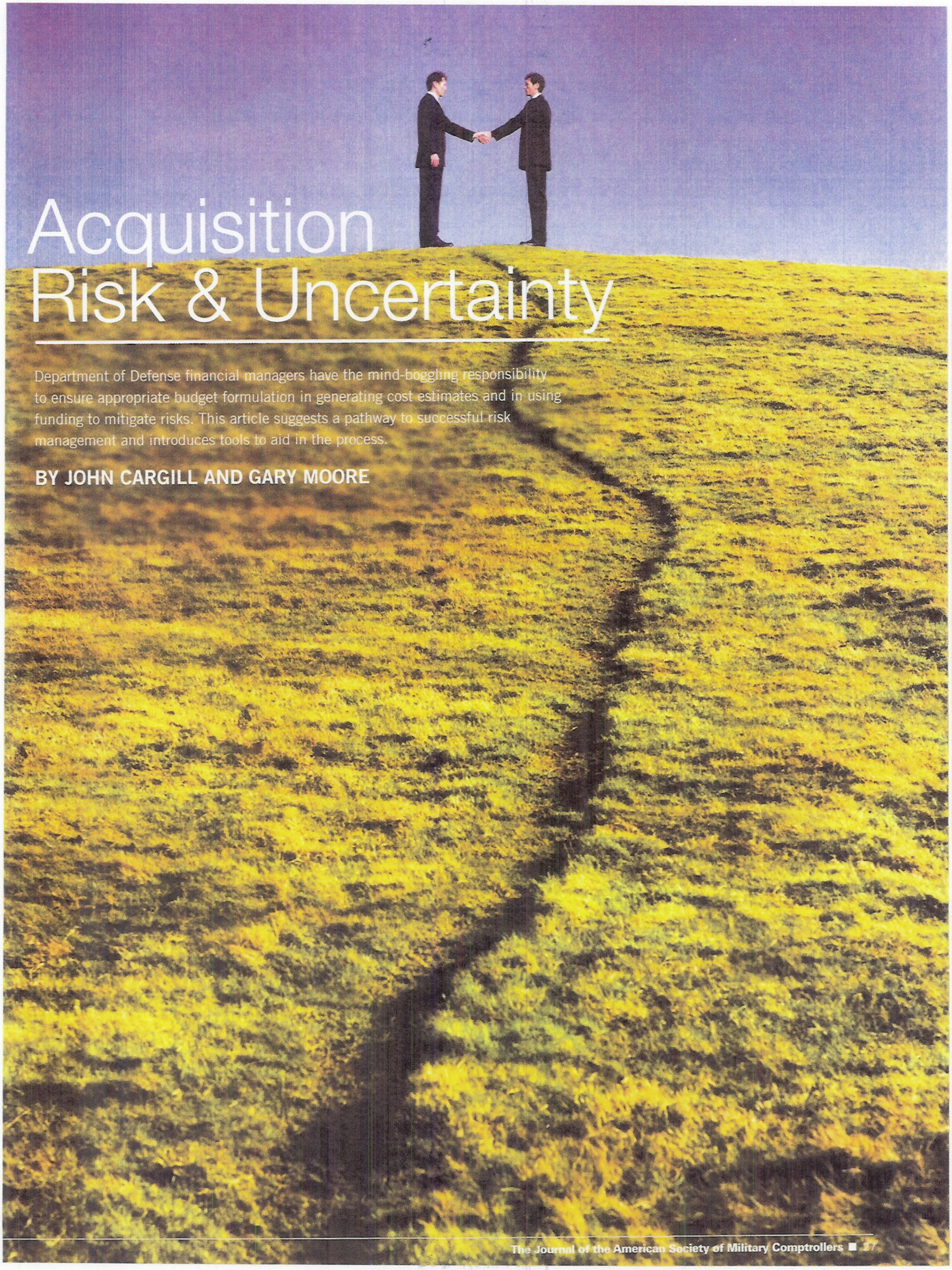

Acquisition Risk & Uncertainty

Department of Defense financial managers have the mind-boggling responsibility to ensure appropriate budget formulation in generating cost estimates and in using funding to mitigate risks. This article suggests a pathway to successful risk management and introduces tools to aid in the process.

BY JOHN CARGILL AND GARY MOORE

> "Through cost and schedule estimates, [military financial managers] influence the shape and size of tomorrow's Department of Defense and Coast Guard programs. As reports of alarming cost growth continue, the credibility of acquisition program cost estimates has been called into question."

American Society of Military Comptrollers members currently serving in decision-support positions, whether knowingly or unwittingly, play an important role in risk management. Through our cost and schedule estimates, we influence the shape and size of tomorrow's Department of Defense (DoD) and Coast Guard programs. As reports of alarming cost growth continue, the credibility of acquisition program cost estimates has been called into question. This article explains what we know about the factors influencing cost growth on past programs and what we, as analysts, can do to enhance the credibility of future program cost estimates. We also share insights for financial managers whose function is to execute programs within funding and what may be a slightly new perspective on risk and uncertainty analysis. We address how estimates are generated to include risk, highlight some of the pitfalls in developing risk ranges, and suggest an improved Air Force best-practice approach that can be used to identify risks.

Evidence of Cost Growth

According to Dr. Jamie Morin, Assistant Secretary of the Air Force for Financial Management and Comptroller, "Over a 40-year period, actual costs for Major Defense Acquisition Programs between Milestone B (the transition between technology development and engineering development) and program completion exceeded the original estimates by an average of 45% for the DoD overall and 54% for the Air Force. A recent independent study shows that the situation has gotten worse rather than better. Across the 15 active Air Force programs, program costs are estimated to be 114% over the initial projected cost, and for the eight programs that have been initiated since 2000, the estimated average increase jumps to 195%."[1] At the time of his statement, in 2010, cost growth trends were not improving over that history. As financial managers with a role in providing decision support, we must play a key part in answering the question, Why are Air Force weapons systems (or those of other Services) continuing to exceed threshold cost and schedule estimates?

The Government Accountability Office (GAO) has opined for years that it is common for DoD programs not to have gone "through system development meeting the best practices standards for mature technologies, stable design, or mature production processes by critical junctures of the program. . ."[2] More specifically, GAO suggests that "programs that began with immature technologies have experienced average research and development cost growth of 34.0%."[3] "Programs consistently move forward with unrealistic cost and schedule estimates, use immature technologies in launching product development, and fail to solidify design and manufacturing processes at appropriate points in development."[4] The GAO has concluded that proceeding with immature technology is a major cost-growth driver. One assessment of 62 weapons systems with a total investment of more than $950 billion found that (1) "[f]ully mature technologies were present in 16% of the systems at development start"[5] and (2) "[p]rograms that began development with immature technologies experienced a 32.3% cost increase, whereas those that began with mature technologies increased 2.6%."[6]

We have to ask why so many programs proceeded with immature technologies while overlooking the risks and uncertainties associated with that reality. The short answer is that planners succumb to the so-called "conspiracy of hope" caused by natural program optimism, subjectivity, and bias. The detailed answer is far more complex and points to the fact that there often are an infinite number of elements that can impact a program's risk and uncertainty, most of which are not fully analyzed at program initiation.

This article introduces some concepts that facilitate improved program evaluations and instill an objective, honest broker counterbalance to the inherent planning optimism often exhibited by program advocates. We will explore useful tools, such as expected value estimating and the Air Force RI3 Guide, as means to provide insights and realistic assessments of program risks and uncertainty. We contend that the use of these tools to influence funding decisions will help mitigate and contain much of the cost growth associated with risk and uncertainty.

Methods Used to Measure Risk

"Technology Readiness Levels (TRLs) are measures of the maturity of the technologies for which a program is pursuing."[7] (See the boxed information for TRL definitions.) Low TRLs are inherently more risky than high TRLs. From analysis of past acquisition programs, it is clear that inflation of TRLs has been a problem. Consequently, it is incumbent upon financial managers at all levels to assist in discerning unbridled optimism and aspects of the acquisition culture that contribute to unrealistic ratings of TRLs. New guidance and independent assessments of TRLs already have enhanced our ability to assess the cost risk and uncertainty for programs at different TRLs. Much work, however, is left to be done, and cost estimators must be

Technology Readiness Level Definitions

1 = Basic principle observed
2 = Technology concept and/or application formulated
3 = Analytical and experimental function and/or characteristic proof of concept
4 = Component and/or breadboard validation in a laboratory environment
5 = Component and/or breadboard validated in a relevant environment
6 = System/subsystem model or prototype in a relevant environment
7 = System prototype demonstration in an operational environment .
8 = Actual system completed and qualified through test and demonstration
9 = Actual system proven through successful mission operations

closely in tune with the engineers in understanding the true maturity of technology proposed for the programs being estimated.

Financial managers need to identify when conspiracy of hope and group-think cloud program judgment. There are many simple indicators that trigger the need to do hard investigations. Those triggers most obvious include nonfact-based justifications such as, "That's an unacceptable number" or "If we say it will cost that much, they will cancel the program." Other clues to impending cost overruns are much more subtle. Optimism and a can-do attitude, at both the government program office and the contractor program office, are essential to running a successful cutting-edge, high-technology program. The optimism, however, has to be kept within bounds. Part of our job as financial managers and cost estimators is to inject that objective, honest-broker realism into the discussions to ensure that risks and uncertainties aren't overlooked as we plan and implement these programs.

One of the best ways to visualize risk potential in a wide range of endeavors is the GAO Cone of Uncertainty.[8] Figure 1 shows the change in the range of cost estimates as a program matures. It is important to note that at the beginning of a major project, there generally is greater potential for cost overrun than for underrun; thus, the Cone of Uncertainty is asymmetrical. As the program progresses, some risks are experienced and others fail to materialize as issues are well mitigated. Estimates will tend to grow over time as they are refined to take into account the increased knowledge and understanding of associated program risks. As the program becomes better defined, the uncertainty decreases. As programs mature, some risk can be retired and risk mitigation plans can be improved, thereby reducing the potential for unexpected cost and schedule growth.

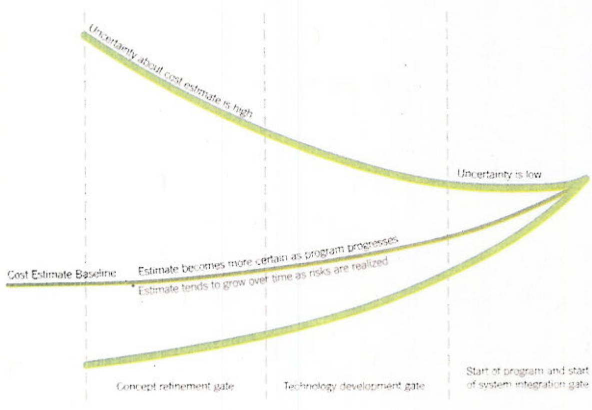

Figure 1. GAO Cone of Uncertainty

Notionally, programs with low TRLs have more risk and uncertainty than programs with high TRLs. Similarly, programs lacking heritage and requiring 100% new design are inherently more risky and have more unknowns than programs highly leveraged from previous programs. Figure 2 shows that by putting the GAO Cone of Uncertainty Concept together with TRLs, we see stabilization over time of program phases and risk across a range of program cost outcomes.

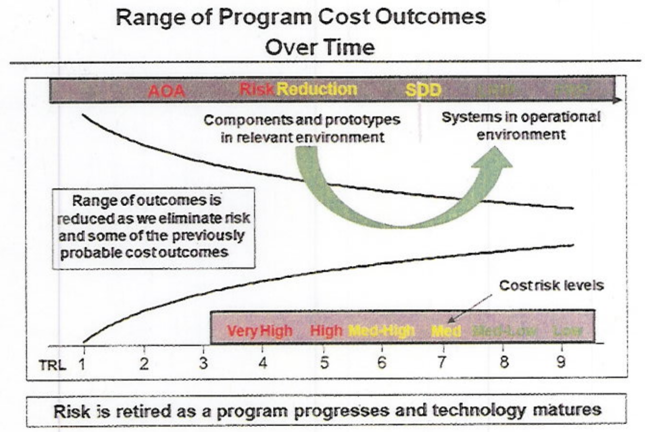

Figure 2. Range of Cost Outcomes Related to TRL, Risk, and Acquisition Phases Over Time

Cost-Estimating and Funding Approaches to Account for Risk

Ensuring that programs enter the acquisition with mature technology is a substantial step in the right direction, but that action alone will not prevent program cost growth. For programs to be successful, we rely on cost estimators to produce estimates having cost realism. Cost realism is a term of art that focuses on developing an expected value for the cost of a program. Four important cost-estimating components for developing this expected value are:

- Technical Baseline Estimate (TBE) – the sum of the "most likely" components, subsystems estimates.
- Risk – the chance of "loss" or "injury." Acquisition programs require a quantified assessment of program risks, to include technical, schedule, organizational, and cost elements.
- Uncertainty – the indefiniteness about the outcome of a situation. No program office can identify the complete range of potential outcomes before they happen. The literature suggests that subject matter experts are doing well if they can identify 70% of the probable range of potential outcomes.
- Skew – the amount of the probability density function (distribution) either below or above the most likely estimate.

In most cases, the outcome of these analyses is a cost-confidence probability distribution commonly known as the S-curve. This distribution actually is a range of costs where higher confidence levels on the S-curve typically yield higher costs. A confidence level (CL) is the probability that one will achieve values up to and including the given value. So an 80% CL of $1.1 billion means an 80% probability actual cost of the system will be less than or equal to $1.1 billion. The mean of the cost-estimate distribution (in large acquisitions typically found to be 50 to 65% CL) includes expected levels of risk, uncertainty, and external influences. In other words, the mean or expected value of an acquisition program is defined as the very best cost estimate (not the TBE sum of the most likely costs of the system components) adjusted for levels of expected risk and uncertainty. It is the output value that takes into consideration all underlying risk and uncertainty distributions incorporated in the estimate.

Now let's switch from discussing the formulation of the expected cost estimate to the connections between the cost estimate and the funding level. To be clear, the funding level for a program should be based on the cost estimate. Often, however, there are other drivers separating funding from the cost estimate. These include:

- portfolio dependency;
- added funding to ensure delivery for critical "must have soonest at any cost" programs;
- items added that engineers, program managers, and cost estimators are thought to have missed; and
- changes due to emerging threat assessments.

The Air Force, Office of the Secretary of Defense, and the Congress at various times have experimented with mandated funding at certain confidence levels. To simplify for this article, current practice is for programs typically to be funded at the expected program cost (the mean of the cost risk distribution). Note that the expected value is linked integrally to program content; thus, no funds are available to be removed without reducing program scope. The goal is to ensure proper funding for the program to execute under normal (realistic or on average) conditions. Or as GAO puts it:

- "The main purpose of risk and uncertainty analysis is to ensure that a program's cost, schedule and performance goals can be met."[9]
- Cost estimators should "[r]ecommend sufficient contingency reserves to achieve levels of confidence acceptable to the organization."[10]

So we've summarized two aspects of analysis that cost estimators and financial managers must perform to ensure that risk and uncertainties are not overlooked: (1) Start with a realistic assessment of technology maturity and (2) develop the expected cost for the program incorporating risk and uncertainty assessments. Now let's explore further exactly how the cost estimator and financial manager can avoid some pitfalls associated with performing risks and uncertainty assessments and why this knowledge and approach make them better financial advisors to the program manager (PM).

> "Two aspects of analysis that cost estimators and financial managers must perform to **ensure that risk and uncertainties are not overlooked:** Start with a realistic assessment of technology maturity and **develop the expected cost for the program** incorporating risk and uncertainty assessments."

Subject Matter Expert Bias

"There is a natural tendency to be aggressive with assumptions early in a program to make the program appear attractive."[11] "Therefore, bias gets introduced not through flaws in any of the methods but through bias in the assumptions and factors used in the analysis."[12] Thus, we need to delve more deeply into opinion, bias, and subjectivity, and how these factors influence our risk and uncertainty assessments.

Senior financial managers and cost estimators often have the benefit of experience from working multiple programs; they naturally bring this experience to the table. This experience makes them wary of subjective assumptions and claims, since they have seen the bias drivers and the resulting history of program cost growth play out in real life. They know how powerful good risk and uncertainty assessments can be in providing the PM with a clear picture of realistic outcomes, counterbalancing optimistic tendencies. A clear understanding of bias and subjectivity may predispose the two communities to greater openness about the potential pitfalls lurking "outside the door" of every acquisition program. We caution readers not to take this as a proposal for cynicism or skepticism; rather, view it as a more positive attitude of "show me" and "make your case."

One of the better statements about confidence levels and certainty comes from Trevor Van Atta: "Can we calculate the precise probability of a project overrunning one million dollars through the use of statistics? ...Unfortunately, no matter how much data gathering and analysis we do, we cannot place limits on the real world. We could calculate a 70% confidence interval for an estimate, but some expected events could occur that alter the odds and throw the cost of the project spinning out of control.... Real world scenarios are not decompositional such that we can account for all possible outcomes, so an objective probability calculation of risk isn't possible."[13]

The cost estimator's risk distribution and uncertainty assessments bridge the gap between intuition and mathematics. This bridging process requires an understanding of the powers and flaws of our intuition (subject matter expert opinion) and boundary interpretation before we can formulate any meaningful statistical inferences about cost. The U. S. Air Force Cost Risk and Uncertainty Analysis Handbook has a good example of boundary interpretation, depicted in Figure 3.[14]

In this example, expert opinion was sought to identify a range for cost (could be hours, scrap, rework percentage, etc.) of 90 to 170 with a value of 145 at the 80% confidence level. The cost estimator estimated that the expert was able to identify 70% of the range of outcomes (typical at Milestone B). This resulted in a right-skewed triangular distribution. To capture the fact that the range of 90 to 170 identifies only 70% of the range of outcomes, the estimator must adjust for the "uncertainty not captured." By extending the tails of the distribution by relative amounts, the resulting distribution is more in the range of 60 to 250 with a value of 170 at the 80% confidence level (rather than a value of 145).

From the preceding example, we see that as uncertainty and risk increase, the range of outcomes becomes significantly wider. This lends credence to the premise in this article that we have been underestimating the true risk and uncertainty in Air Force cost estimates. As technical and mathematically

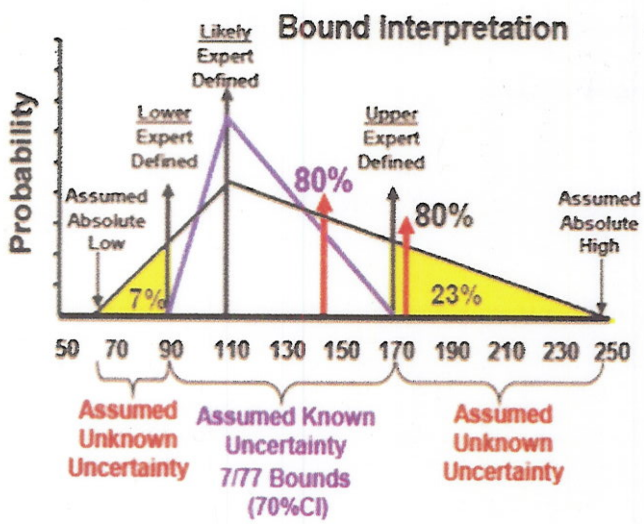

Figure 3. Triangular Distribution Accounting for Uncertainty

correct as the cost estimators attempt to be, we have to recognize that intuitive judgment significantly influences projected cost ranges.

Every cost forecast is a hypothesis based on a series of reasoned judgments about the future. While this does not inspire great confidence in the accuracy of a point estimate, it is better than the alternatives. Daniel Kahneman, the 2002 Nobel Prize Winner for Economics, suggests that "[o]bjective measurements of probability are often unavailable, and most significant choices under risk require an intuitive evaluation of probability."[15] Put another way, whenever there are risk and uncertainty, there is subjectivity. The pertinent question becomes, What do we need to address this subjectivity? The approach described above offers one improvement over how we currently assess this subjectivity. There is, however, more to this subjectivity than just the boundary interpretation just shown.

Under conditions of commitment and optimism for the future, no one likes to recognize risk. We have explored some general characteristics of risk relative to uncertainty, heritage, program phase, and bias to include group-think and over-optimistic acquisition culture. Since we are so highly influenced by expert subjective judgment, it is important to look at three of the psychological processes distorting that judgment: conjunction fallacy, representativeness, and anchoring and adjustment. To illustrate these concepts, we will draw upon Kahneman's famous "Linda Problem."

Conjunction Fallacy

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice and also participated in anti-nuclear demonstrations.

Based on the preceding description, which of the following statements about Linda is more probable?

a. Linda is a bank teller.
b. Linda is a bank teller and is active in the feminist movement.

Results: In Kahneman's study, 85% of respondents chose answer (b) in clear violation of the conjunction rule. In short, if you did not answer (a), then you brought your own biases to what was a very simple statistical question. The very, very simple and patently obvious conjunction rule of probability states:

For all X and Y, P(X & Y) <= P(Y)

In other words, for all X and Y, the probability of X and Y occurring together is less than or equal to the probability of Y occurring. Figure 4 depicts this outcome in the Venn diagram of the Linda Problem. Many psychological researchers have tried to explain why so many people choose the intersection of bank tellers and feminists (shown in green) over either bank tellers or feminists. It well may be that "[w]hen faced with a difficult question, we often answer an easier one instead, usually without noticing the substitution."[16] This means that we can expect experts to take a shortcut and replace the real question at hand with one that is simpler; these heuristics (rules of thumb) lead to systemic and predictable biases.

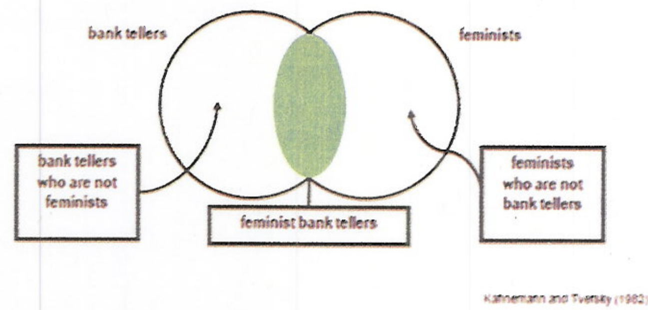

Figure 4. Venn Diagram of the Linda Problem

Gary Charness, et al., put a new twist on this old Linda Problem that business majors are exposed to in undergraduate and graduate studies. Those authors introduced collaboration and incentives to see whether those additions to the mix could improve the quality of the answers.[17] The results were astounding and are shown in Figure 5. Using teams of three, it was possible to cut conjunction rule violations by 56% through collaboration and 82% by using collaboration and a minor, almost immaterial, incentive. We submit that this collaboration between independent financial managers, cost estimators, program engineers, and program managers can go a long way toward avoiding the conjunction fallacy pitfall. Collaboration works best when the collaborators are relatively independent of each other so that group-think and conspiracy of optimism are not dominant.

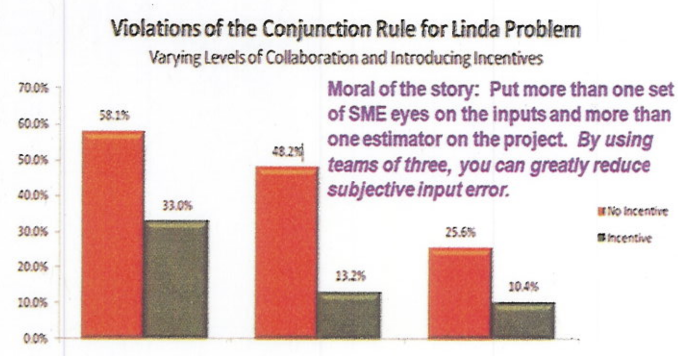

Figure 5.

Representativeness

Kahneman proposed the concept of "representativeness," where we tend to draw a large inference on what we perceive to be a "representative" story or sample. This perception causes us to favor narrow distributions over wider distributions. When looking for analogous programs from which to extrapolate or interpolate, it is important to keep in mind the tendency toward representativeness. Respondents were told the U.S. Federal Government spent $22.5 billion on education in 1987. They were asked to answer the following question accordingly:

What amount of money was spent on education by the U.S. Federal Government in 1987?

a. $18 to $20 Billion
b. $20 to $40 Billion

Results: Eighty percent of the respondents chose answer (a.) because they favored narrow distributions over wider distributions. This experiment has been repeated many times by many researchers with different questions but has yielded similar results. Ilan Yaniv has one of the better examples:[18]

Respondents were given, for instance, estimates of the number of United Nations member countries (in 1987). In one case, the following two estimates were given:

a. 140 to 150
b. 50 to 300

Respondents were also told that the correct answer was 159 and then asked to indicate which answer was better. Most (90%) of the respondents preferred estimate (a.) over (b.), even though only the latter included the correct answer. Thus, respondents were willing to accept some error in order to obtain more informative judgments. The suggestion has been made that the desire for a minimum error may have played a part in the choice of answers.

Anchoring and Adjustment

Kahneman has suggested we tend to estimate an uncertain value by clinging to a prominent reference point that we know to be wrong (anchoring), then adjusting to a more likely value. That is a wild statement. Margaret Neale and Gregory B. Northcraft tested it on their own with four groups of participants independently viewing the same Tucson, Arizona, home for 20 minutes.[19] Each group was provided a different listing price and asked to estimate a "reasonable" sales price. The results are shown in Figure 6. Each group anchored on the listing price and adjusted downward. A regression line drawn through these four data points, described as "Estimate = 40.91 + 0.3587*Listing Price," with an R-Square of 81.95%, shows the consistency of the downward adjustments from the listing price. The respondents knew they did not like the listing price and, in all cases, made a downward adjustment. A similar bias can creep into cost estimates when evaluating contractor bids.

RI3 Approach

What questions should be asked to qualify and quantify risk? The Air Force has a method for that, called the Risk Identification, Integration and "ilities" (RI3) Guide.[20,21] It is a lessons-learned guide providing a concise set of questions to highlight key risk "areas that seem to have traditionally been underrated in terms of risk, [and] interestingly, literature from the field of cognitive psychology generally suggests that people often have difficulty in characterizing the relative risks of various activities appropriately (possibly due to 'group-think'), thereby resulting in underestimation of their effects."[22] The RI3 questions are based on those needed to be asked in past programs that ran into trouble, but were not asked. It uses a systems engineering strategy to enable sound decisions and avoid cost and schedule growth. The questions are separated into the following categories:

- Design Maturity and Stability
- Scalability and Complexity
- Integrability
- Software
- Reliability
- Maintainability
- Human Factors
- People, Organization, and Skills

The RI3 questions are designed for use by the chief engineer of program engineers, by the PM of the chief engineer, by the cost estimator of the program office, and by auditors. The Air Force Audit Agency has incorporated some RI3 questions into its audit worksheets, and the GAO is evaluating RI3 for questions that organizations may incorporate in its audit guide. We believe it is advantageous for financial managers to be familiar with the questions—and to be prepared to ask them when it appears that risk questions have not been asked or the answers have not been forthcoming.

University of Arizona Study (1987) *Participants viewed a Tucson, AZ Home for 20 Minutes*				
List Price	$65.9K	$71.9K	$77.9K	$83.9K
Estimated Reasonable Price	$63.6K	$67.6K	$70.1K	$69.5K

Figure 6. Results of Neal and Northcraft Study Showing Anchoring and Adjustment

Conclusion

Financial managers have an integral part to play in the identification, qualification, and quantification of program risks and uncertainty. They have a fiduciary responsibility to ensure appropriate budget formulation in the generation of cost estimates and in the use of funding to mitigate risks.

This requires accuracy in cost estimates, which must include accurate risk and uncertainty assessments throughout the program and should avoid the common mistakes that may bias risk analyses when using expert opinion. It is time that program financial managers aggressively start asking risk questions in order to provide credible analysis and advice to PMs.

We have pointed out ways to understand and to guard against past mistakes in assessing cost risk due to bias and subjectivity. The Air Force's RI3 Guide is one tool that can be used to identify and qualify risks having the potential to become issues on future programs. While no estimate is made with perfect data, and no expert opinion is completely accurate, we end with two thoughts:

> *The trouble with forecasting is that it is right too often for us to ignore, and wrong too often for us to completely rely upon it.*
>
> *"It is far better to foresee even without certainty than not to foresee at all."*[23]

> "It is time that program financial managers **aggressivey start asking** risk questions in order to **provide credible analysis and advice** to program managers."

ENDNOTES

[1] Morin, Jamie. "Achieving Acquisition Excellence in the Air Force: A Financial Management Perspective," Armed Forces Comptroller, Spring Issue 2010, p .11

[2] GAO-08-467SP Defense Acquisitions: Assessments of Selected Weapons Programs

[3] GAO-06-391 Defense Acquisitions: Assessments of Selected Weapons Programs

[4] Ibid.

[5] GAO-07-406SP Defense Acquisitions: Assessments of Selected Weapons Programs

[6] Ibid.

[7] Department of Defense, Technology Readiness Assessment Guidance, April 2011, pp. 2-13, Section 2.5

[8] GAO-09-3SP GAO Cost Estimating and Assessment Guide, pp. 37-38

[9] GAO-09-3SP GAO Cost Estimating and Assessment Guide, p. 173

[10] Ibid., p. 159

[11] Arena, Mark; Obaid Younossi, Lionel Galway, Bernard Fox, John Graser, Jerry Sollinger, Felicia Wu, and Carolyn Wong. "Impossible Certainty: Cost Risk Analysis for Air Force Systems." MG-415, Rand Corporation, 2006, p. 123

[12] Rand MG-415, p. 127

[13] Van Atta, Trevor, Chair of the Army Cost Risk Working Group, "The Unseen: Statistical Inference with Limited Data" in Society for Cost Estimation and Analysis, 2012

[14] Air Force Cost Risk and Uncertainty Analysis Handbook, p. vii

[15] Kahneman, Daniel and Tversky, Amos, "Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment" in *Psychological Review*, Volume 90, Number 4, pp. 293-315 (October 1983)

[16] Kahneman, Daniel: *Thinking, Fast and Slow*, Macmillan, 2011

[17] Charness, Gary et al.; "On the Conjunction Fallacy in Probability Judgment: New Experimental Evidence Regarding Linda"; 19 May 2009

[18] Yaniv, Ilan, Foster Dean P.; Precision and Accuracy of Judgmental Estimation; *Journal of Behavioral Decision Making*, Vol. 10, pp. 21-32 (1997)

[19] Neale, Margaret, Northcraft, Gregory, "Experts, Amateurs, and Real Estate: An Anchoring-and-Adjustment Perspective on Property Pricing Decisions," *Journal of Organizational Behavior and Human Decision Processes* 39(1), 84-97 (1987)

[20] Air Force Risk Identification, Integration and "illities" Guide, Version 1.2, 15 December 2008

[21] Air Force Pamphlet 62-128, Guide to Acquisition and Sustainment Life Cycle Management, 5 Oct 2009

[22] Air Force Risk Identification, Integration and "illities" Guide, Version 1.2, 15 December 2008, p. 5

[23] Poincaré, Henri in *The Foundations of Science*, p. 129. Poincaré is one of the foundation builders of chaos theory.

JOHN CARGILL

John Cargill is an operations research analyst focusing on air armament weapons systems costs at the Air Force Cost Analysis Agency, Eglin Air Force Base Operating Location. Previously, he served four years estimating costs of space weapons systems. He started as a financial manager in the Marine Corps and served in progressively responsive comptroller positions. He holds an MA in human resources from Pepperdine University, an MBA with accounting emphasis from National University, and an MS in Management from the Naval Postgraduate School. He is DAWIA Level III certified and a member of ASMC's Gulf Coast Chapter.

GARY MOORE

Gary Moore is an operations research analyst focusing on aircraft weapons systems costs at the Air Force Cost Analysis Agency, Wright-Patterson Air Force Base Operating Location. He has over eight years of hands-on experience estimating aircraft weapons systems costs and nearly five years of academic cost-estimating experience, both as a student and as an instructor. He holds a bachelor's degree in psychology from North Carolina State University and an MS in cost analysis from the Air Force Institute of Technology. He is a Certified Cost Estimator/Analyst and is DAWIA Level III certified.