



Annie Oakleying the Risk Cube

Management Goes Better with Analysis

Peter J. Braxton-PBraxton@technomics.net

David H. Brown-DBrown@technomics.net

Robert G. Fatzinger-RFatzinger@technomics.net

Sean M. Wells-SWells@technomics.net

International Cost Estimating and Analysis Association (ICEAA), Atlanta, GA

5/13/2025

Abstract

Despite advances in both Risk Management and Risk Analysis over the past two decades, there remains a persistent divide between the two disciplines. Risk Management, with its predisposition to discrete risks (“what could go wrong?”) and central construct of the Risk Cube, tends to have a myopic focus on Issues and struggles to accurately assess and portray Likelihood and Consequence. Risk Analysis, with its predisposition to continuous risks (“what is the range of possible outcomes?”) and central construct of the S-Curve, brings the mathematical tools needed for fidelity in risk calculations but struggles to put a recognizable face on these necessary abstractions. This paper seeks to unify these two fields by drawing the best from each. It identifies fundamental deficiencies in the Risk Cube, such as lack of basis, distortion of linear scales, misleading color codes, and junk math, and remedies them with principles from Risk Analysis to better inform decision-makers.

<https://www.iceaaonline.com/atl25sessions/#MESTrack>

Keywords: *Communication, IPM, Methods, Modeling, Risk, Risk Management, Risk Analysis*

Table of Contents

Abstract.....	i
Prologue	1
1.0 Risk Management (RM) vs. Risk Analysis (RA).....	1
1.1 Risk Management (RM).....	2
1.2 Risk Analysis (RA).....	3
2.0 Risk Current State of Practice	4
3.0 RM Notional Example	1
3.1 Notional Assumptions	2
3.2 Notional Risk Register	2
3.3 Notional Risk Viewer	3
3.4 Notional Risk Cube.....	4
4.0 Risk Framework	8
4.1 Framing Assumptions	8
4.2 Discrete Risks	11
4.2.1 Typical Metadata.....	13
4.2.2 Nested Risk Registers.....	14
4.3 Uncertainties	15
5.0 Risk Register Deficiencies and Remedies.....	18
5.1 Risk Viewer Improvements	18
5.2 Risk Register Improvements.....	18
6.0 Risk Cube Deficiencies and Remedies.....	19
6.1 Likelihood and Consequence.....	19
6.2 Severity and Junk Math	20
6.3 Ordinal, Interval, and Ratio Scales.....	24
7.0 Risk Data Visualization.....	28
7.1 Risk Management Visualizations	28
7.1.1 Centered Risk Cube Boundaries	28
7.1.2 Centered Risk Cube Granularity.....	31
7.1.3 Centered Risk Cube with Opportunities.....	34
7.1.4 Centered Risk Cube with Thresholds	35
7.2 Risk Analysis Visualizations.....	36
8.0 Risk Model Exemplar	37
8.1 Calibrating the Risk Cube	37

8.2 Enhanced Scenario Based Methodology	38
8.3 Implementing the Risk Framework.....	38
8.4 Risk Model Case Study	39
9.0 RM and RA: The Best of Both Worlds	40
9.2 Adopt the Risk Framework	40
9.3 Adopt the Centered Risk Cube	41
9.4 Maintain a Living Risk Model	42
Next Steps	44
ICSRA	44
Portfolio Management	44
Risk Benchmarks	45
Track Record.....	45
Bayesian Tripwires	46
Epilogue.....	47
Bibliography	48
Appendix A: Acronym List	49
Appendix B: Agency State of Practice Details	51
B.1 Department of Energy	51
B.2 Department of Homeland Security.....	52
B.2.1 Risk Management Current State of Practice at DHS.....	53
B.2.2 Risk Analysis Current State of Practice at DHS	56
B.3 Intelligence Community	60
B.4 Government of Canada	61
B.5 Industry / Shipbuilding	63
B.6 Department of Defense	64
Appendix C: RM According to ChatGPT.....	66
Appendix D: RM Notional Example Additional Details	69
Appendix E: Risk Aphorisms	72
Figure 1: 5-Step Risk Management Process	4
Figure 2: Risk Viewer Notional Example	4
Figure 3: Risk Cube per Mathematica (via ChatGPT)	7

Figure 4: Example Risk Register as presented in DOE Risk Management Guide	14
Figure 5: Centered Risk Cube	26
Figure 6: Centered Risk Cube Boundaries	29
Figure 7: Centered Risk Cube (Extended)	30
Figure 8: Center Risk Cube (Deciles).....	32
Figure 9: Centered Risk Cube (Icosatiles).....	33
Figure 10: Centered Risk Cube with Opportunities	35
Figure 11: Centered Risk Cube with Thresholds	36
Figure 12: DHS Risk Score Mapping.....	55
Figure 13: DHS Risk Priority Matrix and Actions	55
Figure 14: Risk and Issue Management Process Overview	64
Figure 15: Likelihood Ratings per ChatGPT	66
Figure 16: Consequence Ratings per ChatGPT	66
Figure 17: Conflated Risk Scores per ChatGPT.....	67
Figure 18: Notional Risk per ChatGPT	68
Figure 19: Risk Viewer Improvements.....	69
Table 1: Differing Views of Risk	2
Table 2: Federal Agency Survey of RM and RA Policy Maturity	1
Table 3: Notional Risk Cube.....	6
Table 4: Candidate Framing Assumptions by Commodity.....	10
Table 5: Candidate Discrete Risks by Commodity	12
Table 6: Candidate Uncertainties by Commodity	16
Table 7: Severity by Expected Cost Impact.....	21
Table 8: DHS Probability Risk Scale	54
Table 9: DHS Impact Assessment Scores	54
Table 10: Severity by Expected Schedule Impact.....	70

Prologue

*“Anything you can do, I can do better /
I can do anything better than you”*

-Annie Oakley, Annie Get Your Gun

Cost estimators have the annoying habit of thinking that we can outdo practitioners from other fields at their own disciplines. We believe we can develop a better Estimate At Complete (EAC) model than the Earned Value Management (EVM) folks. We believe we can model contract terms and conditions (Ts & Cs) better than the Contract Management folks. We believe we can develop more sophisticated regression models than the Data Science folks. (There we may have finally met our match, but that is the subject of another paper.)

Ultimately, healthy competition is a good thing. Significant contributions have been made by cost estimators indulging in interdisciplinary research, generally to the benefit of both fields.

This paper represents the latest foray by estimators into another discipline in the acquisition process, where we will make the case for “building a better Risk Cube.” The typical implementation of a Risk Cube within the Risk Management (RM) process has fundamental mathematical and logical flaws, for which we prescribe remedies. Even better, we put forth an overarching Risk Framework under which RM and Risk Analysis (RA) can work in tandem to more effectively support high-fidelity decision making.

1.0 Risk Management (RM) vs. Risk Analysis (RA)

It is well-established that there exist divergent views of risk and that parties often end up talking past each other. See for example Module 9 Cost and Schedule Risk Analysis of CEBok. (Cost Estimating Body of Knowledge (CEBoK), 2013) At the risk of

oversimplification, we will call the two primary viewpoints Risk Management (RM) and Risk Analysis (RA). The former is associated with program managers (PMs), engineers, and – naturally – risk managers. The latter is associated with cost analysts, statisticians, and other data analysts. These two fundamental viewpoints are summarized in Table 1 below. Keeping them in mind will help us craft the overarching Risk Framework that combines the best of both and answers the questions each party is interested in.

Table 1: Differing Views of Risk

Engineers / Deterministic / RM	Statisticians / Stochastic / RA
Variability is to be controlled (Lean Six Sigma)	Variability is to be expected (Stuff Happens)
Framing Assumptions and Conditional Estimates	Ground Rules and Assumptions (GR&A) and Unconditional Estimates
Root Cause Analysis (RCA)	Cost and Schedule Growth Analysis
Inputs Risk	Outputs Risk

The authors are decidedly from the RA camp, but we will do our best to “walk a mile in the shoes” of the RM viewpoint. Our inherent bias notwithstanding, we ultimately believe that each has much to gain from the other.

1.1 Risk Management (RM)

Risk Management is often expanded to “RIO” Management, taking into account Risks, Issues, and Opportunities (RIO). We will touch on all three, but our focus remains Risk, and we’ll stick with the RM shorthand throughout. PMs often adhere to the MAIMS principle, “Money Allocated Is Money Spent,” also known as the Self-Fulfilling Prophecy. They may use RM to determine a contingency or management reserve (MR), a sort of self-insurance retained at the management level. The key to managing risk is then, “Budget high, allocate low.” The more MR the PM retains, the more they are prepared for eventualities, but the more aggressive the allocated budgets, the more likely the teams executing the work are to fail initially and be disrupted by replanning. There is clearly a balance to be struck. RM tends to think in terms of discrete risks, “What could go wrong?”

The strengths of the RM process as described in the .0 RM Notional Example section below include consistency, visibility, and identification of dependencies, which are designed to enable action to affect outcomes. Risk reviews and configuration control boards are convened with a regular cadence. The Risk Cube has generally made its way from PowerPoint slides to live dashboards. For information technology (IT) systems of systems (SoS), we have even seen a formal DRIO process with a focus on Dependency Risks, Issues, and Opportunities.

The weaknesses of the RM process as it is typically implemented include its myopic focus on Issues and near-Issues, ordinal scoring, inconsistent mapping, and the resultant “garbage math” (or no math at all). These will be expanded and remedied below. There tends to be an endemic lack of historical data-driven standards. Each program, together with its associated risk scores, is treated as a unique snowflake.

1.2 Risk Analysis (RA)

RA embraces the Risk Paradox or Tail Chase wherein even though our cost estimates are based on final actuals that should include *all* growth, we continue to observe cost growth from those estimates. The schedule corollary is known as Hofstadter’s Law: “It always takes longer than you expect, even when you take into account Hofstadter’s Law.” RA practitioners seek to predict a growth factor or range of growth factors based on the risk profile of a particular program. Issues of retention and allocation aside, RA agrees with a healthy MR, enabling the program to “under-promise and over-deliver.” RA tends to think in terms of continuous risks, “What is the range of possible outcomes?” It is often important to pay attention to tail events, “What could *really* go wrong?!”

One weakness of RA is that, while it can identify cost and schedule drivers and other parameters that influence risk, it may not have the level of specificity to enable the formulation of actionable mitigation plans. Another weakness of RA is lack of timeliness. A significant amount of time and effort may be put into the development of a risk model for a milestone estimate, but then it may not be fully leveraged for the ongoing management of the program between then and the next milestone.

2.0 Risk Current State of Practice

Before formulating recommendations, we want to step back and assess the current state of practice for both RM and RA. In a nutshell, RM tends to adhere very well to policy and guidance, but the policy and guidance are themselves poor¹, allowing and even perpetuating many of the deficiencies we will discuss. RA is largely the opposite, with very thorough and detailed guidance, but adherence to that guidance is generally poor, with many analysts “driving” Monte Carlo models “without a license.” There has been both convergence and divergence over the past two decades in the Risk community when it comes to standard practice. In RM, the five-step process has become universal (see Figure 1).



Figure 1: 5-Step Risk Management Process

In RA, a handful of common approaches are prevalent, including Monte Carlo Simulation vs. Method of Moments (MoM) and Inputs vs. Outputs Risk. While the 5x5 Risk Cube is all but ubiquitous, there is wide variance in how the ordinal 1 through 5 scales for Likelihood (or Probability, Pf) and Consequence (or Impact, Cf) are mapped and implemented, most in need of improvement.

For some seminal guidance documents, see Appendix B: Agency State of Practice Details.

¹ Here we are specifically focused on guidance related to implementation and analysis of the Risk Cube.

Table 2 summarizes the current state of practice, comprising both *de jure* formal policy and guidance and *de facto* implementation. The rows represent key elements of the RM and RA processes. The columns reflect agencies with which the authors have direct recent experience. It is by no means intended to be an exhaustive government survey, but it should be diverse enough to give a sense of where and how practices vary due to “local custom.”

The survey attempted to examine the prevalence of RM and RA policy across a sample of Agencies facing some of the largest acquisition and project delivery initiatives in the United States and Canada. The survey examined the large project and acquisition risk policy and guidance of the Department of Defense (DoD), Department of Energy and the National Nuclear Security Administration (DOE/NNSA), Department of Homeland Security (DHS), the Intelligence Community (IC), the Royal Canadian Air Force major acquisitions, and the Canadian Shipbuilding and their National Shipbuilding Strategy (NSS). The authors examined and scored each Agency based on numerous detailed policy positions and then summarized these results into major categories. The rating areas included, but were not limited to, the following aspects of RA and RM practice:

- Risk Cube Application
- Cost and Schedule Risk Analyses
- Use of risk-adjusted estimates to inform Budgets, Acquisition, and Contingency
- Existence of clear guidance

Area of RA and RM Maturity	DoD	DOE/ NNSA	DHS	IC	Canada
				NGA, NRO	Future Aircrew Training (FACT)
RM Process	4.0	4.0	4.0	4.0	4.0
Risk Cube Application	4.0	4.0	4.0	3.0	3.0
Cost Risk Uncertainty Analysis (CRUA)	4.0	4.0	3.0	3.0	2.0
Schedule Risk Analysis (SRA)	3.0	4.0	2.0	2.0	2.0
Integrated Cost Schedule and Risk Analysis (ICSRA)	2.0	4.0	1.0	1.0	2.0
Budgets based on risk adjusted estimates	3.0	4.0	4.0	4.0	2.0
Acquisition based on risk adjusted position	4.0	4.0	4.0	4.0	4.0
Acquisition level (ACAT) determined by risk adjusted estimates	3.0	3.0	4.0	3.0	NA
Acquisition Affordability Determination	2.0	2.0	4.0	3.0	3.0
Management Reserve and Contingency established through risk adjusted estimates	2.0	4.0	2.0	2.0	3.0
Formal Risk Policy Established and Maintained	3.0	4.0	4.0	3.0	0.0
Risk Management Guidance Established and Maintained	4.0	4.0	4.0	4.0	4.0
Cost Risk Uncertainty Analysis Guidance	4.0	4.0	3.0	4.0	2.0
Schedule Risk Analysis Guidance Documented	2.0	4.0	0.0	1.0	2.0
Training Available	4.0	4.0	4.0	4.0	4.0

NA	0.0	1.0	2.0	3.0	4.0
Measure not applicable for this organization	Policy or practice is non-existent	Policy is ill defined (leveraged from outside agency) regarding acquisition by milestone	Policy is defined or required for use within the Agency; it may cover acquisition by milestone, and is slanted to a singular RA or RM view but not both	Policy is well defined, and agency tailored; it covers acquisition by milestone and includes some RA and RM techniques and tactics but lacks clarity and direction as to the proper application of the policy	Policy is well defined and integrated with related financial management policy; it covers acquisition by milestone and has both practices and guidance for RA and RM throughout the project lifecycle

Table 2: Federal Agency Survey of RM and RA Policy Maturity

The authors found that the Agencies and programs reviewed in our survey maintained some prevailing guidance related to both RA and RM practices. In most cases, Agencies had implemented, or were striving to implement, RA and RM practices in acquisition or cost estimating policy to support budget positions. Although difficult to score precisely in a clear and fair manner, many Agency policies did not provide clear guidance as to when specific RA practices are more appropriate (e.g., when attempting to provide a cost range around an early milestone, budgetary program estimate), compared to when specific RM practices are more useful (e.g., implementing risk management practices and budgeting reserve dollars for mitigation activities during execution). Note that the final scores included in Table 2 are qualitative and applied by the authors; however, they are the result of observations related to lower-level survey attributes for each category that were reviewed, evaluated, and tabulated to arrive at a rating displayed.

3.0 RM Notional Example

While we cannot catalog every variant of the Risk Cube Pf/Cf mapping, it is instructive to provide a typical implementation of RM from which to launch our critique. What follows is an amalgam of the basic RM implementations we have observed in our collective experience. Where there are variations on a theme, we will comment on common departures from this singular example.

The key components of this RM example are:

- **Assumptions:** framing ground rules and statements for the model
- **Risk Register:** the structure of the table itself, what data and metadata are included
- **Risk Viewer:** a means to view the detailed data and metadata for an individual risk without having to scroll down the rows (individual RIOs) and across the columns (data fields) of the table
- **Risk Cube:** The display of risks on a 5x5 square-of-squares (not really a cube!)
- **RIOs:** The contents of the Risk Register

3.1 Notional Assumptions

The risk register (described in detail below) should always be reviewed in context of the list of programmatic assumptions. Often referred to as ground rules, these assumptions help to bound the scope of risk analysis by excluding “*force majeure*” scenarios. This ensures that program managers can focus on the quantifiable and leave the unquantifiable to the contract clauses. In a more practical sense, having clearly defined assumptions can streamline conversations about risk mitigation and prevent devolvement into “what-ifs.”

In the notional example, the assumptions are associated with a major production contract and are included in the offeror’s (sole source) proposal. As such, they are focused on exclusions and don’t necessarily reflect the full risk posture of the government buyer. They are documented as part of the Master Data Assumptions List (MDAL).

3.2 Notional Risk Register

While “Risk Cube” and “Risk Register” are often used interchangeably, we will consistently treat the Risk Register as the tabular data and metadata related to RIOs and the Risk Cube as the graphical means of binning and/or displaying those data. A typical risk register contains several fields, including a Risk Description, Risk Type (dependent on the system being proposed), Status (open, accepted, mitigated, etc.), Responsible Party (contractor, government), and an If-Then statement. Impacts are scored across cost, schedule, and technical areas with a rated probability for each major contractor. The Risk Register also includes Mitigation Scores and Strategies.

Correspondence between risk score and Pf/Cf should be accomplished as lookups in the database, not hardwired into the single flat table of the Risk Register. In our notional example, the Risk Register is at a program level, shared jointly between the government and prime contractor. As such, it draws from lower-level risk registers, as those of major subcontractors, and it may also inherit up to a portfolio-level risk register for Enterprise Risks on the government side, though we have limited insight into both.

Issues, which are common in proposal risk registers, include:

- New risks are specified within an already identified risk. This causes a loss of traceability by not having singular and distinct risks per line item
- Mitigation strategies appear to be in columns other than the appropriate Mitigation column, such as the Description field, causing an issue with consistently documenting within the Risk Register
- Risks are in a variety of unclear statuses including “Rejected” risks
- Risk IDs have been skipped, and duplicate risks are included.

3.3 Notional Risk Viewer

The Risk Register is essentially a flat database table, with as many rows as there are RIOs (including those already closed and those drafted but not yet approved) and as many columns as there are data and metadata fields. This tends to make it less than ideal for viewing the details of an individual RIO. Accordingly, a Risk Viewer provides the means to see the details of any selected individual RIO via lookup in a more compact array, ranging from a laundry list to a form where related fields are grouped together. For a more robust and colorful example, see Appendix D: RM Notional Example Additional Details.

The notional Risk Viewer example is shown in Figure 2 below.

Title
Integration Risk
Project: PMO XYZ
RIO Type: Risk
Priority Level: Medium
Probability: Medium Impact: High
Escalation Data: 10/31/2024
Unique ID: XYZ-139
POC(s): Genevieve Desjardins – XYZ GOV
Ernest Thayer – XYZ GOV
Risk Owner: Dave Davidson – XYZ CTR
Risk Statement
IF subsystem interfaces are not well-documented, *THEN* integration may be delayed, incurring rework and additional cost.
Mitigation: XYZ Systems Engineering is working to identify and document all subsystem interfaces and other dependencies.

Figure 2: Risk Viewer Notional Example

Note that many RM implementations include a Risk Dashboard implemented in business intelligence (BI) software such as Tableau or Qlik, or in a custom RM tool. In those cases, the dashboard usually comprises both Risk Viewer and Risk Cube functionality. We split them out here for the sake of clarity.

3.4 Notional Risk Cube

The typical risk cube has adjectival ratings for both Pf and Cf for the ordinal 1-through-5 scales. For Likelihood, a range of probabilities may be given for each bucket, starting at

0% (or 1%) for the low end of 1 and ending at 100% (or 99%) for the high end of 5. For Consequence, ranges may be given for Cost, Schedule, and/or Technical impacts. Cost is almost always included. Cost and Schedule are usually given in absolute terms, such as millions of dollars of cost growth and months of schedule slip, respectively, but occasionally they will be in relative terms (i.e., percentage growth) instead. Technical impacts may be called Performance or Mission instead, and their descriptions are almost always qualitative.

There is usually a “junk math” attempt to combine Pf and Cf into a Severity or Priority, which ranges from a three-point scale (Low, Medium, High) to a five-point scale (akin to the two component rankings) to a 25-point scale (sorting all of the smaller squares in the Risk Cube “in order”). Even when this combining ranking isn’t shown as a separate field in the Risk Register, the three-point scale is commonly reflected as the green, yellow, and red shading of the Risk Cube, though without community consensus as to which squares should be which color.

The Risk Cube for our notional example is shown in Table 3 below. The Likelihood scale uses letters (A through E) to distinguish it from the Impact scale (1 through 5). This enables an alphanumeric “coordinate system” of sorts. For example, 5C denotes a “Severe” impact with a “Possible” likelihood. This avoids having to remember which rating comes first between (5,3) and (3,5). In fact, this risk cube puts Likelihood on the horizontal and Impact on the vertical, which is the reverse of the typical implementation. There is a combined Risk Score on a five-point scale: Very High (bright red), High (dark red), Significant (orange), Moderate (yellow), and Low (green). As usual, there is no basis for which squares are assigned to which score, and in fact, we will later see that this color-coding is inconsistent with the expected values of the respective squares.

Table 3: Notional Risk Cube

IMPACT / CONSEQUENCE				Likelihood					Score	
				Rare <20%	Unlikely >20% < 40%	Possible >41% < 60%	Likely >60% < 80%	Almost Certain >81%		
Cost	Schedule	Technical		A	B	C	D	E		
>= \$100M	> 52 weeks slips to an identified cardinal milestone	Platform- lose a major capability	5 Severe	5A	5B	5C	5D	5E	Very High	
\$25M to <\$100M	26 to 52 weeks slip to an identified cardinal milestone	Platform - degradation of a major capability	4 Major	4A	4B	4C	4D	4E	High	
\$5M to <\$25M	12 to 26 weeks slip to an identified cardinal milestone	Platform - loss or degradation of a system	3 Moderate	3A	3B	3C	3D	3E	Significant	
\$1M to <\$5M	4 to 12 weeks slip to an identified cardinal milestone	Subsystem - experience a reduction in technical/expected performance	2 Minor	2A	2B	2C	2D	2E	Moderate	
<\$1M	<4 weeks slip to an identified cardinal milestone	No significant reduction	1 Insignificant	1A	1B	1C	1D	1E	Low	

While not always clear, the prevailing practice is to bin risks into the 5x5 Risk Cube *first*, and then (if at all) map the bin to actual Pf and Cf values. When asked to generate a typical example of a Risk Cube, ChatGPT actually produced Figure 3 below from Mathematica! Only an ordinal scale is evident, with adjectival descriptions for both Pf (ranging from *Rare* to *Almost Certain*, vertically) and Cf (ranging from *Negligible* to *Catastrophic*, horizontally). There is a four-point Severity scale (*Low*, *Moderate*, *High*, and *Critical*) shown on the Risk Cube with no basis, but the strictly diagonal bands seem to imply sum Pf and Cf. Based on a separate prompt, ChatGPT produces a similar four-point Severity scale based on the product of Pf and Cf. These are both instances of junk math and dangerously misleading. The analyst must resist any temptation to add or multiply risk scores meaninglessly. We will explore this further in section 6.3 Ordinal, Interval, and Ratio Scales below.

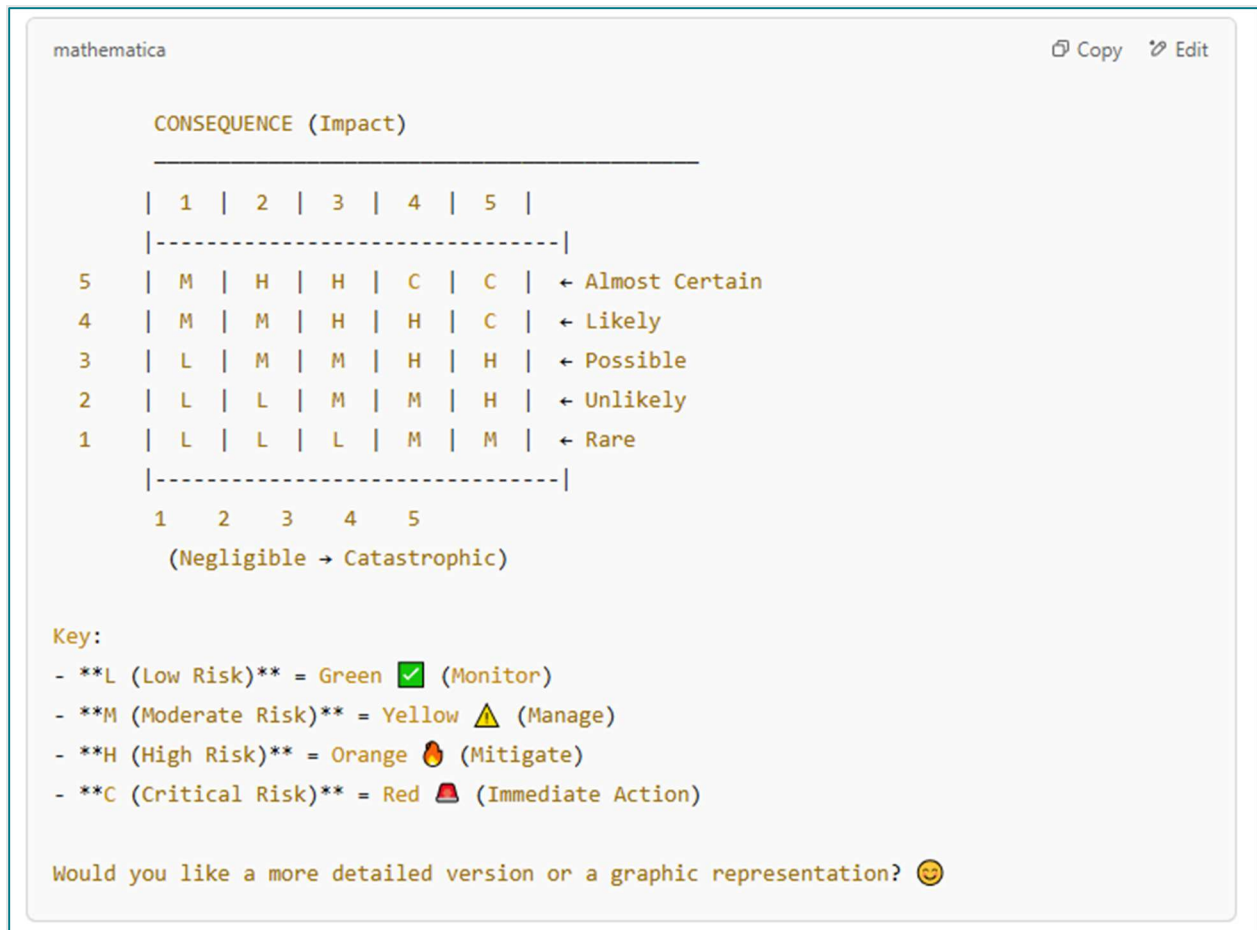


Figure 3: Risk Cube per Mathematica (via ChatGPT)

For more on ChatGPT's responses to Risk Cube prompts, see Appendix C: RM According to ChatGPT.

As a notable exception, one government organization uses a 15-unit Impact scale on the horizontal axis and shows a more granular green/yellow/red shading, but it does so for the wrong reasons: the junk math of adding 1-5 scores for Cost, Schedule, and Technical together!

4.0 Risk Framework

The Program Manager's Serenity Prayer:

Grant me the serenity to accept the risks I cannot mitigate; the courage to mitigate the risk I can; and the wisdom to know the difference.

The following general framework for identification and treatment of potential risks, issues, and opportunities is grounded in experience with shipbuilding proposals. This framework supports both those building up the risk register (i.e., the risk managers) and those evaluating the consequences for a proposal and for the government (i.e., the risk analysts). The framework groups the common components of a risk register into three groupings:

- *Framing Assumptions* should not be included in the Risk Register and are instead more appropriate for the Ground Rules and Assumptions list
- *Discrete Risks* should always be included in the Risk Register
- *Uncertainties* should generally be addressed in the probabilistic cost estimate, though they can be included in the Risk Register instead if the range of values / probability distribution allows for quantifying an impact

The three lists below are not intended to be comprehensive, but they do highlight many of the common items in each grouping:

4.1 Framing Assumptions

Framing Assumptions are the problem level assumptions upon which cost and schedule estimates fundamentally rest. Therefore, any substantive violation of a framing assumption would call for a new estimate. While the spirit of risk and uncertainty analysis is to be unconditional (considering all possible outcomes, however unlikely), it is far more practical to be conditional on a set of high-level framing assumptions. Put another way, many of these items would not be included in contract risks, but rather in insurance policies.

These items are not to be included in the Risk Register at hand but may need to be bubbled up to the portfolio or enterprise level to be monitored, managed, and self-insured against. A prime contractor may propose building some of these Framing Assumptions into contract Terms & Conditions (Ts & Cs) to take the associated cost risk “off the shareline.” Some risk analysis may need to be conducted during negotiation to determine what degree of liability would accrue on the government side, were these Ts & Cs to be accepted.

Typical framing assumptions include the following:

- **Acquisition Strategy:** the government will adhere to the shipyard workshare for the program (usually building all ships at one yard or a back-and-forth delivery schedule between two yards)
- **Procurement Profile:** the government will adhere to planned quantities for the program
- **Funding and Contracting:** the government will ensure program is adequately funded and will complete contracting actions in a timely manner so as to avoid unnecessary delay and disruption.
- **Base Design:** the base design, which is often military off-the-shelf (MOTS), is well-established. This framing assumption does not obviate all design-related risk, but it does stipulate that the program won't have to “go back to the drawing” and start over with a blank-sheet-of-paper design.
- **Black Swans:** the production program will not be adversely affected by a global pandemic, terrorist attack, extreme weather event, or other calamity. Sadly, in the current millennium, this assumption seems to be getting more and more tenuous – hence the popularity of the term.

A more detailed breakout of candidate Framing Assumptions by commodity is shown in Table 4 below.

Table 4: Candidate Framing Assumptions by Commodity

Area	Subarea	Shipbuilding	CapAx	SW/IT
Acquisition Strategy	Competition	Ship type and IB; program swaps	Open Competition VS. Limited specialized competition or Sole Source	Generally competitive. No impact on cost.
Acquisition Strategy	Work Share	Dual yards	Design-Build VS. Design-Bid-Build	Generally n/a because multiple competing awards are rare
Procurement Profile	Quantity	Blue collar labor (learning)	n/a (subsystem level?)	Software requirements and associated size
Procurement Profile	Build Sequence	Preserve build sequence for compartments, blocks, and grand blocks	n/a (sourcing subsystems across multiple facilities?)	SW release schedule
Program Stability	Funding	Retaining skilled workforce	Affordability and funding profile (front loading VS. level programming)	Affordability and funding profile
Program Stability	Contracting	Avoid stop work or out-of-sequence work	The volume of capable trades available in geographic locations	Possible cost implication for redundant contracts during a transition period
Technical Solution	Requirements Stability	Avoid costly changes	Scope growth through concept maturity	Software growth; scope creep; agile re-prioritization
Technical Solution	Design Approach	Military off-the-shelf (MOTS) vs. new design	Standard Constructions VS. Specialized Capability or Unique Hazards	Developed versus COTS SW
Black Swans	Pandemic	Retaining skilled workforce	Supply chain interruption	Work form home impact
Black Swans	Terrorist Attack	Shipyard, supply networks	Varying levels of security and potential interest	Cybersecurity breach
Black Swans	Extreme Weather	Waterfront facilities	Varying climates, always a potential factor	n/a

4.2 Discrete Risks

Discrete risks are to be included in the Risk Register delivered with the proposal for the purposes of informing the government's risk reserve. They are *not* priced into the base estimate. It helpful to parse into Cost-driven and Schedule-driven risks as this enables the dollarization of Schedule-driven risks via a "standing army" monthly burn rate under the Cost Informed by Schedule Method (CISM). If a true integrated cost and schedule risk analysis (ICSRA) is being performed, risk items need to be linked to the Integrated Master Schedule (IMS) and cost impacts tagged as time-dependent or time-independent.

- **Requirements Changes:** If there are fundamental requirements changes, resultant redesign will both drive up non-recurring efforts and affect production cost, depending on the nature of the change. Generally, increased capability and/or disruption would lead to higher cost, whereas a concerted design for affordability effort might lead to lower cost
- **Standards/Classification:** If the class authority changes or standards are not well understood, it may result in redesign and/or additional cost and schedule for documentation and reviews
- **Digital Engineering:** Adoption of new design tools may increase design cost/schedule (learning curve) but decrease production cost/schedule (less rework, more efficient early outfitting, ability to virtually "walk the space" and spot conflicts and design flaws).
- **Workforce:** Inability to hire the requisite blue-collar workers causes increased cost (having to pay overtime premiums) and/or schedule delays (reduced yard capacity)
- **Concurrent Programs:** Other government and/or commercial programs may disrupt scheduled facilities (aka the traffic jam problem) and/or labor resources
- **Supply Chain:** Supplier issues may drive price increases and/or delays in Material
- **Integration:** If interfaces are not well-documented, integration of Contractor- and/or Government-Furnished Equipment (CFE/GFE) subsystems may take longer and be more expensive than planned

- Capital Improvements:** Planned improvements may fail to yield projected benefits, constituting a risk if the savings are baked-in to the point estimate or may yield additional benefits, constituting an opportunity if savings are not baked-in

A more detailed breakout of candidate Discrete Risks by commodity is shown in Table 5 below

Table 5: Candidate Discrete Risks by Commodity

Area	Subarea	Shipbuilding	CapAx	SW/IT
Requirements Changes		Disruption, out-of-sequence work	Requirements growing or changing as the project matures through natural phase-gates.	Captured as uncertainty in software size/growth
Standards/Classification		"Build-to" standards drives effort	Standards in code and environmental regulations changes due to the long project planning and construction cycles.	Captured as uncertainty in software size/growth
Digital Engineering		Manufacturing integration	Potential for opportunities in efficiency	Captured as uncertainty in productivity
Workforce		Competition with other industries (e.g., Oil and Gas for welders)	Large scale projects greatly exceed local trade capacity - especially when considering more sensitive and secure facilities requiring craft resources to hold full citizenship or clearances.	Can be a risk for highly specialized and/or staff requiring security clearance. Captured in uncertainty on labor rates.
Concurrent Programs		Delicate balance of yard capacity	Multiple building projects in one location occurring simultaneously	n/a
Supply Chain	Throughput	Steel, long-lead items	Building material shortages	n/a

Supply Chain	Chain of Custody (Cyber)	Combat systems	Building cleared facilities in an unclassified environment to secure in a later stage commissioning turnover to a fully secured environment.	n/a
Integration		Manifests during Test and Trials	Specialized material processing equipment integration into a new or refurbished facility.	Captured as uncertainty in software size/growth
Capital Improvements		Covered spaces, cranes, robotic welding	Other supporting capital investments or improvements failing to come on-line as scheduled, or failing to integrate properly.	n/a
Black Swans	Terrorist Attack	Shipyard, supply networks	Varying levels of security and potential interest	Cybersecurity breach
Black Swans	Extreme Weather	Waterfront facilities	Varying climates, always a potential factor	n/a

4.2.1 Typical Metadata

As shown in section 3.3 Notional Risk Viewer above, RIO metadata should include:

- Project Title
- Categorization as Risk, Opportunity
- Priority Level (i.e., Low, Medium, High, Urgent)
- Probability (Low-Medium-High) and Impact (Low-Medium-High)
- A unique identifier

- POC(s) / Risk Owner
- Risk Statement (in If-Then format)
- Mitigation Statement

Figure 4 further illustrates the functionality and application of common metadata in a risk register. This example from DOE Order DOE G 413.3-7A Chg 2 (LtdChg), Risk Management Guide includes standard risk identification and grading metadata.

Risk #	Owner	Risk Description	Residual Risk									
			Likelihood	Consequence	Risk Score/Rank	Probability %	Cost Impacts (\$)			Schedule Impacts (mo)		
							Best Case	Most Likely	Worst Case	Best Case	Most Likely	Worst Case
T47	Federal	Nonperformance of contract to provide shielded overpack containers leads to project delays and cost.	Unlikely	Significant	Moderate	40%	850,000	3,000,000	5,000,000	\$0	\$0	\$0
T52	Federal	Oversight organization's interpret requirements different than implementation, leading to cost and schedule impacts.	Likely	Significant	Moderate	60%	-	3,000,000	5,000,000	\$0	1 month	3 months
T12	Contractor	Failure of crane results in delayed removal of canisters, impacting schedule.	Unlikely	Marginal	Low	40%	100,000	200,000	1,400,000	1 day	2 days	2 weeks
T61	Contractor	Calibration services are unavailable causing shut down of operations.	Very Unlikely	Marginal	Low	10%	100,000	410,000	715,000	1 day	4 days	7 days
T266	Contractor	Hot cell cannot be designed to meet active ventilation strategy increasing design and construction cost.	Very Unlikely	Critical	Moderate	10%	3,200,000	7,000,000	20,000,000	1 month	2 months	5 months

Figure 4: Example Risk Register as presented in DOE Risk Management Guide

4.2.2 Nested Risk Registers

It is advantageous if risk registers can be maintained at the portfolio, program, project (prime contractor), and sub-project (subcontractor) levels. In most cases, individual RIOs will “live” at a single level, wherever is deemed most appropriate. Mechanisms can be put in place to escalate RIOs to a higher level, if they affect multiple components and can be more effectively mitigated there, or conversely to flow them down to where impacts and mitigation strategies are more tailored. A digital engineering (DE) tool, such as Jira for agile development, facilitates tagging using “Parent of” and “Child of” relationships.

Risks should be sized in absolute terms but then displayed on a relative Risk Cube aligned to the purview of the decision maker (“a big fish in a small pond”). Galaxy

Charts are a stellar visual for demonstrating the absolute and relative size of risks.
(Nehring, Mann, & Jones, 2012)

4.3 Uncertainties

These are continuous risks that are better included in the uncertainty of the base estimate. At heart, it is not a matter of *whether* they will happen, only a question of the range of potential impacts. Any good estimator's intent is that the point estimate reflects a "most likely" value of these parameters, so that there is a near-net-zero impact from a risk perspective. However, if the parameter value associated with the point estimate is aggressive (optimistic), then the uncertainty will induce a net risk. Conversely, if the parameter value is conservative (pessimistic), then the uncertainty will induce a net opportunity.

- **Inflation:** Realized inflation will be lower or higher than steady state projections. Also, escalation, reflecting the specific basket of goods comprising the shipbuilding program, will differ from inflation due to real price change.
- **FOREX:** Fluctuation in foreign exchange rates will affect prices paid in one dollar type for any goods and services priced in other currencies
- **Productivity:** Realized productivity rates will differ from baseline value in proposal.
- **Learning Curve:** Realized learning curve slope and T1 adder will differ from baseline values in proposal.
- **ECPs:** Type I and Type II changes will vary from the planned/budgeted amount.; often shown as an analogy factor from other programs
- **Rework:** Realized rework rates will differ from baseline value in proposal.
- **Supervision:** Realized factors for supervision and other support labor will differ from baseline values in proposal.
- **Labor Rates:** Realized blended labor rates will differ from baseline values in proposal. This is due not so much to variation in hourly rates (largely covered by the Inflation item above) as departure from planned allocation across labor categories. The risk is that a richer labor mix is required, whereas the opportunity is that the work can successfully executed with a greener (and hence less expensive) labor mix.

A more detailed breakout of candidate Uncertainties by commodity is shown in Table 6 below.

Table 6: Candidate Uncertainties by Commodity

Area	Subarea	Shipbuilding	CapAx	SW/IT
Economic	Inflation / Escalation	Steel, long-lead items	Price sensitive building materials and labor	No commodity-specific indices, but uncertainty applies to general inflation
Economic	FOREX	Multinational project team	For the most part, we are dealing with federal infrastructure projects within the continental United States	n/a
Economic	Labor Rates	In-demand trades (welding, electrician)	Fluctuations often manifest in the cleared labor in isolated locations	Rate uncertainty due to market variance; clearance level; domain knowledge (AI, algorithms)
Labor	Labor Mix	White Collar	Lower impact in classic construction trades	LCAT uncertainty due to relative mix of PM, SE, developers
Labor	Staffing Level	White Collar	Highly variable	Calculated based on the estimating method. Always contains uncertainty.
ECPs		Lead ship, block upgrades	Higher potential to overlap or double count with discrete risks	Highly uncertain early in the lifecycle.
Rework		Shipyard maturity	Varies with contract type and processing equipment complexity	Highly uncertain early in the lifecycle.
Sizing	Sizing (Point Estimate)	tons of steel; ft of cable; sq ft of pipe	Sq ft; equipment sizing, equipment layout	FP, ESLOC, Epics, etc.

			assumptions, scope growth	
Sizing	Growth	Design Margin	Law out and occupancy variances	Scope Creep; Code Growth; Reuse Optimism
Productivity		Performance to Standards	Trade productivity Weather, location, hazard are all contributing factors to consider in productivity	Highly uncertain due to varying productivity based on size, development method, tools used, amount of re-use, type of software
Cost Improvement	Learning Curve	Complexity of design, spacing of hulls	More limited for CapAx unless there is a significant machinery or plant installation / construction sequence that can benefit from learned efficiency	n/a
Cost Improvement	Compute & Store	Command and Control	n/a	Moore's Law / Kryder's Law
Supervision		Chargehands	The amount of Supervision to the necessary Craft labor can get out of alignments with estimates when project delays occur or staffing shortages occur	Highly uncertain, and generally not visible to anyone outside the development contract organization

5.0 Risk Register Deficiencies and Remedies

The immediate issues with our notional Risk Register are missing risks, redundant risks, risks in uncertain statuses, and risks that should be split out or combined. For the risks that remain, many are schedule-driven, where we don't expect direct cost impacts, the existence of a Cost Cf notwithstanding. Conversely, a few are cost-driven, where we deem to ignore the Schedule Cf.

Most crucially, there is considerable ambiguity in how the risk scores are generated and how they are applied.

5.1 Risk Viewer Improvements

The Risk Viewer should provide easy access to all relevant data and metadata for a given RIO. Raw Pf and Cf values and their associated mappings should be color-coded. For more, see Appendix D: RM Notional Example Additional Details.

5.2 Risk Register Improvements

RM often employs a configuration control process to determine when new RIOs are officially added to the Risk Register and subsequently mitigated or retired. The team should do its utmost to ensure that the Risk Register is a complete representation of Discrete Risks on the program at a given point in time. In addition to the "push" wherein a PM or Control Account Manager (CAM) comes forward with RIO candidates, there should be a proactive "pull" on the part of the Risk Manager, using the Risk Framework checklists, facilitating risk workshops, and leveraging other engineering and management fora. As much as possible, RIOs should be mutually exclusive and exhaustive ("no gaps or overlaps").

For RIO data, Pf and Cf values should be provided directly to Risk Register data fields using appropriate units: decimals or percentages for Probability; \$, \$K, or \$M for Cost Impacts; days, weeks, or months for Schedule Impact. Only then should Likelihood and Impact ratings be determined as calculated fields in the database. If the register does allow ratings to be entered, they should be mapped to the center of the corresponding square on the Risk Cube (i.e., average of the Probability and Impact interval endpoints). If possible, where Cost and Schedule Impacts are both included for a risk, they should

be synchronized with an integrated cost and schedule risk analysis (ICSRA). (See also the ICSRA section under Next Steps.) Free text fields should be included for capturing the basis and rationale for Pf and Cf assessments.

Where multiple ratings are entered (e.g., Cost, Schedule, and Technical), they should be neither added, nor averaged, nor multiplied. The ratings represent apples and oranges; what does it mean to average a \$10M cost growth and a 4-week schedule slip?!

Data fields should enable linking the risk directly to work breakdown structure (WBS) elements for Cost impacts and integrated master schedule (IMS) activities for Schedule impacts. This would enable an alternate relative scale for impacts. For example, the risk could be expressed as a percentage growth to the cost of that WBS element or duration of that activity.

Finally, there should be adequate fields to capture information related to Risk Mitigation, essentially a mini-business case to demonstrate that the resources needed to invest in the mitigation actions in the first place are more than outweighed by the resultant risk avoidance. (Put another way, don't spend more to Mitigate a risk than it would cost you simply to Accept it!) There should be a well-defined status field to indicate whether the Mitigation actions are complete, in process, or not yet started. Reduced post-Mitigation Pf and Cf values should be captured.

6.0 Risk Cube Deficiencies and Remedies

Now that we've addressed how to improve the Risk Register itself, we turn to the Risk Cube and its associated calculations.

6.1 Likelihood and Consequence

One of the virtues of the Risk Cube is that it enables most or all of the active program risks to be compactly plotted on one display. Dashboards will often apply "jitter" to points so that risks with identical Pf and Cf ratings won't directly overlay each other. A better approach is to plot using the actual Pf and Cf values, like a Cartesian plane,

which avoids putting all the risks in the center of squares. Before we delve into the visual aspects of the Risk Cube, we first address the computational aspects.

The Notional Risk Cube is problematic in several ways, but we are able to remedy some of these by splitting the parameters into Probability (Likelihood) and Impact (Consequence) tables. Each of the five bins was assigned consistent quantitative boundaries, where the high end for one bin is equal to the low end of the next bin. For example, Likelihood C starts at 40%, not 41%. (As originally labeled, a probability assessment of 40.5% would have no bin!). Keep in mind that probability is a continuous scale, and we use percentages as a common frame of reference. There is nothing special about one percentage point, and it certainly should not be used as a gap between

Midpoints were used for the purposes of expected value calculations. For example, a “4C” would have an Impact of \$62.5M, splitting the difference between \$25M and \$100M, and a Probability of 50%, splitting the difference between 40% and 60%. Because the highest Impact bin, “Category 5,” is open-ended, we set the high end as twice the low end for purposes of expected value calculations. That is, Cost Impact is between \$100M and \$200M, with a midpoint of \$150M, and Schedule Impact is between one year (52 weeks) and two years (104 weeks), with a midpoint of 78 weeks.

A major deficiency in risk analysis frequently observed by the team is the use of the Risk Cube in Impact and Probability assessments. Individual risk items should *not* be plotted on the Risk Cube first, in which case their bin (e.g., “4C”) is used to determine their Impact and Probability. Rather, the proper order of operations is to assess their Impact and Likelihood first, *with Basis and Rationale*, and then plot on the Risk Cube for visualization purposes only. This display is somewhat useful in helpful in focusing management attention on the most critical risks but is useless for Risk Analysis purposes.

6.2 Severity and Junk Math

Even in its role as a risk triage tool, the Risk Cube falls short. The 1 through 5 bins on either axis give the illusion of a linear scale. While the Probability bins on the horizontal

axis are more straightforward, they are not a ratio scale as you might expect from the top-end probabilities of 20%, 40%, 60%, and 80%. When you look at the representative midpoints (10%, 30%, 50%, 70%, and 90%), it is evident that this is only an interval scale, and the ratios are quite unexpected. A score of 2 (or B) is not twice a 1 (or A), but rather three times! Similarly, a 5 (E) is three times a 2 (B). The problem is even worse on the vertical axis, where the Cost and Schedule Impacts are not at all linear and are not even proportional to each other. As previously noted, the open-endedness of the top score of 5 also causes ambiguity in the expected value calculation (infinite, on average?!).

These foibles manifest in the attempt to rank the 25 squares of the risk cube in order and to further group them into yet another 1 through 5 set of buckets that the contractor calls “Severity.” What we quickly see is that their color coding is inconsistent and matches neither the expected Cost Impact nor the expected Schedule Impact.

Table 7: Severity by Expected Cost Impact

1A	1	A	0.10	\$	0.5	\$	0.050	1	Low	Green
1B	1	B	0.30	\$	0.5	\$	0.150	2	Low	Green
1C	1	C	0.50	\$	0.5	\$	0.250	3	Low	Green
2A	2	A	0.10	\$	3.0	\$	0.300	4	Low	Green
1D	1	D	0.70	\$	0.5	\$	0.350	5	Moderate	Yellow
1E	1	E	0.90	\$	0.5	\$	0.450	6	Moderate	Yellow
2B	2	B	0.30	\$	3.0	\$	0.900	7	Low	Green
3A	3	A	0.10	\$	15.0	\$	1.500	8.5	Low	Green
2C	2	C	0.50	\$	3.0	\$	1.500	8.5	Moderate	Yellow
2D	2	D	0.70	\$	3.0	\$	2.100	10	Moderate	Yellow
2E	2	E	0.90	\$	3.0	\$	2.700	11	Significant	Orange
3B	3	B	0.30	\$	15.0	\$	4.500	12	Moderate	Yellow
4A	4	A	0.10	\$	62.5	\$	6.250	13	Moderate	Yellow
3C	3	C	0.50	\$	15.0	\$	7.500	14	Significant	Orange
3D	3	D	0.70	\$	15.0	\$	10.500	15	Significant	Orange
3E	3	E	0.90	\$	15.0	\$	13.500	16	High	Red
5A	5	A	0.10	\$	150.0	\$	15.000	17	Significant	Orange
4B	4	B	0.30	\$	62.5	\$	18.750	18	Significant	Orange
4C	4	C	0.50	\$	62.5	\$	31.250	19	High	Red
4D	4	D	0.70	\$	62.5	\$	43.750	20	High	Red

5B	5	B	0.30	\$	150.0	\$	45.000	21	High
4E	4	E	0.90	\$	62.5	\$	56.250	22	Very High
5C	5	C	0.50	\$	150.0	\$	75.000	23	High
5D	5	D	0.70	\$	150.0	\$	105.000	24	Very High
5E	5	E	0.90	\$	150.0	\$	135.000	25	Very High

In **Table 3 above**, we see the undesirable fact that the Severity categories as labeled are interleaved. For example, some “Moderate” bins (1D and 1E) have *lower* expected values than some “Low” bins (2B and 3A), while other “Moderate” bins (3B and 4A) have *higher* expected values than a “Significant” bin (2E). Similarly, two “Significant” bins (5A and 4B) outpace a “High” bin (3E), and a “High” bin (5C) outpaces a “Very High” bin (4E). Of course, any calculations involving an Impact score of 5 are sensitive to the somewhat arbitrary 2x top end.

Even the center of each square can represent other point in that square only so well. (We might call this the southwest tip of Virginia problem. Fun fact: The southwestern extremity of the Commonwealth of Virginia is closer to a handful of *other* state capitals than it is to Richmond!) The upper right (NE) corner of one square is likely to have a higher expected value than the lower left (SW) corner of either adjacent square (i.e., slightly lower Likelihood but much higher Consequence, or slightly lower Consequence but much higher Likelihood).

Take square 3C in comparison to its “riskier” neighbors 4C and 3D. 3C contains the point (59%, \$24M). 4C contains (41%, \$26M), and 3D contains (61%, \$6M). Crunching the numbers to calculate expected values we have:

$$59\% \times \$24M = \$14.16M$$

$$41\% \times \$26M = \$10.66M$$

$$61\% \times \$6M = \$3.66M$$

Thus, the riskier parts of a less risky “neighborhood” are more risky than the less risky parts of riskier neighborhoods, often by a wide margin! This impact can be minimized by slicing into a larger number of smaller squares, as we’ll see below. In the end, though, swaths of equal expected value cannot be made to follow the jagged stairsteps of the Risk Cube; they are actually circular arcs.

Another concern is the averaging of the three individual Cost, Schedule, and Technical scores to produce an overall Impact score. While this is somewhat intuitive and consistently produces a number between 1 and 5, inclusive, there is no reason to believe that averaging three distinct ordinal scales, the third of which is qualitative only, will produce a meaningful result. Even worse is the multiplying of 1 through 5 Cost, Schedule, Technical, *and* Likelihood scores! Thankfully, this is buried in the bowels of the Risk Register table and doesn't appear to be used anywhere, nor should it be.

As shown on the Viewer tab, the Overall Impact and Probability scores are taken as the maximum values across the assessments of the four contractors. This begs the question of whether these assessments are their view of the overall program risk, in which case they should be used to build consensus or be otherwise combined as in a Delphi method, or whether they are their Ptolemaic view of the impacts to only their scope of work, in which case they should be combined proportionally to those relative scopes. As described above, any effort to combine the 1 through 5 scores in a meaningful way is fraught with peril.

Both the Cost (dollars) and Schedule (weeks) Impacts are expressed in absolute terms, for example a growth of \$5M or a slip of 4 weeks. In performing risk assessments, it would be more useful to use a relative (percentage) scale calibrated to historically observed cost and schedule growth. These percentages can then be applied at any level (program, system, subsystem, contractor, trades, etc.) to determine quantitatively the cost and schedule impacts. This recommended approach is more fully described in the cited paper. (Coleman, Summerville, & Dameron, 2006)

A Risk Register is by its nature predisposed to discrete risks and opportunities. Many items are better suited to continuous uncertainty. In this case, the item could be in the Risk Register with a range of possible Impact values (again, with Basis) or it could be treated in the probabilistic cost estimate, such as via a rate or factor. Either way, care should be taken to avoid double-counting. Any fact-of-life issues should be "baked in" to the cost estimate. The bottom line is that "All risk is relative," and it should be made clear what assumptions comprise the basis for the point estimate and how deviations from those assumptions will affect cost and schedule.

6.3 Ordinal, Interval, and Ratio Scales

ICEAA Lifetime Award winner Dick Coleman originated the analogy of the “fence posts” and the “fence rails” for a set of bins and their endpoints, irrespectively. Picture a split rail fence, and note that if there are n sets of horizontal fence rails, there will always be $n+1$ vertical fence posts. We encounter this all the time when dealing with intervals. For example, if you count off January 15th, February 15th, ..., July 15th, you will have seven months (“fence posts”) but only a six-month span of time (“fence rails”).

In the context of the Risk Cube, take the more straightforward Pf bins with end points 0%, 20%, 40%, 60%, 80%, and 100% -- six fence posts. There are, of course, five bins (fence rails), which we label 1 through 5 ... *in the center* of each bin. Herein lies the fundamental disconnect. The midpoints of the bins are 10%, 30%, 50%, 70%, and 90%. Do you see the problem?

To do so, we need to recall the hierarchy of data types from Module 4 Data Collection and Normalization. (Cost Estimating Body of Knowledge (CEBoK), 2013)

- **Qualitative (Categorical):** Categorical data can be sorted into bins, but those bins don't have an inherent order. A relevant example is Risks, Issues, and Opportunities. Using business rules, we can easily classify each RM item as one of the three. While you may be tempted to say there is an inherent order, *Issues > Risks > Opportunities*, note that a Risk with a small Pf but sizeable Cf may have a much larger expected value than a modest Issue that has already been realized.
- **Quantitative:** Almost all the data we deal with – even binned data – behave as quantitative or numerical. The question is where they fall in the following hierarchy:
- **Ordinal:** At the most essential level, the data have an inherent order. This is certainly true of both Pf and Cf. A higher number represents a greater probability or a greater impact.
- **Interval:** In interval data, the *distance* between numbers (or scores or bins) has a consistent meaning. This holds for the typical Pf scale. For example, going

from a 1 to a 3 (an increase of 2) represents a 40% increase in probability (50% - 10%, using the midpoints), whereas going from a 1 to a 4 (an increase of 3) represents a 60% increase (70% - 10%). These two intervals are proportional ($3/2 = 60\%/40\%$). It is left as an exercise to the reader to verify that this is true for any two pairs of Pf scores. Interval data ensures that relative distances have meaning on the Risk Cube (though this is generally *not* true for Cf as the vertical dimension).

- **Ratio:** This is where everything falls apart! For data to be ratio, the origin (zero value) must have meaning. The classic example is the various temperature scales. Celsius (centigrade) and Kelvin are equivalent when it comes to Intervals, since their respective degrees represent the same delta in temperature. However, while 0K has meaning in thermodynamic terms – absolute zero, 0 degrees Celsius is arbitrary, chosen to be the freezing point of water.

The Risk Cube is clearly intended to have an origin, the southwest corner of the lower left square (1,1). But the risk scores do not behave as Ratio relative to that origin. A 3 should be thrice a 1, and a 4 should be twice a 2, but when we look at Pf, a 3 (50%) is five times a 1 (10%), and a 4 (70%) is seven thirds of a 2 (30%). The same can be seen to be the case for a Cf scale with even intervals (say, 1 = \$0M-\$2M, 2 = \$2M-4M, etc.)

In fence terms, the fatal flaw is that the ratio scale (0%, 20%, 40%,...) is placed at the fence posts and not the fence rails.² This causes issues with the color coding of the squares, among other things. If this is the case, there seems to be an easy remedy: simply move the fence rails to be centers on the ratio scale. For Pf, scores of 1, 2, 3, and 4, would be aligned with proportional center points of 20%, 40%, 60%, and 80%. So far so good. The corresponding fence posts are 10%, 30%, 50%, 70%, and 90%. Now comes the problem. We need Pf scores of 0 and 5 to align probabilities of 0% and

² In order for the traditional risk cube to be ratio, the “representative point” for each component square needs to be counterintuitively the upper right corner, whereas our eyes are naturally drawn to the center of the square. For cost analysts, this may seem eerily familiar. It is the equivalent shortcoming seen in Cumulative Average (CUMAV) Learning Curve theory, wherein the representative point for a Lot is the last unit number, and not the more intuitive “middle” unit.

100%, respectively, but the corresponding fence rails would extend from -10% to 10% and 90% to 110%. Negative probabilities have no meaning, nor those above one (sports cliches about “giving 110%” notwithstanding).

The Centered Risk Cube looks something like this:

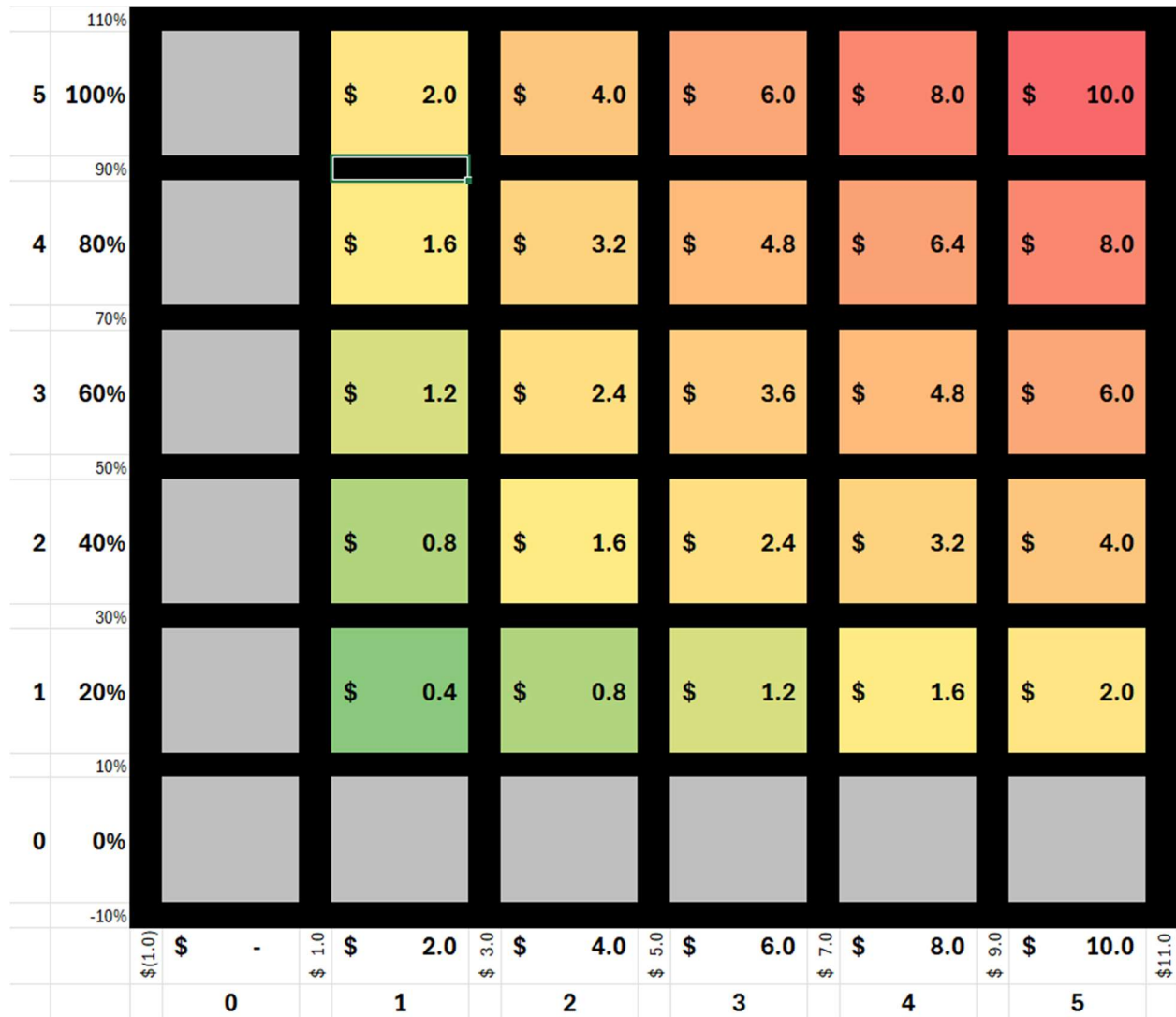


Figure 5: Centered Risk Cube

The ratio scale is easily seen in the constant proportionality of the risk scores, the Pf values (multiples of 20%), and the Cf values (multiples of \$2.0M). (For simplicity, we

chose an arbitrary \$10.0M Cost impact to correspond to a risk score of 5.) The expected values and resultant color codes exhibit symmetry as well.

We will further explore the Centered Risk Cube in the 7.1 Risk Management Visualizations section below.

7.0 Risk Data Visualization

The importance of visual displays of information in effectively conveying qualitative and quantitative data is well established. We will address data visualization in the context of RM and RA.

7.1 Risk Management Visualizations

Our focus remains on the inescapable Risk Cube and how it can be adjusted to be a more faithful representation of the data behind it.

7.1.1 Centered Risk Cube Boundaries

Alas, it seems we are precluded from having our cake (ratio scale) and eating it, too (bins for highly probable and highly improbable Risks). This difficulty can be mitigated by increasing the number of bins, but the bottommost and topmost bins, centered on 0% and 100% on the Pf scale, respectively, will always be problematic. On further reflection, excluding these marginal bins might not be the worst thing. The Risks with near 100% probability become Issues and are shunted into that separate workflow. The Risks with near 0% probability are either negligible or Black Swans that are ground-ruled out by our Framing Assumptions.

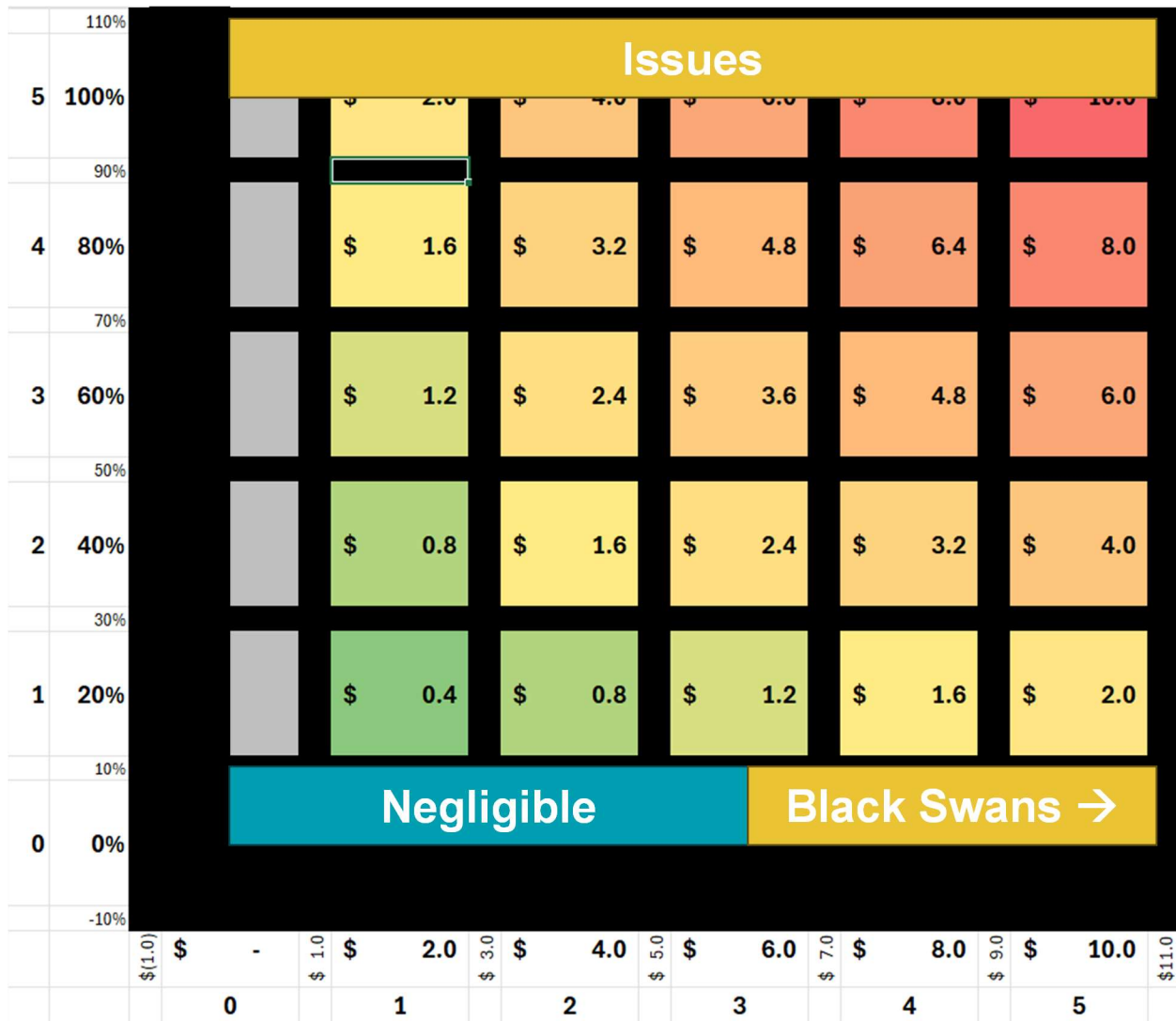


Figure 6: Centered Risk Cube Boundaries

This bottom band brings up another difficulty: the Risk Cube is actually unbounded to the right, since cost and schedule impacts are practically unlimited for the sorts of large, complex projects governments typically undertake. (This is why the upper left square of the traditional 5x5 Risk Cube is often colored green – near certain limited impact – whereas the lower right square is often colored yellow – very unlikely unlimited impact.) One possible remedy is to significantly increase the “maximum” of the Cf scale, but this approach has the disadvantage of “smushing” the majority of risks with much more modest impacts to the far left, where they are harder to see and manage.

Another approach is to let go our of collective fascination with squares – it was never a Risk Cube to begin with! – and extend the right side to create a Risk Rectangle of sorts. While this does not completely address the problem, it at least reminds us of the unboundedness of impact. An example with a doubling of potential impacts (\$0M-\$20M) and their associated Cf risk scores (0-10) is shown in .

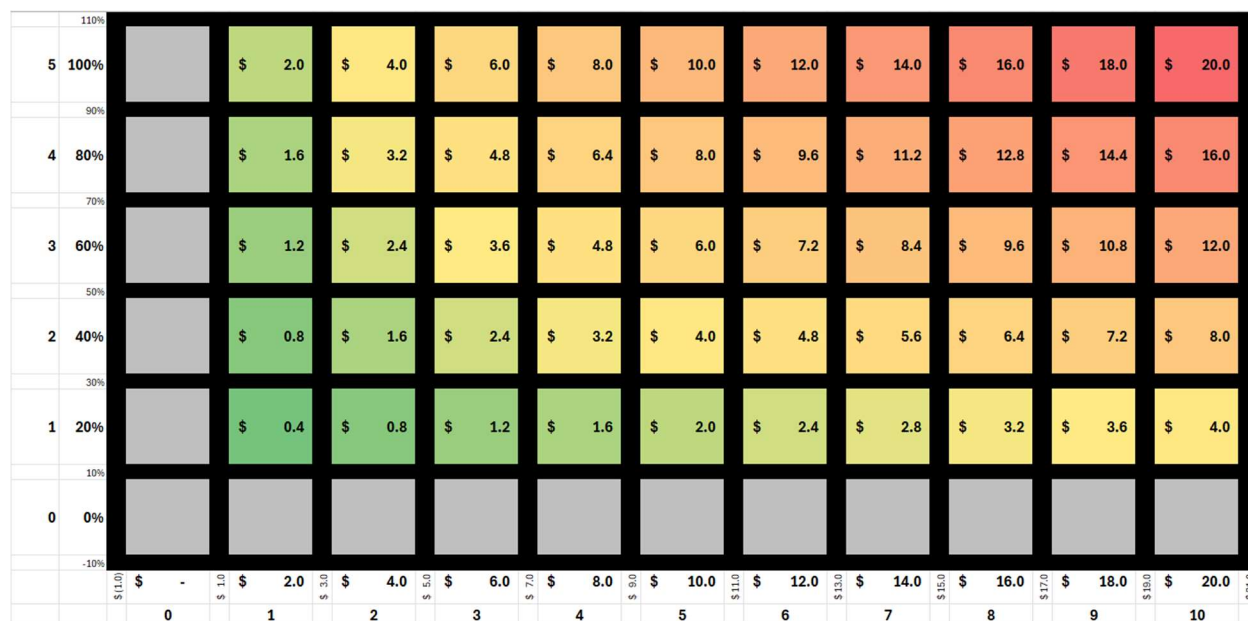


Figure 7: Centered Risk Cube (Extended)

Note that the symmetry is maintained only in the original square. The color coding helps us match squares with identical expected values, which exposes the challenge of trying to manage all risks above a certain threshold. As Pf decreases, Cf must increase proportionally to offset, making the Risk Rectangle – or at least the bottom rows of it – wider and wider. For example, the light orange square with Pf = 80% and Cf = \$10.0M has an expected value of \$8.0M, as does the identically colored square two rows down and five columns over. Pf was halved, to 40%, meaning Cf had to double, to \$20.0M. If Pf were cut in half again, to 20%, Cf would have to double again, to \$40.0M, requiring the bottom colored row to be twice again as wide as shown.

When risk analysts are confronted with these problems of scale, their first instinct is to apply a logarithmic scale, but we worry that that would destroy the visual of the ratio scale we worked to hard to achieve in the first place. Still, that remains an area for

further research. For the moment, we will explore increasing the granularity of the Centered Risk Cube.

7.1.2 Centered Risk Cube Granularity

In RA, it is common to accompany an S-curve depicting a range of possible outcomes for cost or schedule with a table of either deciles (i.e., multiples of 10%) or icosatiles (i.e., multiples of 5%), so named because they divide the 100% probability into 10 or 20 pieces³, respectively. There's no reason that the Risk Cube, representing the Discrete Risks input to the risk model, can't have at least as much granularity. These two versions of the Centered Risk Cube are shown in Figure 8 and Figure 9 below.

³ You guessed it, fence rails! The tables of deciles or icosatiles actually have 11 or 21 values (fence posts), respectively, if we count 0% and 100%, not that we put much stock in those endpoints in the output of a Monte Carlo simulation.

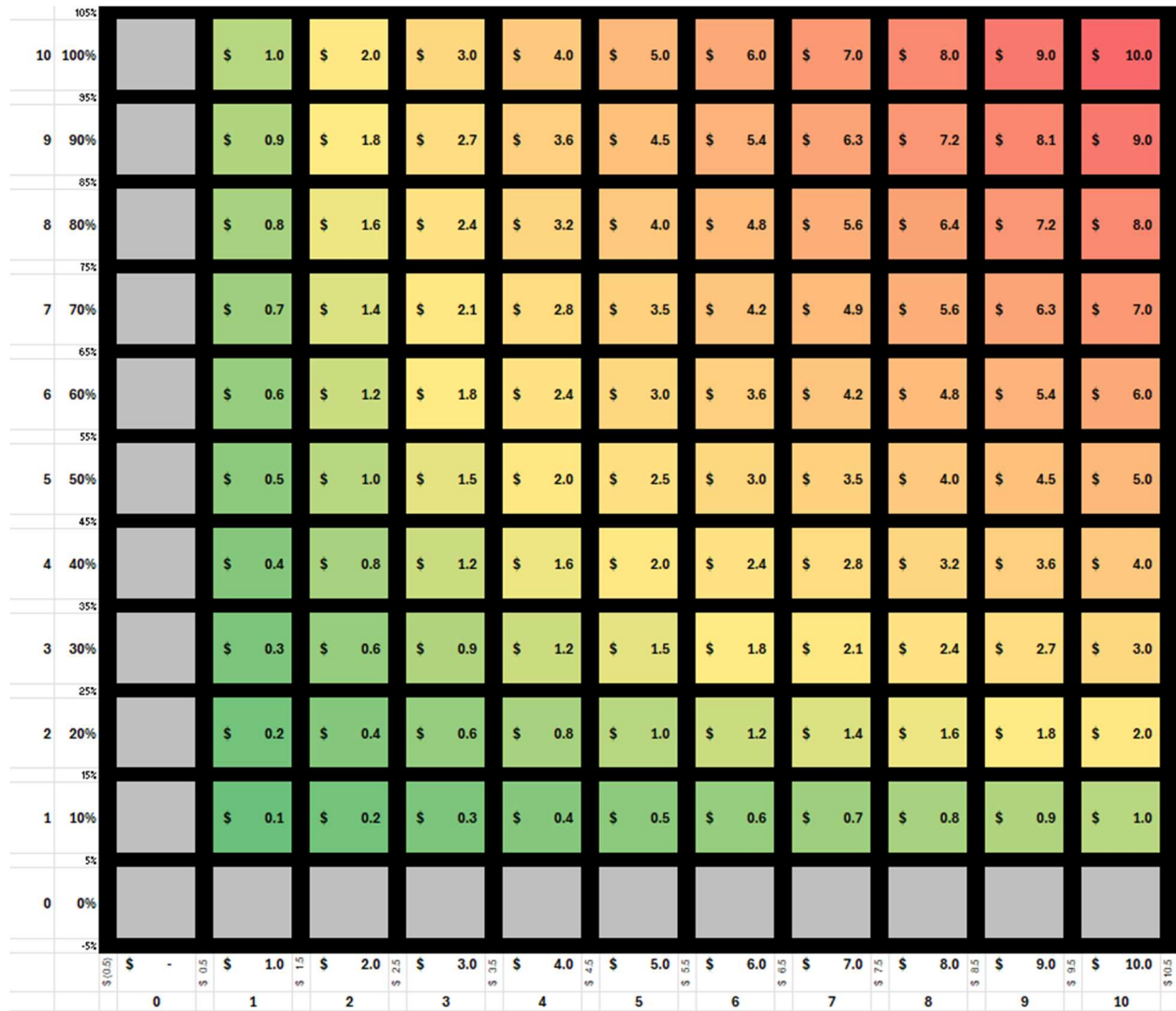


Figure 8: Center Risk Cube (Deciles)

Note that we have reverted to the arbitrary \$10.0M maximum Cf value. For the deciles version, nonzero Pf and Cf scores both range from 1 to 10. By doubling the granularity, we have quadrupled the number of colored squares, from 25 to 100.

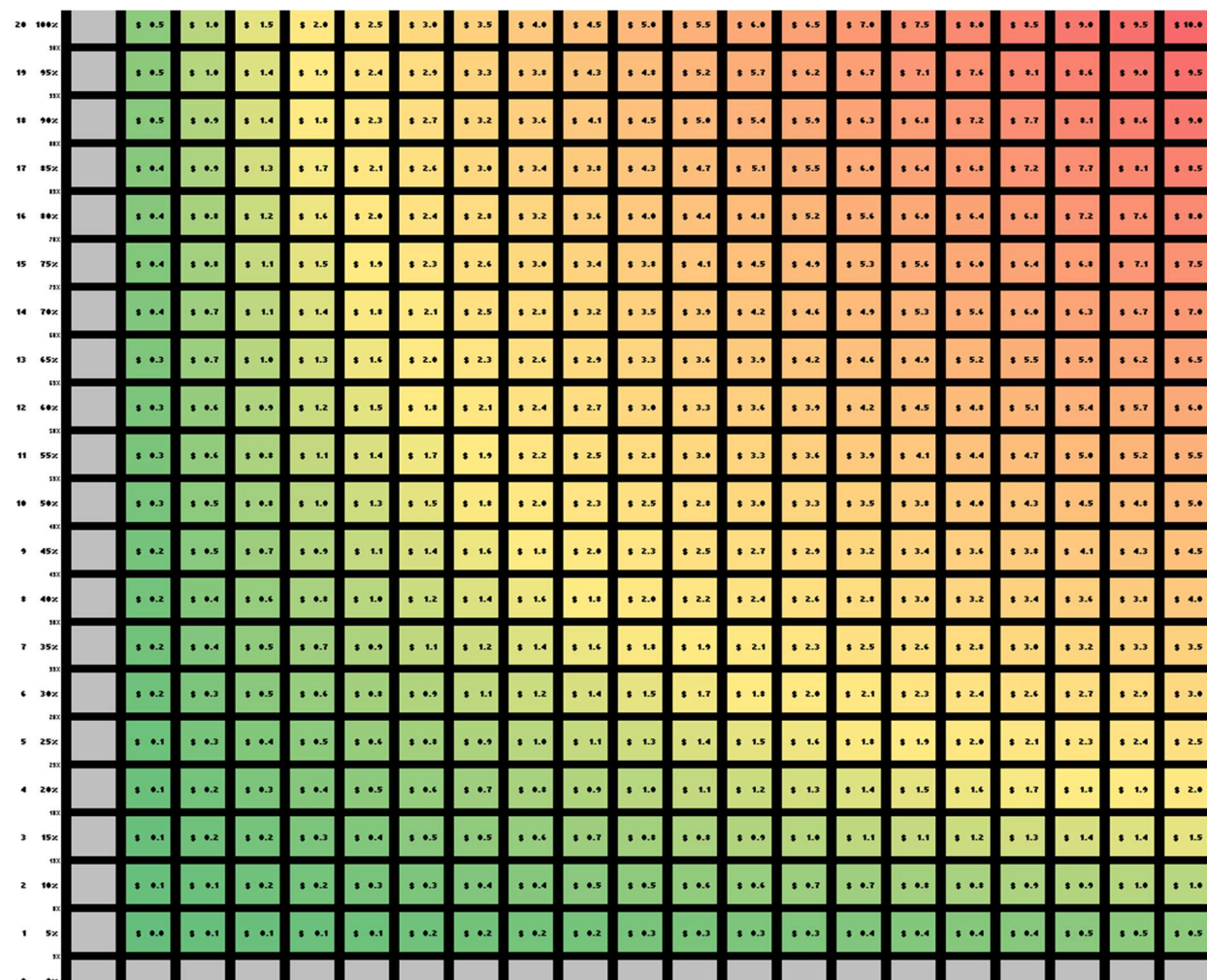


Figure 9: Centered Risk Cube (Icosatiles)

For the icosatiles version, nonzero Pf and Cf scores both range from 1 to 20. We won't make you count them all,⁴ but the number of colored squares has quadrupled again, from 100 to 400.

This may begin to feel like Zeno's Risk Cube (or perhaps Leibniz's Risk Cube). We're on the verge of pixelating, but the point here is not *reductio ad absurdum*. The most important thing to remember is that Pf and Cf should both be assessed on continuous scales first and only then mapped to whatever set of (coherent) bins is being used. The Risk Cube is just a convenient way of visualizing the risk items.

⁴ Insert joke about three kinds of cost estimators here....

Arguably, a Pf of 3 and a Cf of 4, say, on a 1-5 scale is easier to digest as a medium-likelihood risk with an above average impact. Maybe this turns into Pf of 5, Cf of 8 on a 1-10 scale, or Pf of 11 and Cf of 15 on a 1-20 scale, but are the more granular scales really any more complicated? It's not as if we're trying to sketch it on a cocktail napkin. With modern dashboarding tools, there's no reason the dimensions of the Risk Cube even need to be static. Viewers could adjust the "resolution" according to their preference. An additional benefit of going beyond a five-point scale is that we're less tempted to come up with silly adjectives to describe the bins. (Try coming up with 20 different English terms that can be reliability sorted in order of likelihood! "Moderately commonplace"? "Somewhat unlikely"?)

7.1.3 Centered Risk Cube with Opportunities

We've addressed three of the four sides of the Risk Cube, but not the fourth. First, we noted that the top edge where Pf approaches 100% is where Issues are to be found. Conversely, the bottom edge where Pf approaches 0% is where Negligible risks are to be found, except insofar as we worry about the right edge. That third edge is not really an edge at all, and we need to worry about the unbounded nature of potential cost and schedule impacts. In the Risk Framework, Framing Assumption tend to ground rule out low-probability high-impact risks such as Black Swans and other extreme events.

The fourth edge is the left one, where Cf goes to zero. In fact, the gray squares centered on the zero line (vertical axis) actually "go negative." At first blush, this seems nonsensical, but unlike Pf, a negative Cf simply indicates a reduction in Cost (i.e., overrun) or Schedule (i.e., early delivery). It turns out the left edge is not so much an edge as a spine joining the Risk Cube on the right and the much-neglected Opportunity Cube on the left. Or the body of a butterfly joining two beautiful wings.

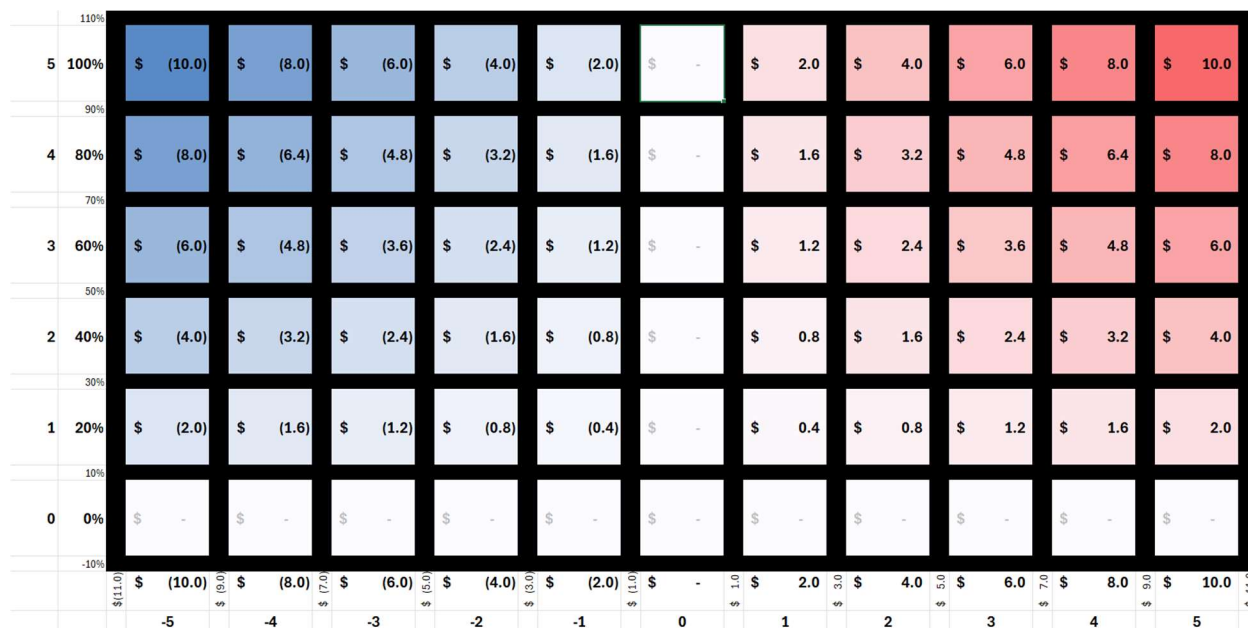


Figure 10: Centered Risk Cube with Opportunities

We’ve shifted to a two-color scale to highlight the symmetry between the two halves. The Cf scale for Opportunities is a mirror image of that for Risks, and we denote as negative one (-1) through negative five (-5) so that the expected value math works out.

The right side of the top edge is for Issues, Risks that have “come true.” There is no commonly accepted designation for the left side of the top edge, Opportunities that have come true. We propose *Windfalls*.

7.1.4 Centered Risk Cube with Thresholds

Finally, we circle back to the seemingly irresistible desire to paint the Risk Cube with a limited color palette. If the decision-maker insists on stoplight colors (i.e., red, yellow, green) or a similar three-, four-, or five-color Severity breakout, we should do it consistently based on expected value.

Figure 11 below shows our same baseline Centered Risk Cube with two distinct color schemes. The version on the left uses Red (High Risk) for any square with an expected value greater than \$5.0M and Yellow (Medium Risk) for any square with an expected value greater than \$2.0M. The version on the right lowers these thresholds to \$4.7M and \$1.8M, respectively. This results in the two previously yellow squares at (3,4) and (4,3) turning red, and the two previously green squares at (1,5) and (5,1) turning yellow.

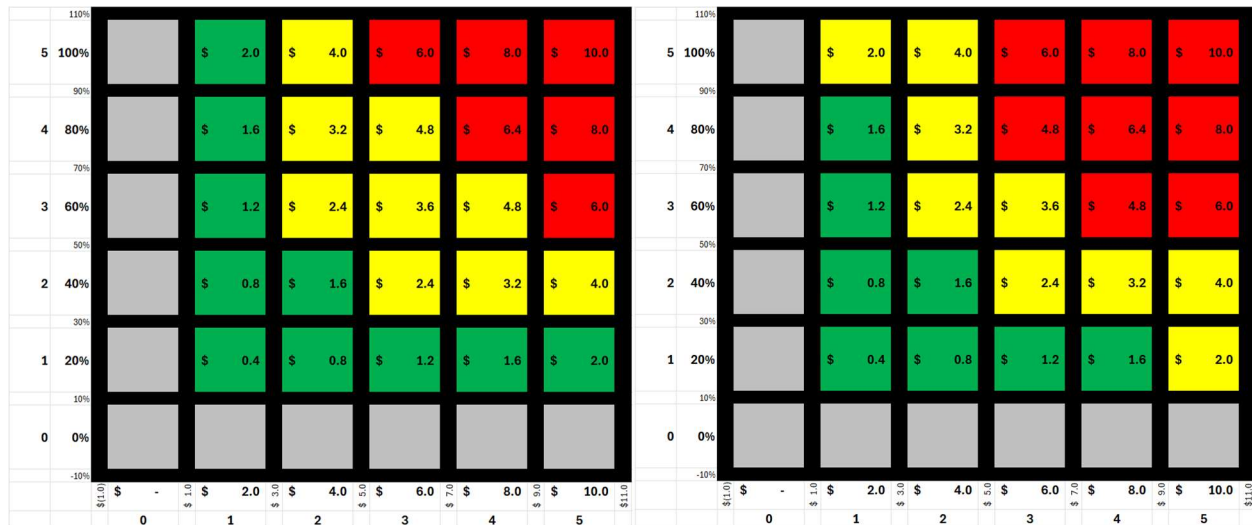


Figure 11: Centered Risk Cube with Thresholds

7.2 Risk Analysis Visualizations

The S-curve is the touchstone graphic for RA and gets a thorough treatment in other papers. The NCCA S-Curve Tool remains a versatile means of generating, annotating, and comparing S-curve.

As RM and RA become better integrated, producing “live” S-curves may require implementing cost estimates as a data pipeline, another variant of the proposed Continuous Estimating / Continuous Analysis (CE/CA) companion to the Continuous Integration / Continuous Deployment (CI/CD) of DevSecOps.

8.0 Risk Model Exemplar

While a detailed functional model is beyond the scope of this paper, the basic principle is simple and elegant: If you're doing RM alone, or RA alone, you're missing a significant piece of your program risk. Only by doing the two effectively in conjunction will you get the most comprehensive picture of risk.

8.1 Calibrating the Risk Cube

RM with no RA involves relying solely on the Risk Register and Risk Cube for information. In this case, any uncertainties are inappropriately modeled as a discrete risk, or more likely omitted altogether. Increasingly higher stages of consciousness are:

- Ignore the Risk Register information, and stick with the point estimate. The Risk Cube is used for management and mitigation only, and is completely divorced from cost and schedule estimates.
- Calculate total exposure in the Risk Register and use it to risk-adjust point estimates.
 - **Unfactored risk:** The sum of all Impacts across the entire Risk Register. This represents an extreme upper bound, a completely impractical worst-case scenario.
 - **Factored risk:** The sum of pairwise products of Probability and Impact. This is the appropriate expected value to add. It captures the mean impact but not the statistical variation about that mean.
- Conduct a Monte Carlo simulation to conflate all the risks and produce a single distribution to capture the collective contribution of the Risk Register. The mean and standard deviation of this distribution can be used to produce a coefficient of variation (CV), which captures variability in unitless (percentage) terms.

Suppose the point estimate is \$100M and the factored risk in the Risk Register is \$10M, for an expected growth of 10%. If that is in the form of a single massive risk with $P_f = 50\%$ and $C_f = \$20M$, the mean value of \$110M will actually never be realized. The “coin flip” produces the point estimate of \$100M the lucky half of the time, and the estimate with full growth of \$120M the unlucky half of the time. The so-called S-curve, or

cumulative distribution function (CDF), will simply look like one big stair step, connecting those two values at the 50th percentile. If instead, the risks were divided up into five risks of (Pf, Cf) = (50%, \$4M), or ten risks of (50%, \$2M), the stair step would start to smooth itself out, as these are binomial distributions of more and more trials, which approach Normality in the limit.

Suppose the conflated risk has both a mean and a standard deviation of \$10M. This is a CV of 100% when viewed as just the risk distribution, or a more reasonable value of 10.0% in relation to the point estimate, or 9.1% in relation to the risk-adjusted estimate. If historical benchmarks show that risk and uncertainty are understated, one approach is to simply double all the component risks, say. This would result in a mean and standard deviation of \$20M, resulting in a CV of 20.0% relative to the point estimate, or 16.7% relative to the risk-adjusted estimate. There are two concerns with this approach. One is that the shape of the S-curve will be wrong, since we're using the "pumped up" discrete risks as a proxy for the missing risk. This could lead us to draw unsound conclusions for percentiles of interest. The second concern is that the source of risk is misattributed. We are missing the root cause for the risk and uncertainty for which we substituted.

8.2 Enhanced Scenario-Based Method

RA with no RM may take several forms, including the enhanced Scenario-Based Method (eSBM). Conversely, risk modeling relies on uncertainties – which may be symmetric or have a net nonzero risk or opportunity – and either omits discrete risks or hopes they're implicitly included like the old Prego commercial: "It's in there!" By modeling the continuous distributions associated with these uncertainties, the analyst produces a smooth S-curve, but risk and uncertainty may still be understated due to the set of missing risks. eSBM focuses on combinations of parameters tied to optimistic or pessimistic scenarios, but they are not tied to specific discrete risk items, nor related mitigation actions.

8.3 Implementing the Risk Framework

RM and RA working in conjunction promise to produce the most robust set of Discrete Risks and Uncertainty per the Risk Framework. This gives us the best chance of

generating cost growth and uncertainty in line with historical benchmarks. The Naval Center for Cost Analysis (NCCA) S-Curve Tool contains benchmarks for a wide range of commodities, services, and eras. See the Risk Benchmarks section under Next Steps.

Even with a comprehensive combination of Discrete Risks and Uncertainties, our RA results may fall short of historical benchmarks. This could be attributed to unknown unknowns, and the difference could be backed out and retained as a separate contingency. (The mathematics involved is beyond the scope of this paper.) This might be held *above* the program level, depending on the Framing Assumptions portion of the Risk Framework.

8.4 Risk Model Case Study

Bringing analysis and management together in a typical phase-gate project management lifecycle increases the potential for more predictable outcomes. Program management organizations that endeavor to use their risk register and risk cube in their development of early-phase program estimates and the associated range of potential outcomes, often develop a risk adjusted range that is overly optimistic and fails to capture the likely total costs of their project.

Cost estimates that only examine historic project results at the aggregate level miss an opportunity to explore threats and opportunities that need to be identified, monitored and mitigated throughout the project lifecycle.

Often, project organizations are decentralized or limited on cross-functional talent. These project organizations typically excel in particular phases of lifecycle RM and RA. For example, an organization or team that is heavy on schedulers will typically do an excellent job in schedule development and management but may not spend the required time on developing a defensible BOE to support many elements of their cost estimate. A program team with a robust risk management organization may develop a far superior risk register but fail to include uncertainty ranges or fully establish a healthy schedule.

When using more discrete RM practices to support cost estimates, it is absolutely critical that all aspects ICSRA are included or it is nearly impossible to capture the full range of potential outcomes of a project. Using RA techniques and including RA practitioners in the estimating process increases the project team's potential of properly calibrating their RM tools and inputs so that their point estimate, and program contingency properly reflect the full range of program outcomes.

To support cross-functional teams in their efforts to implement the Risk Framework, we endeavor to provide detailed practical training. While outside the scope of this paper, a case study based on the Bradley Fighting Vehicle (BFV) has been approved for use as an open-source model, and the authors are developing training materials on how to successfully combine RM and RA approaches.

9.0 RM and RA: The Best of Both Worlds

This paper recommends a powerful combination of RM and RA approaches, rationalized and more effective when applied in conjunction than separately. While the authors come mostly from the RA perspective, the intent is the proverbial two-way street. We've addressed how RM can better inform RA, and how RA can better inform RM.

This approach builds on prevalent practices in RM and RA, and other than studiously avoiding a number of common pitfalls of RM, it is recommended more in the spirit of continuous improvement than remediation. There is no need to abandon the five-step RM process, for example. The application of RA best practice in quantifying and displaying risks will lead to more accurate estimates and more effective triage and mitigation of identified risks.

For program managers and risk managers alike, implementation boils down to the three steps summarized in the following sections.

9.2 Adopt the Risk Framework

The Risk Framework provides not only a clear delineation of Framing Assumptions, Discrete Risks, and Uncertainties, but also handy starter lists for each of those three

groupings, which you can further tailor to your environment (commodity, acquisition pathway, organization, etc.). Use these lists as the foundation of a Track Record, wherein you iteratively collect data on the accuracy of your assessments and use them to improve future risk identification and analysis.

The Risk Framework improves RM by focusing the five-step process on the Discrete Risks where it is most effective. Framing Assumptions are (for the moment) outside the purview of the PM. Uncertainties should be quantified in the Risk Model and monitored through dashboards and other mechanisms within companion processes such as EVM, Manufacturing, and Agile Software Development. Piloting of tripwire risks within the Risk Register and Risk Cube would better inform whether this approach is fruitful.

The Risk Framework improves RA by guiding the deconfliction of discrete and continuous risks, both of which are quantified in the Living Risk Model. The intent is that these two groups together are mutually exclusive and exhaustive, painting a complete risk picture of the program. The cost analysis mantra of “no gaps or overlaps” is no less important in risk, though arguably more challenging to implement. The Risk Framework seeks to address that challenge. Risk assessment and analysis should be more accurate for discrete and continuous risks alike, because we aren’t trying to fit a square peg in a round hole.

9.3 Adopt the Centered Risk Cube

The Risk Cube currently in wide practice misrepresents the risks plotted on it via the illusion of a ratio scale. Pf and Cf scales have wide variation, and long-recommended calibration to historical experience is thwarted by the fact that the 5 x 5 layout of the component squares make it *impossible* to align to a ratio scale. Inconsistent color-coding exacerbates the problem.

The Centered Risk Cube is aligned to ratio scales for both Pf and Cf and comes in quintile, decile, and icosatile varieties, mapping to 0-5, 0-10, and 0-20 ordinal scores, respectively. It provides for perfectly consistent and symmetric color-coding of Severity (i.e., Expected Value), which in turn leads to better triage of discrete risks. It better highlights Issues (Risks as Pf approaches 100%); dovetails with Opportunities (via a

zero-impact spine connecting the two mirror images); and enables the classification of low-probability items into negligible Risks and Opportunities or Black Swans.

The Centered Risk Cube improves RM by accurately depicting RIOs on a unifying graphic with minimal (but crucial) changes to what most RM practitioners are used to. By aligning with the ratio scale, it correctly puts risks in order by expected value, which should lead to more attention being paid to the right risks and a better return on investment for mitigation efforts. In parallel, the Risk community can explore alternative visualization approaches that address the unbounded nature of impacts, for example.

The Centered Risk Cube improves RA by ensuring that the Pf and Cf quantification of RIOs is as accurate as possible. By insisting that risks are quantified first and *then* mapped to the Risk Cube, instead of vice versa, the joint process will provide better inputs to be conflated in the Living Risk Model (see next section). This avoids any guesswork or approximation on the part of the risk analyst, helps ensure a comprehensive set of inputs to the risk model, and avoids the garbage in, garbage out (GIGO) problem.

9.4 Maintain a Living Risk Model

Most RM implementations never even bother to sum up the expected values of their RIOs and ask how that total compares to expected cost growth on the program. For those that do, discrete risks are misquantified (due to pitfalls noted above) and uncertainties are either ignored or mischaracterized as discrete risks. If the total is then used to determine management reserve (MR), the program's risk position is not what the PM thinks it is. Furthermore, while budgeting to a risk-adjusted mean is a best practice – and infinitely better than ignoring risk altogether – the expected value is only the first moment of the distribution and does not provide any indication of the range of possible outcomes. This is where a Risk Model comes in.

The Living Risk Model improves RM by replacing the junk math to which Risk Registers so often fall prey with a means of continually updating the cost estimate. This applies RA best practices, usually Monte Carlo simulation to conflate both discrete and continuous distributions with an appropriate application of correlation. This can be done

for the whole program or a subset of the program, such as a single project or contract. The results are not just a smattering of points on a Risk Cube, but summary statistics such as standard deviation, coefficient of variation (CV), and percentiles of interest, graphically displayed via an easy-to-read S-curve. This directly informs the establishment and use of MR by the PM.

Note that we focus on Cost, but the Living Risk Model may incorporate Cost, Schedule, or both. In fact, a reliable Risk Model requires consideration of all of these. Put another way, you can't have a good Risk Model without a good Cost (and Schedule) Model. See the ICSRA discussion under Next Steps below.

The Living Risk Model improves RA by keeping input risks to that model, especially the discrete risks that "live" in the Risk Register, continually up to date. No longer is the risk model a dusty artifact that gets put on a shelf and trotted out when it's time for an Acquisition Milestone or an ICE. It directly benefits from the cadence of (customarily monthly) reviews as part of the RM, EVM, Business Management, and/or Agile Software processes. Risk analysts can benefit from direct access to Subject Matter Experts (SMEs), as they can both identify relevant data sources and apply Expert Judgment in leveraging those sources and making direct risk assessment, where necessary. With a more robust Living Risk Model, the risk analyst can apply historical benchmarks and back into an appropriate wedge for unknowns unknowns, capturing by neither the Discrete Risks nor the Uncertainties in the Risk Framework. (See the Risk Benchmarks section under Next Steps below.)

Next Steps

This paper has focused on remedying the deficiencies of the Risk Cube to make it a more useful tool in service of both RM and RA. In addition to the development of a more detailed and nuanced (open source) example, the following areas would benefit from further study.

ICSRA

On the RA side, we have focused on what J-CRSUH calls Cost Risk Uncertainty Analysis (CRUA), as distinct from Schedule Risk Analysis (SRA) and Integrated Cost Schedule Risk Analysis (ICSRA). While CRUA and SRA can yield benefits when practiced independently, ICSRA can be quite powerful when a resource-loaded integrated master schedule (IMS) is available. In the context of the Risk Cube discussion, ICSRA holds promise as a remedy for the junk math of trying to combine Cost and Schedule Impact scores. Risk parameters must still be carefully determined (and in a data-driven manner, where possible), but the identification and modeling of time-dependent and time-independent cost will provide greater fidelity in estimating cost and schedule impacts simultaneously. As described, for schedule-driven items in the Risk Register, a companion cost impact can be determined via the Cost Informed by Schedule Method (CISM) by multiplying a projected schedule slip by the burn rate of the “standing army” expected to be continuing to work during that delay.

Portfolio Management

While “Semi-infinite Risk Rectangle” doesn’t have the same ring to it as “Risk Cube,” the unbounded nature of unfavorable cost and schedule impacts remains problematic. In practical terms, we use Framing Assumptions to ground rule out Black Swans and other extreme events. In essence, we’re not asking the PM to manage beyond the 80th or 90th percentile, say. However, appropriate modeling of tail behavior at the portfolio level is both difficult and crucial to the success of managing that portfolio.

For a treatment of these issues, see Christian Smart’s excellent book *Solving for Project Risk Management*. (Smart, 2021)

Risk Benchmarks

The calibration and benchmarking of a myriad of risk methods, from the Risk Cube to eSBM, rely on analysis of historical data across a wide range of programs. The Selected Acquisition Reports (SARs) submitted annually to Congress up through 2019 cover all Major Defense Acquisition Programs (MDAPs), making its standard format ideal for deriving cost growth and uncertainty benchmarks at the commodity (e.g., Ships and Submarines), MILDEP (e.g., DON), and Department-wide (i.e., DoD) levels. Unfortunately, the analysis behind these benchmarks available via the Naval Center for Cost Analysis (NCCA) S-Curve Tool has not been updated in about a dozen years. (Garvey, Braxton, Flynn, & Lee, 2012)

On the bright side, the Cost and Schedule Variance datasets from the SAR are now available as exportable flat files from the SAR application within the Cost Assessment Data Enterprise (CADE) portal (<https://cade.osd.mil/>). As part of the updated analysis, we would want to incorporate the emerging Modernized SAR (MSAR) format covering 2021-2024. It would also be a good opportunity to revisit alternative methodologies for factor derivation, as presented in the paper “Perils of Portability.”

Track Record

While SAR risk benchmarks constitute a track record of sorts at the program level – how did we do against baseline cost and schedule estimates? – we also need a track record of our lower-level risk assessment that feed into tools like the Risk Cube. For both expert- and data-driven Pf and Cf estimates, how did we do? We have month upon month, year upon year of risk dashboard data, but we generally have not done a good *postmortem* on it. This is easier done on the Impact side, where it should be possible to tease out Material and Labor costs and Schedule delays attributable to a Risk that became on Issue. On the Probability side, it’s much more sticky, and like the program track record, we need to view things *on average*. If you think a coin is fair, and you flip it and get tails, that proves nothing. But if you flip it 10 times, and it comes up tails every time, you have pretty strong evidence that heads or tails was not a 50/50 proposition! (The sign test or the chi square test for distribution will have you reject the null hypothesis at a very stringent *alpha*.)

Statisticians may turn their nose up at anecdotal information, but retrospective Expert Judgment is still useful. By borrowing from root cause analysis (RCA), we should be able to create a positive feedback loop and increase the fidelity of our risk assessments over time.

Bayesian Tripwires

We have provided a thumbnail sketch of a have-our-cake-and-eat-it-too approach to including an Uncertainty in the Risk Register by converting its tail probability to a discrete risk using a tripwire value (the 80th percentile, say). For many uncertainties, such as Sizing or Productivity, the realization of the conceptual distribution doesn't happen in one fell swoop, but gradually over time, week by week, sprint by sprint, month by month. (In fact, it is the sum or product of many small changes that gives us a probabilistic basis for a Normal or Lognormal distribution, respectively, in the first place!) With this viewpoint, the tripwire risks in the Risk Register should not be a "fire and forget" but should be periodically adjusted based on returns to date.

Here we can borrow from both EVM, where EACs are continually updated, and from Bayesian statistics, wherein our prior distribution is informed by the accumulation of data. Providentially, there is a paper at this very conference on a Bayesian approach to EVM EACs! (Cantor & Smart, 2024) We are looking forward to assessing whether that approach could be more broadly application to RM.

Epilogue

Ever since “Wiggly Sticks,” it has been posited that every good ICEAA paper needs a catchy title (and a more to-the-point subtitle). This particular paper generated an unusually large number of possible title/subtitle combinations, and they were too good not to share. (Shout out to “Weird Al” Yankovic and Linda Richman.)

- “Everything You Know Is Wrong: How Current Risk Management Practices Mislead (And What to Do About It)”
- “The Risk Management Cube is Neither Risk, Nor Management, Nor a Cube ... Discuss!”
- “The (Cube) Root of All Evil: How 25 Little Squares Ruined Risk (And How to Fix It)”
- “Beyond Calibration: How the Risk Cube Can Better Serve Risk Management and Analysis”
- “Risk Management vs. Risk Analysis: Bridging the Great Divide”
- “Blame It On RIO: How to See Beyond the Risk Cube”
- “Fence Posts and Fence Rails: Why Your Risk Cube Should Be Ratio, Not Ordinal”
- “Getting the Risk Crew Together: Risk, Opportunity, Windfall, Issues, and Negligible Gains (ROWING) Management”

Bibliography

- Cantor, M., & Smart, D. C. (2024). The Best of Both Worlds: Combining Parametric Cost Risk Analysis with Earned Value Management Using Bayesian Parameter Learning. *DoDCAS Proceedings*. Alexandria, VA: OSD CAPE.
- Coleman, R. L., Summerville, J. R., & Dameron, M. E. (2006). Making Risk Management Tools More Credible: Calibrating the Risk Cube. Washington, DC: Society of Cost Estimating and Analysis (SCEA).
- Cost Estimating Body of Knowledge (CEBoK)*. (2013). Annandale, VA: ICEAA.
- Garvey, P. R., Braxton, P., Flynn, B., & Lee, R. (2012). Enhanced Scenario-Based Method for Cost Risk Analysis: Theory, Application, and Implementation. *Journal of Cost Analysis and Parametrics*, 98-142.
- Nehring, R. K., Mann, K. A., & Jones, R. (2012). Galaxy Charts: The 1,000-Light-Year View of the Data. *SCEA/ISPA National Conference Proceedings*. Orlando, FL: ICEAA.
- Office of the Executive Director for Systems Engineering and Architecture. (2023). *Department of Defense Risk, Issue, and Opportunity (RIO) Management Guide for Defense Acquisition Programs*. Washington, D.C.: Office of the Under Secretary of Defense for Research and Engineering.
- Smart, C. B. (2021). *Solving for Risk Management: Understanding the Critical Role of Uncertainty in Project Management*. New York, NY: McGraw Hill.

Appendix A: Acronym List

abbr	expansion
APB	Acquisition Program Baseline
BOE	Basis Of Estimate
CAD	Cost Analysis Division
CAM	Control Account Manager
CapAx	Capital Acquisition
CBP	Customs and Border Protection
CDF	Cumulative Distribution Function
CFO	Chief Financial Officer
CISA	Cybersecurity and Infrastructure Security Agency
CL	Confidence Level
CRUA	Cost Risk Uncertainty Analysis
CV	Coefficient of Variation
DE	Digital Engineering
DHS	Department of Homeland Security
DoD	Department of Defense
DoE	Department of Energy
EAC	Estimate At Completion
ETC	Estimate To Completion
EVM	Earned Value Management
FEMA	Federal Emergency Management Agency
FMPM	Financial Management Policy Manual
IC	Intelligence Community
ICE	Immigration and Customs Enforcement
ICSRA	Integrated Cost and Schedule Risk Analysis
IMS	Integrated Master Schedule
IT	Information Technology
LPTA	Lowest Price Technically Acceptable
MDA	Missile Defense Agency
MDAL	Master Data Assumptions List
MDAP	Major Defense Acquisition Program
NASA	National Aeronautics and Space Administration
NCCA	Naval Center for Cost Analysis
NGA	National Geospatial-Intelligence Agency
NNSA	National Nuclear Security Administration
NRO	National Reconnaissance Office
ODNI	Office of the Director of National Intelligence
PARM	Program Accountability and Risk Management
PM	Program Manager
RA	Risk Analysis
RIO	Risk, Issue, Opportunity

RM	Risk Management
RMP	Risk Management Plan
TCPED	Tasking, Collection, Processing, Exploitation, and Dissemination
TPC	Total Project Cost
TSA	Transportation Security Agency
USCG	United States Coast Guard
USCIS	United States Citizenship and Immigration Services
WBS	Work Breakdown Structure

Appendix B: Agency State of Practice Details

B.1 Department of Energy

The National Nuclear Security Administration (NNSA) is the semi-autonomous agency within the Department of Energy (DOE) responsible for the nation's nuclear arsenal. It has two major categories of projects: Infrastructure and Facilities, known collectively as Capital Acquisition (CapAx); and Weapon Systems. Both categories of projects rely more heavily on the attributes of RM in their programmatic development of estimates and project execution. The examples and discussion presented within are primarily associated with DOE/NNSA's risk policies and practices used in the pursuit of CapAx projects. Weapons projects are substantially similar in policy and practice but there is currently a significant number of CapAx projects as DOE and NNSA attempt to revitalize severely aging Cold War infrastructure needed to support the safe and secure management of the nuclear stockpile.

DOE and NNSA develop their estimate of Total Project Cost (TPC) for large CapAx projects based on a series of comprehensive guides including but not limited to:

- DOE Order 413.3B Chg 6 (LtdChg), Program and Project Management for the Acquisition of Capital Assets.
- DOE G 413.3-7A Chg 2 (LtdChg), Risk Management Guide.
- NNSA Policy (NAP) 413.6, Confidence Levels and Escalation for Cost Estimating.

The Risk Management Guide is for use by all DOE elements. The Guide intends to provide non-mandatory risk management approaches for implementing the requirements of DOE O 413.3B. The guide includes Attachment 12, the cost and schedule contingency development process. DOE requires that project estimates be developed based on qualitative and quantitative analysis of project risks and other uncertainties. The DOE qualitative and quantitative analysis process begins in the project's planning stage with the identification of project risks during the initial project planning phase. DOE's contingency development process is based on all relevant best practices, prescribes the integration of uncertainty, schedule, correlation, as well as a

comprehensive risk register. Proper application of this process results in a credible estimate where the point estimate and range of outcomes is based a complete analysis of uncertainty and risk.

Even with the proper application of this detailed process, DOE has found that their projects often exceed the upper bound of their early phase estimates. To account for early phase scope and cost growth, DOE has implemented supplemental guidance that prescribes the application of additional cost growth factors for projects in the earlier milestone phases. These cost growth factors or additional range estimates are developed from historic data and are meant to counter early-phase project unknowns and programmatic optimism. DOE's practice of developing their estimating range from a comprehensive risk identification and estimating process is vital to their overall risk management (RM) process but the need for additional contingency factors displays a potential weakness in the process's ability to effectively capture the full range of potential project outcomes.

B.2 Department of Homeland Security

The Department of Homeland Security (DHS) was established in 2002, combining 22 different federal departments and agencies into a unified, integrated Cabinet agency. DHS is composed of multiple operation and support components. Acquisition programs and the primary responsibility for cost estimation, resides in eight (8) acquisition components: United States Citizenship and Immigration Services (USCIS), United States Coast Guard (USCG), Customs and Border Protection (CBP), Cybersecurity and Infrastructure Security Agency (CISA), Federal Emergency Management Agency (FEMA), Immigration and Customs Enforcement (ICE), Secret Service, and Transportation Security Agency (TSA). These acquisition components execute a wide variety of missions, and acquire products and services covering a wide variety of commodity types including IT systems, ships, aircraft, facilities, and security systems. Cost estimating policy and support are managed by the DHS CFO Cost Analysis Division.

DHS cost analysis policy and best practice, including guidance on risk analysis, is found within the Financial Management Policy Manual (FMPM), a document published by the

DHS CFO. There is additional guidance on risk management available in the Risk Management Planning Handbook, a document published by the Program Accountability and Risk Management (PARM) office. This section provides an overview of risk management and risk analysis policy and guidance within DHS.

B.2.1 Risk Management Current State of Practice at DHS

The DHS Risk Management Planning Handbook defines risk management as follows:

Risk management is the method by which a program minimizes its exposure to risks that negatively affect how the program fulfills its defined objectives. The Risk Management Plan (RMP) defines risk processes and tools that facilitate the program's efforts to execute projects, ongoing operations, and support services successfully.

The foundation of risk management at DHS is a process that can be grouped into three phases: risk identification, risk analysis, and risk handling. There is additional guidance on ongoing risk monitoring, control, and communication. Programs are required to document the risk management results in a Risk Management Plan (RMP). The RMP also defines the approach and practices a program will use to help minimize exposure to risks that negatively affect how the program fulfills its defined objectives.

Risk identification guidance starts by providing a list of program artifacts that typically contain relevant information. Risk categories are also listed and defined, which provide a standardized way of grouping similar risks together. For example, cost risks, schedule risks, technical obsolescence, etc. For each identified risk, the RMP documents narrative “if/then” statements, where the “if” portion describes the risk and the conditions and circumstances that define the risk. The “then” portion documents the consequences/impact of occurrence. Risk identification also involves documentation of risk owners, source of risk information, and processes, forums, and frequency of risk assessment, reassessment, and monitoring.

The risk analysis phase begins with an assignment of probability to each risk. A five-point scale is used, based on the following probability ranges:

Table 8: DHS Probability Risk Scale

Probability Description	Probability Range
Very High (Extremely likely)	81% – 99%
High (Probable)	61% – 80%
Medium (Possible)	41% – 60%
Low (Unlikely)	21% – 40%
Very Low (Highly improbable)	1% – 20%

Risks are then reviewed and scored for impact, based on the following guidance:

Table 9: DHS Impact Assessment Scores

Project Objective	Very Low 1	Low 2	Medium 4	High 4	Very High 5
Cost	Insignificant increase	Increase < 2% of budget baseline	Increase 2–5% of budget baseline	Increase 6–10% of budget baseline	Increase > 10% of budget baseline
Schedule	Insignificant slippage	Slippage < 2% of project baseline schedule	Slippage 2–5% of project baseline schedule	Slippage 6–10% of project baseline schedule	Slippage > 10% of project baseline schedule — OR — Slippage past a milestone mandated by Congress
Scope	Scope decrease barely noticeable	Minor areas of scope affected	Major areas of scope affected	Scope reduction unacceptable to sponsor	Project outcome is effectively useless
Quality	Quality degradation barely noticeable	Quality degradation noticeable, but does not fail acceptance criteria	Quality reduction requires sponsor approval	Quality reduction unacceptable to sponsor	Project outcome is effectively useless
Cost	Insignificant increase	Increase < 2% of budget baseline	Increase 2–5% of budget baseline	Increase 6–10% of budget baseline	Increase > 10% of budget baseline

Once probability and impact have been scored, risks score can be calculated by multiplying probability by impact. The resulting score, which can range from 1 to 25, is mapped into a matrix and given a red, yellow, green color code as follows:

B.2.2 Risk Analysis Current State of Practice at DHS

Policy and guidance for risk analysis at DHS resides in the FMPM. This document includes an 8-step process for risk analysis, as well as the governing DHS policy and guidance. The FMPM defines cost risk and uncertainty analysis as follows:

Cost risk and uncertainty analysis is a technique used to quantify the cumulative impact of uncertainty and risk on a cost estimate. It is a critical part of the development of a credible LCCE and accounts for the inherent risks and uncertainties present in all estimates.

Guidance is provided on three types of cost risk analysis: inputs-based, outputs-based, and scenario-based methods. Input-based methods are defined and recommended as the preferred technique. The risk analysis process comprises these steps:

1. Determine Program Cost Drivers and Associated Risks
2. Develop Probability Distributions to Model Uncertainty
3. Account for Correlation among Cost Elements
4. Perform Uncertainty Analysis Using Monte Carlo Simulation
5. Identify Cumulative Probability of Point Estimate
6. Recommend Contingency Reserves to Achieve Acceptable Levels of Confidence
7. Allocate, Phase and Convert a Risk-Adjusted Cost Estimate to Then-Year Dollars
8. Identify High-Risk Elements to Help in Mitigation Efforts

Step 1 includes suggested risk types and categories, in the same way that risk categories are defined in the management process. Although there is some overlap (schedule risk appears in both lists), the list from each guidance document differs.

Step 2 describes the process for identifying probability distributions. In cases where this is based on expert opinion, it is suggested that a triangular distribution is used, and that expert-provided low and high bounds be placed at the 15th and 85th percentiles.

Step 3 addresses correlation among cost elements. Both functional and injected correlation are addressed. For injected correlation, in situations where precise correlations are not known, guidance is based on the Air Force Cost Risk and Uncertainty Handbook, which recommends minimum correlation values of 0.50 for estimates with five or less WBS elements, 0.25 (10 elements) or 0.10 (20+ elements).

Step 4 describes the process for using Monte Carlo simulation. Suggested tools for this step include ACEIT, Crystal Ball, and @Risk. Guidance states that results should be displayed using an S-curve. The S-curve should also identify the value of the point estimate, confidence level (CL) of the point estimate, value of the 50% CL, and the CV of the entire S-curve. Specific guidance on the CV metric states:

Higher CVs represent higher cost risk. As a general rule of thumb, CVs are typically 35-45% for high-risk programs, 25-35% for medium risk programs, and 10-20% for low-risk programs. Unusually low CVs (i.e., less than 10%) should alert the analyst to reassess the analysis, data inputs, sources, and GR&A for inconsistencies, omissions, and/or errors. Low CVs can also be a result of an inappropriate lack of correlation in the risk model.

Step 5 describes the process for calculation and usage of the cumulative probability of the point estimate.

Step 6 describes the usage of risk analysis to calculate and recommend contingency reserves. This process starts with the calculation of a risk-adjusted life cycle cost estimate. This number can be calculated for any CL on the risk curve. However, the 50% CL value plays a significant role in DHS acquisition guidance:

- All DHS programs are categorized by acquisition level, with levels 1-2 being the largest programs that are subject to the most visibility and oversight. Level 3 programs are generally managed at the component level. The acquisition level is

determined by threshold values, based on the 50% CL value from the risk analysis.

- All DHS acquisition programs must publish an Acquisition Program Baseline (APB), which contains values for cost objective and cost threshold. These values determine whether a program is in cost breach. The objective value is calculated using the 50% CL value for investment costs. A sustainment cost threshold also uses the 50% CL value, calculated on a per-year basis. Threshold values are calculated relative to the objective values, and must be less than 15% above the objective and less than the 99% CL value.
- Program submitted budgets must be calculated using risk-adjusted values. Guidance states that the 50% CL values should be used, unless DHS CFO approval is obtained for a different value.
- Affordability analysis, which is also a requirement for DHS acquisition programs, compares available funding against the 50% CL value to determine program affordability.
- Calculation of contingency reserve / management reserve is based on the difference between the point estimate and the 50% CL value.

Step 7 describes the process for allocation of risk-adjusted results to WBS elements, and phasing of these results over each Fiscal Year.

Step 8 describes how risk analysis results should inform risk management best practices:

The final step in cost risk and uncertainty analysis bridges the gap between cost estimating and risk management. Results from the cost risk and uncertainty analysis should be used to prioritize risk mitigation focus, based on probability and impact to the program. The program manager should analyze elements that pose most risk and identify requisite risk mitigation strategies. Finally, once a risk mitigation plan is developed and in place, risks should be tracked throughout the life of the program.

Given the subject of this paper, this guidance is particularly noteworthy. It acknowledges the fact that risk analysis is different from risk management, and that it is desirable for one to inform the other. **It is also noteworthy that the parallel guidance in the Risk Management Planning Handbook contains no such guidance on how to bridge the gap from the other direction.** This paper offers practical advice on this subject.

B.3 Intelligence Community

The Intelligence Community (IC) includes satellites procured by the National Reconnaissance Office (NRO) and corresponding ground systems to facilitate the end-to-end Tasking, Collection, Processing, Exploitation, and Dissemination (TCPED) process.

NRO's Instruction 130-1-4, Risk, Issue, and Opportunity Management (dtd 14 February 2020) falls under Business Function 130, Systems Engineering and Systems Engineering Management and Directive 130-1, Systems Engineering Management. The process is referred to as RIOM, and different handling options are broken out by item type. *Accept* applies to Risks and Issues. *Watch* applies to Risks and Opportunities. *Transfer/Share* applies to all three. The analogue for *Mitigate* for Risks is *Resolve* for Issues and *Exploit/Invest* for Opportunities. The steps of the five-step process are somewhat different:

- Identify, Define and Validate R/I/O
- Assess and Validate Impacts
- Develop and Validate Plan, and Prioritize R/I/Os
- Execute and Assess Plan Execution
- Close and Identify Lessons Learned

Lessons Learned shunt off from Step 5 to an unnumbered Continuous Process Improvement spur.

There is a thorough process for tracking mitigation actions. Each item has an *Inherent Assessment* (i.e., pre-mitigation) for Probability and Impact. The *Current Response* field documents any ongoing mitigation actions, with the *Residual Assessment* capturing reduced Pf or Cf as a result. The *Proposed Response* field can be used to show a subsequent round of mitigation actions. There is an extensive process flow for identifying, vetting, and approving RIOs.

The National Geospatial-Intelligence Agency (NGA) RM process is governed by the Information Technology (IT) Risk Management Plan (RMP) Revision A (dtd 11 May 2022). It defines IT RM at the Enterprise Level, together with the flow down to the

Programs, Projects, and Segments, and the flow up to the Agency and Inter-Agency level (including Mission Partners). The usual five steps are augmented with Reporting and Closure. The plan defines the decision boards responsible for the configuration control of RIOs and touches briefly on the associated DE tools. It includes a Sample Risk Matrix with typical R/Y/G shading but no definitions and defers to the IT RM Procedure (dtd 31 January 2023) for more detail.

The Procedure provides ranges for Pf and Cf. For Probability, scores of 1, 2, and 3 represent quartiles (1-25%, 26-50%, and 50-75%), but the fourth quartile is split between 4 (75-90%) and 5 (90-99%). This piecewise linearity might make more sense on the Cf side, as a nod to the unbounded nature of impacts, except that basis to which the relative values (percentages) should be applied is not specified. A third quantity called Risk Priority is specified with five adjectival bins (Low, Minor, Major, Critical, and Blocker), but the definitions are ambiguous and reference Pf more than Cf. There is minimal information on the treatment of Issues and Opportunities as distinct from Risks. RIO item data and metadata fields are defined in an appendix.

On the RA side, the Office of the Director of National Intelligence (ODNI) Cost Risk Analysis Handbook (Version 1 – Fiscal Year 2014) provides guidance on approaches ranging from Monte Carlo simulation to Method of Moments to eSBM. It has substantive mathematical detail on some of the thornier constructs in RA, including correlation and the lognormal distribution.

B.4 Government of Canada

Just like its friendly neighbor to the south (or the north, in the case of Detroit and Alaska), the sovereign nation of Canada has a diverse portfolio of projects across both Defence and Civilian Departments. Two case studies shed light on the state of risk analysis on Canadian projects:

First, the National Shipbuilding Strategy (NSS) is an approach to revitalize the nation's shipbuilding industrial base (IB) while providing much-needed ship to both the Royal Canadian Navy (RCN) and the Canada Coast Guard (CCG). Prime contractors submit large, complex proposals made in commercial software, necessitating the use of

innovative techniques to create miniature versions of the proposal for Monte Carlo analysis. Much of the notional example of risk is built on client experience with these proposals.

Second, the Future Aircrew Training Program (FAcT) is a 25-year contract comprising a comprehensive aircrew training program that provides aircraft, simulators, civilian instructors and classroom training systems, as well as other essential services, such as aircraft and airfield maintenance, accommodations, and food services. The program has a lifetime estimated contract value of \$11.5B.

During the recent procurement of these services, Canada required bidders to adopt, document, and present significantly more defensible estimating and risk management practices in their bids. For the price evaluation, Canada required bidders to perform a cost maturity assessment that required significant depth and breadth in their basis of estimate (BOE) development and submission. In more typical fashion, the offerors were giving higher scores in the price evaluation for lower pricing.

To offset the benefit of competitive pricing, Canada also required offerors to perform a comprehensive examination of risk and submit the documentation in a formal risk maturity assessment. In the cost maturity and risk maturity assessment, bidders were required to justify all aspects of their complex BOEs. They were rewarded for being lower on the evaluate price scale; however, they were then penalized if their quantitative risk assessment failed to show a thorough examination of risk and if their ultimate sale price failed to include a reasonable return that had the potential to cover the impact of the identified risks. Based on the solicitation rules, failing to properly identify, estimate the impact of, and include additional risk in the bid would trigger a government adjustment to bid price to reduce the offeror's overall evaluated score.

The novel approach by the Canadian government exposed both strengths and weaknesses in the application of RA and RM. This approach had the potential to incentivize industry to provide a competitive price while still requiring realism based on effective application of risk management and analysis. In practice, the required RM processes were very limited, focused primarily on the identification and inclusion of discrete risks, and did not fully account for the potential uncertainty in the development

of an offeror's final risk-encumbered fee position. The result was a far more realistic acquisition price position than a classic lowest price technically acceptable (LPTA) evaluation but likely still an understatement of total program cost.

B.5 Industry / Shipbuilding

In addition to NSS and the Canadian yards, the U.S. shipbuilding portfolio and industrial base, spanning both USCG and the United States Navy (USN), provide a view of RM and RA typical of capital-intensive design and production programs. Private contractors develop internal S-curves of their proposals to understand the most impactful areas for negotiation and "setting the share-line." These S-curves are typically built up from a myriad of internal data sources, such as parametric tools or detailed engineering buildups, characteristic of the level of detail a private contractor has on its operations. By contrast, the government perspective tends to be significantly more focused on benchmarking against other programs (i.e., how did the last analogous class of ship perform in the build stage, and are there significant design/environment changes since). This bottom-up (shipyard level) versus top-down (government level) clash of approaches is rooted in the incentives for each stakeholder: the government tends to be focused on a long-term portfolio approach and evaluating performance across many programs, while private contractor's quarterly balance sheets necessitate a detailed focus on individual vessel performance.

Placing this duel of incentives in context of the courtship between Risk Analysis and Risk Management, the Risk Register acts as Cupid's arrow. The shipyard is burdened with defending both the impact and probability of the risk's occurrence, as well as discussing cost-effective mitigatory strategies to get at the root cause of risk. The government leverages this analysis (and their own) to make fleet-level decisions on where to invest in the industrial base, provide cost or schedule incentives, or even instigate an engineering change proposal. These trade-off decisions are a direct impact on the military budget (and to a lesser extent the shipyard's balance sheets), highlighting the necessity of a clear and defensible risk register.

B.6 Department of Defense

Within the Department of Defense (DoD), policy is by its nature top-down, but guidance often takes form bottom-up. An example of the former in the realm of RM is the *Risk, Issue, and Opportunity (RIO) Management Guide for Defense Acquisition Programs* (hereinafter, the “RIO Guide”), a recent update to which was published in September, 2023. An example of the latter in the realm of RI is the *Joint Cost Schedule Risk Uncertainty Handbook (J-CSRUH)*, which was developed iteratively by the Service Cost Centers (SCC) and other DoD organizations such as the Missile Defense Agency (MDA).

The seminal 5-step RM process from the RIO Guide is consistent with versions seen elsewhere. Figure 14 replicates Figure 2-1 from the guide.



Figure 14: Risk and Issue Management Process Overview

J-CRSUH was developed collaboratively over a number of years and is endorsed by all of the DoD Service Cost Centers, the National Aeronautics and Space Administration (NASA), and the Missile Defense Agency (MDA).

Appendix C: RM According to ChatGPT

What follows is a very rudimentary view of RM based on a few quick prompts to ChatGPT. It is a reasonable cross-check of typical practice described in this paper, and (as usual) also serves as a cautionary tale about blind use of ChatGPT.

When asked about Risk Cube implementation, ChatGPT produced these responses.

Assigning Likelihood and Consequence ratings for your Risk Cube depends on the context of your risk assessment, such as business, engineering, cybersecurity, or project management. Here's a general approach:

1. Likelihood Rating (Probability of Occurrence)

Likelihood assesses how often a risk might occur. Use a 5-point scale (or adjust as needed):

Rating	Description	Example
1 - Rare	Unlikely to happen; highly improbable	Less than once in 10 years
2 - Unlikely	Could happen, but not common	Once every 5–10 years
3 - Possible	Might occur occasionally	Once every 1–5 years
4 - Likely	Expected to happen regularly	Annually or more frequently
5 - Almost Certain	Will almost certainly happen	Multiple times per year

Figure 15: Likelihood Ratings per ChatGPT

2. Consequence Rating (Impact Severity)

Consequence measures the impact if the risk occurs. Again, a 5-point scale is common:

Rating	Description	Example (Business Context)
1 - Insignificant	Minimal impact, no disruption	Minor financial loss, no reputation damage
2 - Minor	Some impact, but manageable	Small delays, slight cost increase
3 - Moderate	Noticeable impact, needs response	Operational disruption, moderate financial loss
4 - Major	Significant impact	Project failure, major financial or reputational damage
5 - Catastrophic	Severe impact, potential crisis	Business shutdown, legal violations, fatalities

Figure 16: Consequence Ratings per ChatGPT

3. Risk Score & Cube Mapping

Once Likelihood and Consequence are assigned, multiply them to get the **Risk Score** (e.g., $3 \times 4 = 12$).

Then, map it to your **Risk Cube**:

- **Low Risk (1–5)** → Acceptable, minimal action needed
- **Medium Risk (6–10)** → Monitor and manage
- **High Risk (11–15)** → Requires mitigation
- **Critical Risk (16–25)** → Immediate action needed

Would you like a visual example of the Risk Cube? 😊

Figure 17: Conflated Risk Scores per ChatGPT

When asked for a typical Risk Register risk item, ChatGPT produces the information in Figure 18 below. Data include a Risk Description and Pf and Cf ratings, plus the ill-advised product of those two as a calculated field. Metadata include a Risk Owner and Status. A multi-step Mitigation Plan is also included.

This is comparable in content and format to the example in section 3.3 Notional Risk Viewer.

Risk Title: Project Deadline Overrun

Risk Description:

There is a risk that the project may not meet its planned deadline due to unforeseen delays in software development and testing. This could result from resource constraints, technical complexities, or dependencies on third-party vendors.

Likelihood Rating: 3 – Possible

Based on historical data, similar projects have experienced moderate delays due to unforeseen coding challenges and integration issues.

Consequence Rating: 4 – Major

Failure to meet the deadline could lead to significant financial penalties, reputational damage, and loss of client trust. Additionally, project delays may impact downstream deliverables in related initiatives.

Risk Score: $3 \times 4 = 12$ (High Risk)

Mitigation Plan:

1. Implement an **agile sprint review** every two weeks to monitor progress.
2. Assign a **contingency buffer** of two additional weeks in the timeline.
3. Secure additional **developer resources** to avoid bottlenecks.
4. Escalate potential delays in **weekly risk review meetings** for early intervention.

Risk Owner: Project Manager – John Doe

Status: Under Monitoring



Figure 18: Notional Risk per ChatGPT

Appendix D: RM Notional Example Additional Details

The team devised a more useful way of displaying items from the Risk Register workbook submitted by the offeror.

Risk ID	Risk Type	Status	Raised By	Category	Risk Name	Risk Statement					
2	Threat	Accepted	Original Estimating Risk Log	Certification / Standards	Certification	There is a risk that Industry will be instructed to comply with Standards or agreements/understandings that are different from those specified in the Development Contract. Caused by: The government mandating adherence to different, more stringent standards.					
	Prime Score		Integrator Score		Design Agent Score		Support Lead Score		Overall Score		
	Impact	Probability	Impact	Probability	Impact		Probability	Impact	Probability	Impact	Probability
	1	1	2	3	3		5	1	1	3	5
	Prime Mitigated Score		Integrator Mitigated Score		Design Agent Mitigated Score		Support Lead Mitigated Score		Overall Mitigated Score		
	N/A	N/A	N/A	N/A	1		2	N/A	N/A	1	2
	Prime Mitigation Strategy		Integrator Mitigation Strategy		Design Agent Mitigation Strategy		Support Lead Mitigation Strategy		Final Mitigation Strategy		
	N/A		N/A		Ensure contractual terms and scopes of work are aligned to the Development Contract standards with no deviation therefore allowing the design and the build to continue on the same standards as designed. Any new standards introduced would be treated as a contractual change.		N/A		N/A		

Figure 19: Risk Viewer Improvements

The example 'Viewer' shown in Figure 19 above pulls all the information from one line item in the Risk Register table into a more easily digestible format. Using the basic layout from the contractor, we made several improvements. The drop-down under 'Risk ID' was expanded to encompass all Risk Register items. Erroneous look-ups were corrected so that the data appear in the correct fields as labeled. Color codes were added for Impact and Probability. The former uses a temperature scale to allow for fractional values due to averaging, whereas the latter uses a discrete scale with the same 1 through 5 colors as Severity. While Probability ratings are retained as A (least likely) to E (most likely) in the source table, they are mapped to the more intuitive and consistent numerical 1 through 5 scale here.

The Viewer uses grayed-out “N/A” values to clearly distinguish nulls from zeroes. In the example shown, only one out of the four contractors provided Mitigation Scores and a Mitigation Strategy for Risk ID 2. Finally, some cells were resized to make the table narrower and taller for viewing on a single large screen monitor with no scrolling.

Section 6.2 Severity and Junk Math demonstrated how the five Severity ratings of the Risk Cube in the Notional Example were interleaved when the Cost Impacts were used to determine the expected value of each square. A similar picture emerges when the Schedule Impacts are used instead.

Table 10: Severity by Expected Schedule Impact

1A	1	A	0.10	2	0.20	1	Low
1B	1	B	0.30	2	0.60	2	Low
2A	2	A	0.10	8	0.80	3	Low
1C	1	C	0.50	2	1.00	4	Low
1D	1	D	0.70	2	1.40	5	Moderate
1E	1	E	0.90	2	1.80	6	Moderate
3A	3	A	0.10	19	1.90	7	Low
2B	2	B	0.30	8	2.40	8	Low
4A	4	A	0.10	39	3.90	9	Moderate
2C	2	C	0.50	8	4.00	10	Moderate
2D	2	D	0.70	8	5.60	11	Moderate
3B	3	B	0.30	19	5.70	12	Moderate
2E	2	E	0.90	8	7.20	13	Significant
5A	5	A	0.10	78	7.80	14	Significant
3C	3	C	0.50	19	9.50	15	Significant
4B	4	B	0.30	39	11.70	16	Significant
3D	3	D	0.70	19	13.30	17	Significant
3E	3	E	0.90	19	17.10	18	High
4C	4	C	0.50	39	19.50	19	High
5B	5	B	0.30	78	23.40	20	High
4D	4	D	0.70	39	27.30	21	High
4E	4	E	0.90	39	35.10	22	Very High
5C	5	C	0.50	78	39.00	23	High
5D	5	D	0.70	78	54.60	24	Very High
5E	5	E	0.90	78	70.20	25	Very High

The picture is somewhat better for Schedule, with all the “Significant” bins consistently grouped in the middle, but there is still some inappropriate overlap between “Low” and “Moderate,” and again between “High” and “Very High.”

Appendix E: Risk Aphorisms

The below is a list of aphorisms the lead author frequently uses in teach Cost and Risk Analysis. The astute reader will find that many of them have been echoed throughout this paper.

- No estimate is ever right.
- Better to be approximately right than precisely wrong!
- Our estimate is always the highest, and it's never high enough! [SSG]
- Good risk analysis can't cure bad cost estimating. [RLC]
- All risk is relative. [PJB]
- Monte's never busy. [RLC]
- 0.6 is the new 0.3. [CBS]
- You want the 80th percentile? You can't handle the 80th percentile!
- They're asking for the 80th percentile, but they just want the real 50th percentile. [RLC]
- Beware the conspiracy of hope. [TNL]
- Lessons to be learned. [TNL]
- Hofstadter's Law: "It always takes longer than you expect, even when you take into account Hofstadter's Law" [CBS]
- Risk is something that happens to other programs. [The PM]
- The Program Manager's Serenity Prayer: Grant me the serenity to accept the risks I cannot mitigate; the courage to mitigate the risk I can; and the wisdom to know the difference.
- So you're tellin' me there's a chance! [PJB]
- Opportunities are like donuts: if there are any left on the table, it's because they were overlooked. [RLC]

Attributions:

[SSG] = Shishu Gupta

[RLC] = Dick Coleman

[PJB] = Peter Braxton

[CBS] = Dr. Christian Smart

[TNL] = Dr. Tzee-nan Lo