character in Tolkien's *Lord of the Rings* trilogy)

ACKNOWLEDGEMENTS

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense through the Systems Engineering Research Center (SERC) under Contract H98230-08-D-0171. The SERC is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology.

REFERENCES

[1] DoD. Risk Management Guide for DoD Acquisition, 6th Edition. Aug 2006. http://www.acq.osd.mil/se/docs/2006-RM-Guide-4Aug06-final-version.pdf

[2] Tecolote Research Inc. U.S. Air Force Cost Risk and Uncertainty Analysis Handbook. Air Force Cost Analysis Agency. Apr 2007. https://acc.dau.mil/adl/en-US/316093/file/46243/AF_Cost_Risk_and_Uncertainty_Handbook_Jul07.pdf

[3] GAO. Defense Acquisitions – Where Should Reform Aim Next? GAO-14-145T. Oct 2013. http://www.gao.gov/assets/660/658615.pdf

[4] Karen Richey. Update to GAO's Cost Estimating Assessment Guide and Scheduling Guide (draft). GAO. Mar 2013. http://www.aacei-ncs.org/gaohq-6188401-v1-presentation_to_aace_mar-2013.pdf

[5] Gary Bliss. Observations from AT&L/PARCA's Root Cause Analysis. PARCA. Mar 2012. https://acc.dau.mil/adl/en-US/542861/file/67065/10%20-%202012-11-15%20PARCA,%20Nunn-McCurdy%20Trends%20and%20Lessons%20Learned,%20Bliss.pdf

[6] NDIA Systems Engineering Division. Report on Systemic Root Cause Analysis of Program Failures. NDAI. Dec 2008. http://www.ndia.org/Divisions/Divisions/SystemsEngineering/Documents/Studies/NDIASRCAReportFINA18Dec2008.pdf

[7] Joachim Hofbauer, Gregory Sanders, Jesse Ellman, and David Morrow. Cost and Time Overruns for Major Defense Acquisition Programs. Center for Strategic and International Studies. Apr 2011. http://csis.org/files/publication/110517_DIIG_MDAP_overruns.pdf

[8] Joseph Bolten, Robert Leonard, Mark Arena, Obaid Younossi, and Jerry Sollinger. Sources of Weapon System Cost Growth. RAND. 2008. http://www.rand.org/content/dam/rand/pubs/monographs/2008/RAND_MG670.pdf

[9] Gene Porter. The Major Causes of Cost Growth in Defense Acquisition. Institute for Defense Analysis. IDA Paper P-4531. Dec 2009. www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA519884

[10] Assistant Secretary of Defense for Research and Engineering. Technology Readiness Assessment (TRA) Guidance. DoD. May 2011. https://acc.dau.mil/adl/en-US/154268/file/59527/TRA%20Guide%20OSD%20May%202011.pdf

[11] Lauren A. Fleishman-Mayer, Mark V. Arena, Michael E. McMahon. A Risk Assessment Methodology and Excel Tool for Acquisition Programs. RAND. 2013. http://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR262/RAND_RR262.pdf

[12] NDIA Systems Engineering Division. Top Systems Engineering Issues in U.S. Defense Industry. National Defense Industrial Association. Sep 2010. http://www.ndia.org/Divisions/Divisions/SystemsEngineering/Documents/Studies/Top%20SE%20Issues%202010%20Report%20v11%20FINAL.pdf

[13] Gary Roedler, Donna H. Rhodes, Howard Schimmoller, and Cheryl Jones. Systems Engineering Leading Indicators Guide – Version 2.0. INCOSE-TP-2005-001-03. Jan 2010. http://www.incose.org/ProductsPubs/pdf/SELI-Guide-Rev2-01292010-Industry.pdf

[14] NDIA Systems Engineering Division. System Development Performance Measurement – Phase I Report. NDIA. Oct 2011. http://www.ndia.org/Divisions/Divisions/SystemsEngineering/Documents/Studies/NDIA%20System%20Development%20Performance%20Measurement%20Report.pdf

[15] NDIA System Development Performance Measurement Working Group. System Development Performance Measurement Project 2012 Final Report. NDIA. Dec 2012. http://www.ndia.org/Divisions/Divisions/SystemsEngineering/Documents/Past%20Meetings/December%202012%20Meeting/SDPM%202012%20Final%20Report%20v4.pdf

[16] Paul Kohl and Ronald S. Carson. New Opportunities for System Architecture Measurement. Lockheed Martin Corp. 15th Annual NDIA Systems Engineering Conference. Oct 2012. http://www.dtic.mil/ndia/2012system/ttrack914886.pdf

[17] Mike Ferraro HQ DCMA. Implementing Technical Performance Measurement. PEO/SYSCOM Conference. Nov 2002. http://www.acq.osd.mil/dpap/about/PEOSYSCOM2002/presentations/Track4B-MikeFerraro.ppt

[18] Mike Ferraro. Technical Performance Measurement—A Program Manager's Barometer. DCMA. Program Manager. Nov-Dec 2002. http://www.dau.mil/pubscats/PubsCats/PM/articles02/fer-jf3.pdf

[19] United States Government Accountability Office. GAO Schedule Assessment Guide – Best Practices for Project Schedules. GAO-12-12OG. May 2012. http://www.gao.gov/assets/600/591240.pdf

[20] NASA. 2008 NASA Cost Estimating Handbook. NASA. 2008. http://www.nasa.gov/pdf/263676main_2008-NASA-Cost-Handbook-FINAL_v6.pdf

[21] International Society of Parametric Analysis. Parametric Estimating Handbook, Fourth Edition. Apr 2008. http://www.galorath.com/images/uploads/ISPA_PEH_4th_ed_Final.pdf

[22] Wilson Rosa, Barry Boehm, Brad Clark, Thomas Tan, and Ray Madachy. Domain-Driven Software Cost

Estimation. 27th International Forum on COCOMO and Systems/Software Cost Modeling. Oct 2012. http://csse.usc.edu/csse/event/2012/COCOMO/

[23] DoD. Work Breakdown Structure for Defense Materiel Items. MIL-STD-881C. Oct 2011. http://www.navair.navy.mil/nawctsd/Resources/Library/Acqguide/MIL-STD%20881C%203%20Oct%2011.pdf

[24] DoD. Integrated Master Plan and Integrated Master Schedule Preparation and Use Guide. Oct 2005. http://www.acq.osd.mil/se/docs/IMP_IMS_Guide_v9.pdf

[25] Defense Contract Management Agency. Earned Value Management Implementation Guide. Oct 2006. https://acc.dau.mil/adl/en-US/386074/file/52051/DoD%20EVM%20Implementation%20Guide%20Oct%202006.doc

[26] Dennis Fett, William Pritchett, and Jim Richardson. The JCGV Ground System Architecture Framework. U.S. Army TARDEC. NDIA 16th Annual Systems Engineering Conference. Oct 2013. http://www.dtic.mil/ndia/2013system/W16096_Pritchett.pdf

[27] Defense Acquisition University. Systems Engineering Configuration Baselines. Jul 2014. http://www.acqnotes.com/Career%20Fields/Configuration%20Baselines.html

[28] DoD. Preliminary Design Review (PDR) Checklist. Jun 2007. https://acc.dau.mil/adl/en-US/640005/file/69326/DoD%20PDR%20CheckList%2018%20April%202013.xls

[29] PEO IWS. Technical Review Manual. Dec. 2009. http://www.acqnotes.com/Attachments/NAVSEA%20Technical%20Review%20Manual%2018%20Dec%2009.pdf

[30] DoD. Test & Evaluation Management Guide. Dec 2012. http://www.dau.mil/publications/publicationsDocs/Test%20and%20Evaluation%20Management%20Guide,%20December%202012,%206th%20Edition%20-v1.pdf

[31] Defense Research and Engineering. Manufacturing Readiness Level Deskbook. DoD. Jan 2010. http://www.dodmrl.com/MRL_Deskbook_v1.pdf

[32] US Army AMSAA. Design for Reliability Handbook. TR-2011-24. Aug 2011. http://web.amsaa.army.mil/Documents/Design%20for%20Reliability%20Handbook%20(TR-2011-24).pdf

[33] Pang-Ning, Michael Steincach, and Vipin Kumar. Introduction to Data Mining. Addison-Wesley. 2007. http://www-users.cs.umn.edu/~kumar/dmbook/index.php

[34] Pradeep Mendoza. The TARDEC Advanced Systems Engineering Capability. 2012 NDIA Ground Vehicle Systems Engineering and Technology Symposium. Aug 2012. http://www.dtic.mil/dtic/tr/fulltext/u2/a566803.pdf