How to Catch a Black Swan:
Measuring the Benefits of the Premortem Technique for Risk Identification

by David L. Gallop

B.S. in Air Commerce and Transportation Technology, June 1984, Florida Institute of Technology
M.B.A. in Contract Management, May 1996, Florida Institute of Technology
M.S. in National Resource Strategy, June 2012, National Defense University
M.S. in Systems Engineering, May 2014, George Washington University

A Dissertation submitted to

The Faculty of
The School of Engineering and Applied Science
of the George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

January 31, 2017

Dissertation directed by

Christopher Willy
Professorial Lecturer in Engineering Management and Systems Engineering

John Bischoff
Professorial Lecturer in Engineering Management and Systems Engineering

The School of Engineering and Applied Science of The George Washington University certifies that David L. Gallop has passed the Final Examination for the degree of Doctor of Philosophy as of October 20, 2016. This is the final and approved form of the dissertation.

How to Catch a Black Swan:
Measuring the Benefits of the Premortem Technique for Risk Identification

David L. Gallop

Dissertation Research Committee:

Christopher Willy, Professorial Lecturer in Engineering Management and Systems Engineering, Dissertation Co-Chair

John Bischoff, Professorial Lecturer in Engineering Management and Systems Engineering, Dissertation Co-Chair

Sharam Sarkani, Professor of Engineering Management and Systems Engineering, Committee Member

Thomas Mazzuchi, Professor of Engineering Management and Systems Engineering & of Decision Sciences, Committee Member

E. Lile Murphree, Professor Emeritus of Engineering Management and Systems Engineering, Committee Member

© Copyright 2016 by David L. Gallop
All rights reserved

Dedication

To my family.

Acknowledgments

Over 130 defense acquisition professionals supported my research. Many participated as subjects in the risk identification experiments. Some of the Defense Systems Management College faculty supported the pilot testing or served as expert panel members. I am particularly grateful to the acquisition leaders that helped facilitate my experiments within their organizations. They are:

- Kim Herrington, Textron Defense
- Captain Keith Ingalsbe, United States Coast Guard
- Dean Pat Wills, Defense Systems Management College
- Captain Dan Britzinghoffer, United States Navy, PMS-515 Future Frigate
- Mary Redshaw, National Defense University, Eisenhower School

I could always count on my brother Robert's guidance and expertise. My cohort-mates were often a source of inspiration. I am grateful to Dr. Willy and Dr. Bischoff for their patience and encouragement as my advisors.

Abstract of Dissertation

How to Catch a Black Swan:
Measuring the Benefits of the Premortem Technique for Risk Identification

Risk identification is the most important step in the risk management process because risks that are unidentified are implicitly assumed and unmanaged. Projects face the challenge of increasing complexity and uncertainty. Risk management has taken advantage of powerful quantitative tools to analyze historical data and reports. However, these tools are not enough to identify the unprecedented, hard-to-predict, and potentially catastrophic risks called black swans. To "catch a black swan," we need to understand the capabilities and limitations of creativity-based risk identification tools. Group dynamics threaten the collaborative processes of creativity-based tools, leaving technology projects vulnerable to black swans. The premortem technique is the theoretical opposite of a postmortem. The premortem process starts by declaring that the project has failed. Premortem makes it safe for the team to be critical of each other, question assumptions, share information, study interrelationships and address risk areas that may otherwise go unaddressed. This study used several data collection and evaluation techniques (structured observation, subject surveys, and expert panel review) to study the collaborative power of the premortem technique against the time-tested brainstorming method when considering risks in a technology insertion (system upgrade) plan. The study compared the quality of the products resulting from the two techniques and found the premortem superior to brainstorming in three of the hypothesis tests. The findings are important to senior decision makers, educators, managers, and systems engineers.

Table of Contents

Dedication ..... **iv**

Acknowledgment ..... **v**

Abstract ..... **vi**

Table of Contents ..... **vii**

List of Figures ..... **xi**

List of Tables ..... **xiii**

1 Introduction **1**

1.1 Statement of the Problem 2

1.2 Background 4

1.3 Purpose 7

1.4 Significance 7

1.5 Scope and Limitations 8

1.6 Organization of the Document 8

2 Literature Review **9**

2.1 Establishing the Black Swan Context 9

2.2 Characteristics of the Black Swan 10

2.3 The Origins of the Black Swan 12

2.4 More Black Swans in Our Future 14

2.5 Deliberate Ignorance and the Black Swan 16

2.6 Risk: Objective (Risk as Fact) or Subjective (Risk as Feeling) 20

2.7 Creativity and Collaboration to Catch Black Swans 22

2.7.1 Critique and the Decision Balance Sheet 25

2.7.2 Devil's Advocate 27

2.7.3 Red Team 28

2.7.4 Scenario Planning 29

2.7.5 Brainstorming 32

2.7.6 The Premortem Technique 36

2.8 Premortem and Fault Tree Analysis Compared 47
2.9 Other Techniques for Creativity-Based Risk Identification 52
2.10 Summary of the Literature Review 59
3 Contributions 62
3.1 Great Expectations 62
3.2 A Safer Tomorrow 63
3.3 Ready for Wildness 66
3.4 Close the Gap Between Technical and Programmatic 68
3.5 The Challenge 70
4 Research Methodology 72
4.1 Research Goals 72
4.2 Research Methods 79
4.3 Research Plan 82
4.4 Data Collection Plan 85
4.4.1 Subjects and Expert Panel Members 86
4.4.2 Scenario Selection 89
4.4.3 Design of Experiment 92
4.4.4 Data Collection Instruments 101
4.4.5 The Instructional System Design 112
4.4.6 Peer Reviews/Pilot Tests and Mock Expert Panel 115
4.5 Data Analysis Methods 135
4.5.1 Lilliefors Test for Normality 136
4.5.2 Mann-Whitney Test (Nonparametric) 137
4.5.3 Box-Cox Transformation 137
4.5.4 Parametric Tests 138
4.5.5 Goodman and Kruskal's Gamma 139
4.6 Summary of Research Methodology 139
5 Results 142
5.1 Demographic Data and Response Rates 142
5.2 Null Hypothesis #1 148
5.3 Null Hypothesis #2 153

5.4 Null Hypothesis #3 155
5.5 Null Hypothesis #4 156
5.6 Expert Panel Votes on the Top-3 Deliverable Sets 157
5.7 Subject Survey Results for Items 9 thru 12 158
5.8 Subject Survey Results for Items 13 thru 15 164
5.9 Group Performance (Student, Government, Industry) 168
5.10 Goodman and Kruskal's Gamma , Strength of Association 177
5.11 Data Validity 179
5.12 Anecdotal Subject Reports and Researcher Observations 181
5.13 Summary of Results 184
6 Conclusions and Recommendations 188
6.1 Conclusions 188
6.2 Recommendations 189
6.2.1 Taxonomy of Ignorance Awareness 189
6.2.2 Brainstorming Training and Discipline 190
6.2.3 Senior Decision Makers Expect Creative Solutions 191
6.2.4 More Study Needed on Risk Identification Techniques 191
6.2.5 More Risks Does Not Guarantee Better Risks 192
6.2.6 Teams Think and Act Differently Using Premortem 192
6.2.7 Creativity-Based, Best Defense Against the Black Swan 193
6.3 Recommendations for Future Research 194
6.3.1 Before and After Study 194
6.3.2 Change the Scenario 195
6.3.3 Integrate the Team 195
6.3.4 Longitudinal Study 195
6.3.5 A Different Treatment-A 196
6.3.6 Further Analysis of Structured Observation Data 196
References 197
Appendix-A Expert Panel Member Biographies **220**
Appendix-B Technology Insertion Scenario- Instructions to Subjects **243**

Appendix-C Marlin Torpedo Project Charter **245**

Appendix-D Subject Questionnaire **248**

Appendix-E Team Deliverable Templates **251**

Appendix-F Expert Panel Instructions Sheet **254**

Appendix-G Expert Panel Score Sheet **257**

List of Figures

Figure-1 Three Characteristics of the Black Swan **10**

Figure-2 Where do Black Swans Come From? **13**

Figure-3 Modified Taxonomy of Ignorance **17**

Figure-4 Two Schools of Risk (Objective and Subjective) **21**

Figure-5 Transactive Memory Model of Collaboration **24**

Figure-6 Decision Balance Sheet **26**

Figure-7 Scenario Planning Steps, SRI Model, and Envelope of Uncertainty **30**

Figure-8 Brainstorming Steps **32**

Figure-9 Pitfalls of Brainstorming **34**

Figure-10 Inductive versus Deductive Thinking **38**

Figure-11 Summary of Veinott et al. Findings **45**

Figure-12 Example of a Fault Tree **50**

Figure-13 Implications Wheel **53**

Figure- 14 Six Thinking Hats **54**

Figure-15 Strengths-Weaknesses-Opportunities-Threats (SWOT) Analysis Grid **56**

Figure-16 Layer of Protection Analysis (LOPA) Model of the Challenger Disaster **57**

Figure-17 Swiss Cheese Model **58**

Figure- 18 An Example of the Persistence and Popularity of Brainstorming **73**

Figure-19 Brainstorming vs Premortem **75**

Figure-20 Increasing Collaboration Expected to Aid in Catching Black Swans **76**

Figure-21 "Catchable" Black Swans **78**

Figure-22 Defense Systems Management College (DSMC) Executive Resident Courses **87**

Figure-23 Textron Systems Overview **88**

Figure-24 N-Gram Viewer Results for Words and Phrases Related to the Mark-14 Case **91**

Figure-25 Data Collection Diagram **95**

Figure-26 Treatment Activity Timelines Compared **98**

Figure-27 Structured Observation Data Collection Instrument **102**

Figure-28 Team Deliverables Data Collection Instrument **104**

Figure-29 Components of the Subject Survey/Questionnaire.....**106**

Figure-30 Scoring for Each Deliverable Element.....**109**

Figure-31 Lay-out for the Expert Panel Instructions Sheet **111**

Figure-32 Lay-out for the Expert Panel Grading Sheet **112**

Figure-33 The ADDIE Framework for Instructional Design.....**113**

Figure-34 Peer Review #1 Structured Observation Findings **117**

Figure-35 Peer Review #2 Validated Changes to the SO Data Collection Instrument **121**

Figure-36 Peer Review #2 Validated the Benefit of Providing Templates for the Team Deliverables **123**

Figure-37 Peer Review #3 Validated the Benefit of Providing a Process Worksheet **125**

Figure-38 Mock Expert Panel Member Scoring for Risk Dimensions.....**129**

Figure- 39 Mock Expert Panel Scoring for Changes to the Plan Dimensions.....**131**

Figure-40 Subject Experience Level and Duties.....**144**

Figure-41 Subjects' Most Used Risk ID Technique.....**145**

Figure-42 Subject Survey Items 5 thru 8 .....**152**

Figure-43 Subject Survey Items 9 thru 12 .....**159**

Figure-44 Subject Survey Items 13 thru 15 .....**165**

Figure-45 Number of Experiments by Organization Type and Treatment.....**168**

Figure-46 Total Top-3 Votes by Organization Type and Treatment.....**169**

Figure-47 Average Expert Panel Top-3 Votes by Organization and Treatment per Experiment.....**170**

List of Tables

Table-1 Summary of Known Premortem Technique Application Cases **47**

Table-2 Comparison of Creativity-Based Risk Identification Techniques **60**

Table-3 Subject Teams by Organization, Treatment, and Team Size **143**

Table-4 Hypothesis Test Showing No Difference in Treatment Demographics **146**

Table-5 Testing Null Hypothesis #1 based on Structured Observation **151**

Table-6 Survey Questions #5 through #8 Confirm Structured Observation Findings **153**

Table-7 Test Results for Null Hypothesis #2 Quality Risks **154**

Table-8 Test Results for Null Hypothesis # 3 Quality Changes to the Plan **155**

Table-9 Test Results for Null Hypothesis #4 Black Swan Hits **156**

Table-10 Test Results for Expert Panel Member Top-3 Voting **158**

Table-11 Test Results Subject Survey Item #9 "More Risks" **160**

Table-12 Test Results Subject Survey Item #10 "Better Risks" **161**

Table-13 Test Results Subject Survey Item #11 "Creative Improvements" **162**

Table- 14 Test Results Subject Survey Item #12 "Avoid Organizational Blind Spots" **163**

Table-15 Summary Statistics for Subject Survey Items 13, 14, and 15. ..... **166**

Table-16 ANOVA for Subject Survey Items #13 and #14 ..... **167**

Table-17 ANOVA Results for Average Top-3 Votes **171**

Table- 18 Student Group Scores by Treatment with the Alternative Hypothesis that Brainstorming Scores are Less than Premortem **172**

Table-19 Government Group Scores by Treatment with the Alternative Hypothesis that Brainstorming Scores are Less than Premortem **173**

Table-20 Industry Group Scores by Treatment with the Alternative Hypothesis that Brainstorming Scores are Less than Premortem **174**

Table-21 Comparing the Student and the Government Group Means to the Industry Group Means **175**

Table-22 ANOVA and Box Plot Comparing Group Creativity **176**

Table-23 Test Results Comparing Creativity by Group and Treatment (Alternative Hypothesis is Brainstorming Creativity Scores are Lower than Premortem Creativity Scores).....**177**

Table-24 Gamma Scores for Associations with Expert Panel Rankings Based on Top-3 Votes.....**178**

Chapter 1 - Introduction

What more can be said on the topic of risk? A quick search of the Internet will return thousands of sites that discuss risk in one form or another. Risk management has been a field of practice since the Renaissance (Bernstein 1998). The advancements in computers and software applications have made risk management tools readily available to systems engineers, project managers, and policymakers. With all our knowledge, experience, and tools for managing risk, it does not appear that we are becoming better risk managers. Perhaps our tools for predicting events and their impacts that fit quite nicely in the disconnected and centrally planned world of the Industrial Age have betrayed us in the interconnected and distributed Information Age.

Three books influenced the author to investigate the subject of risk and a practical approach to dealing with the uncertain and complex world in which our systems and technical projects exist. In *The Polar Bear Strategy: Reflections on Risk in Modern Life,* John F. Ross theorizes that we face more risks than prior generations because of the interconnected nature of the modern world. Those risks are more complex and require finesse to extract the big picture before acting. Ross also mentions that human-made risks today are of the low-probability, high-impact variety resulting in rare but devastating conditions (Ross 1999). Nassim Taleb builds on the premise of the low-probability, high impact event in *The Black Swan: The Impact of the Highly Improbable*. He makes the argument that spending resources on predicting and controlling typical risks can be better used to identify and avoid the low probability but devastating event, or

position oneself for the once-in-a-lifetime opportunity (Taleb 2010). Both Ross and Taleb suggest that relying on conventional risk management tools is not enough. In *Streetlights and Shadow: Searching for the Keys of Adaptive Decision Making*, Gary Klein offers the premortem technique as an unconventional way to derive the highly improbable from the minds of a team (Klein 2009). Teams can then change the plan to address the highly improbable once those risks are out in the open.

In 2010, as part of a U.S. Army study, the research team of Veinott, Klein, and Wiggins performed a scientific study on the premortem technique. The study measured the effect the premortem technique had individuals' confidence in the plan. They were studying the common problem within crisis management planning teams having an inflated belief that a plan will be successful (Veinott, Klein, and Wiggins 2010). The study found that the premortem technique did reduce confidence in the plan (causing the team to address weaknesses in the plan) more than other techniques for risk identification (Veinott, Klein, and Wiggins 2010). This dissertation builds on Veinott, Klein, and Wiggins' study by comparing the quality of the risks identified and proposed changes to the plan from a team using the premortem technique to a team using the brainstorming technique for risk identification. The intent of this study is to justify the addition of the premortem technique to the systems engineer and project manager's toolkit.

1.1 Statement of the Problem

As systems get bigger and more complex, we will face problems that have never been encountered before. Our traditional risk identification tools based on historical data

and Boolean logic may not be robust enough in the uncertain and complex environment of 21st-century systems engineering.

The International Council on Systems Engineering (INCOSE) published its strategic vision for systems engineering in 2014 that identified a number of challenges for the profession in the coming decade. *A World in Motion: Systems Engineering 2025*, INCOSE notes:

> *"The evolution of systems is an ever increasing complexity which can be observed in terms of the number of system functions, components, and interfaces and their nonlinear interactions and emergent properties. Each of these indicators of complexity has increased dramatically over the last fifty years, and will continue to increase…"* (INCOSE 2014).

The nonlinear and emergent properties of systems will surely produce problems that are unprecedented (Taleb, Goldstein, and Spitznagel 2009). Again, INCOSE reports that "current systems engineering practice, based on well-defined processes and innovative analytic approaches, has demonstrated significant value to stakeholders, but in the future, the systems community must tackle many new fundamental interdisciplinary and integration-related challenges" (INCOSE 2014). Quoting Todd Bayer from the Jet Propulsion Laboratory, INCOSE identified two emerging systems engineering challenges relative to risk and risked-based decision making. The first challenge is the "technical and programmatic sides are poorly coupled … hampering project risked-based decision making" (INCOSE 2014). Second, "most major disasters such as Challenger and Columbia have resulted from the failure to recognize and deal with risks" (INCOSE 2014). These challenges call for us to look for new tools to identify and communicate

risks. The premortem technique for risk identification may enhance a team's ability to identify and communicate risks.

1.2 Background

Gary Klein is a researcher in the area of naturalistic decision making (NDM), which is the study of how people make decisions in real-world, complex settings. He discovered that people did not make decisions by using the familiar process of generating alternatives and comparing those alternatives against evaluation criteria (Klein 2008). Klein and his NDM colleagues found that people relied on "prior stage of perception and recognition of situations" in addition to a mental simulation of appropriate responses in making decisions (Klein 2008). These studies in NDM led Klein to define the premortem technique.

In the September 2007 Harvard Business Review, Klein connects his work in NDM with Mitchell, Russo, and Pennington's 1989 study on "prospective hindsight" to define the premortem technique. Mitchell, Russo, and Pennington found that a backward perspective on the outcome to convey its certainty generated more insight into events that might lead to that outcome (Mitchell, Edward Russo, and Pennington 1989). From the theory of prospective hindsight and NDM, Klein then describes the steps of the premortem technique:

- Get familiar with the plan
- Declare that the plan has failed
- Team members individually record every reason they can think of why the plan failed

- Each member describes a risk from their list until all items are exhausted
- The team revisits the plan in light of the identified risks

In the same Harvard Business Review article, Klein recalls facilitating a premortem session in which a team member kept silent about a risk during the project kick-off but revealed the risk during the premortem (Klein 2007a). At that point, a different team member offered a solution that negated the risk. In another premortem session, the team identified time constraints that had not been discussed during the kick-off. Many teams perform pre-launch risk analysis, but the premortem approach offers benefits other methods do not. The premortem technique reduces the "damn the torpedoes" attitude and sensitizes the team to pick-up early signs of trouble (Klein 2007a).

After Klein's 2007 article, there were a number of anecdotal references to the successful use of the premortem technique in business strategy, logistics planning, and venture capital scenarios. In 2010, under the U.S. Army's Advanced Decision Architectures Collaborative Technology Alliance program, Veinott, Klein and Wiggins performed a scientific experiment on the effectiveness of the premortem technique in countering overconfidence in a plan. The study involved paid undergraduate students and compared the premortem technique against several other risk identification techniques: critique, pro-con, and con. The study indicated that the premortem countered overconfidence in a plan more so than the alternatives (Veinott, Klein, and Wiggins 2010). Their findings also indicated an increase in the subject's understanding of the plan when the subject assessed the plan using the premortem technique (Veinott, Klein and Wiggins 2010, 7). The authors suggest that future research should use experienced

planners and leaders, and might compare the premortem technique to other methods (Veinott, Klein, and Wiggins 2010).

In my study, I use experienced planners and leaders from the United States Department of Defense (DoD), the United States Coast Guard (USCG), and Textron Corporation to conduct my experiments. The subjects include program managers for some of the DoD's most expensive and complex weapon systems. In my experiments, I compare the premortem technique to the commonly used brainstorming technique for risk identification. I compare the two techniques for team collaboration using structured observation and use an expert panel to compare the two techniques for the quality of risks and changes to the plan. Since I have based the experimental scenario on a historical case in which risks blind-sided the team, the expert panel can determine which treatment catches more risks the historical team missed (i.e., black swans).

Veinott, Klein and Wiggins constructed their study around a crisis planning scenario of an influenza outbreak (Veinott, Klein, and Wiggins 2010). My experiments used the premortem and brainstorming techniques in a systems engineering context. Specifically, I looked at the risks in a technology insertion scenario. That scenario can be generalized to other product improvement plans. I selected this type of scenario because they are more common than "new start" projects. In a systems upgrade project, the low-probability, high-impact risks often go unmentioned because of overconfidence and other team dynamics (Reeves et al. 2013).

At this point, a reader schooled in systems engineering may be wondering how the premortem technique differs from Fault Tree Analysis (FTA). FTA identifies a system failure and all the contributing events to that failure (Stamatelatos and Caraballo

2002). Isn't the premortem technique simply a business application of FTA? Although both techniques are deductive, their similarities end there. The premortem technique is based on *behavioral decision making* while the FTA is based on Boolean logic and probability (Howat 2008). "The premortem technique goes beyond the technical issues (which FTA focuses on) and is sensitive to context—the skills and weaknesses of the team, the organizational structures to be contended with" (Klein 2015). The purpose of this research was not to propose the premortem technique as an alternative to FTA; rather, its purpose was measure the benefits of the premortem as an additional systems engineering tool.

1.3 Purpose

The purpose of this research was to determine if the premortem technique works better than brainstorming for risk identification in a systems engineering context. The thesis was that brainstorming identifies the typical and predictable risks while the premortem technique helps identify the atypical and hard-to-predict risks (black swans). By identifying these black swans, the premortem technique helps teams develop higher quality changes to the plan.

1.4 Significance

Risk identification is the most important step in the risk management process, because the team cannot manage risks it fails to identify and discuss. The team tacitly assumes all unidentified risks. However, it is not enough to have the team identify more risks. The team must be more creative in identifying risks to avoid spending limited resources managing typical risks so that it may better prepare the project for black swan

risks. This study provided engineering managers and systems engineers with an additional tool in the risk identification toolkit. The premortem technique can enable robust risk identification, high-quality changes to the plan, and help bridge the gap between the technical and programmatic sides of the program.

1.5 Scope and Limitations

This study focused on the use of two creativity-based techniques for risk identification: brainstorming and premortem. Teams used these two techniques in experiments based on the same systems engineering scenario. That scenario is a systems upgrade scenario. Although the experiments use experienced leaders, the participants were not working on a plan of their own design.

1.6 Organization of the Document

Chapter 2 will cover the literature related to black swan risks, deliberate ignorance, risk identification, collaboration, and creativity-based risk identification techniques. In Chapter 3, I will describe how this study will contribute to the engineering management and systems engineering fields. I will explain my experimental design, data collection, and analysis methods in Chapter 4. I will present the results from the experiments in Chapter 5 and recommendations in Chapter 6.

Chapter 2 - Literature Review

2.1 Establishing the Black Swan Context

Before exploring the relevant literature, it is important to understand the context in which black swans exist. Black swans are those rare but often disruptive risks (Posner 2010). Black swans arise from our inability to identify the atypical, unprecedented risks and make appropriate changes to our plans to account for their occurrence (Taleb, Goldstein, and Spitznagel 2009). Keep in mind that black swans come in both the positive (beneficial) and negative (harmful) varieties. Do traditional risk management methodologies help teams with the early identification of black swans?

Risk management methodologies are consistent in their recommendation to start by identifying risks. However, those same methodologies provide little advice on how to perform this crucial first step of risk management. This seems to be a major flaw in the practice of risk management and the International Council on Systems Engineering (INCOSE) technical management process. Every risk left unidentified is a risk the team tacitly accepts and leaves unmanaged. In the worst case, risk identification is an ad hoc activity, performed without adequate planning or preparation. In the best case, the project manager and systems engineer will establish cross-functional teams to identify risks using formal risk management processes and tools. Those cross-functional teams will review relevant historical information and analyze the current environment to identify risks in the project plan. The process of reviewing history and analyzing the

environment is likely to yield typical and predictable risks. Since black swans are not typical and often unprecedented risks, we may remain blind to black swan risks.

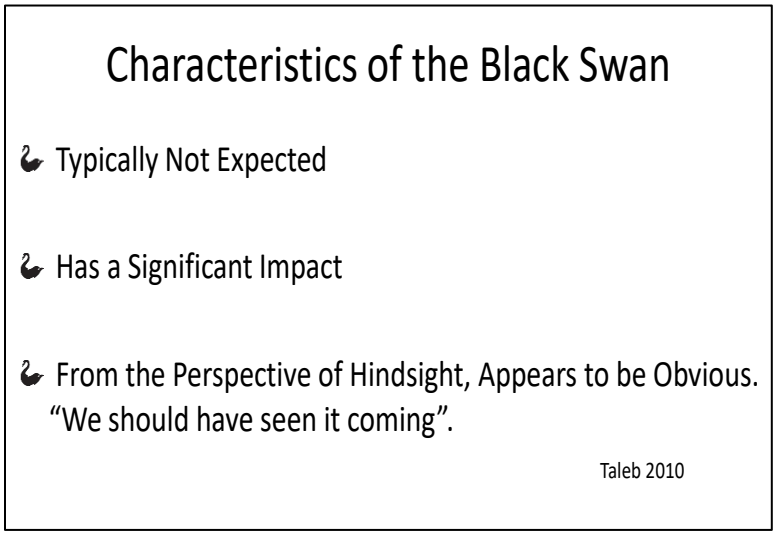


Figure 1 Three Characteristics of the Black Swan

Group dynamics contribute to black swan blindness. We depend on teams to perform risk identification because no single individual has all the knowledge, skills, and experience to identify risks in complex systems. Teams of subject matter experts must communicate information and perspectives to identify risks. All teams struggle with common group dynamics such as explaining away problems, overconfidence, and group think, to name a few. In most cases, it is up to the team leader to prevent or mitigate these dynamics. There are risk identification methodologies that do address these sources of black swan blindness. They include devil's advocate, red teams, scenario planning, fault tree analysis, brainstorming, and premortem. Each risk identification technique has its advantages and disadvantages in catching a black swan.

2.2 Characteristics of the Black Swan

There was a time when European scientists believed all swans were white. They regularly confirmed this through empirical evidence. The only swans Europeans ever encountered were indeed white. Then, in 1697, a Dutch explorer discovered black swans in Western Australia (Taleb 2010). The term black swan has become a metaphor for encountering the unexpected. Nasim Taleb in his book *The Black Swan: The Impact of the Highly Improbable* describes three characteristics of the black swan event (Figure 1). First, the black swan is typically unexpected. Second, the black swan has a significant impact (positive or negative). Third, the black swan is obvious (appears to be a "white swan") given the perspective of hindsight (Taleb 2010). The following are examples of recent projects that failed to anticipate and prepare for black swan risks.

The German Electronic Health Card and the Levi Strauss & Company Enterprise Resource Planning (ERP) projects are two cases in which interdependencies resulted in the emergence of a black swan (Buhl 2012). In the case of the Levi Strauss & Company, technical difficulties integrating a new ERP system with a highly interdependent set of legacy systems resulted in a loss of nearly $200 million U.S. (Buhl 2012). The German Electronic Health Card system's black swan resulted from the system's complex web of stakeholders that included physicians, hospitals, and health insurance companies. When the black swan of unspoken requirements emerged from the stakeholder needs, the initial investment cost estimate rose from €511-million to €1.4-billion (Buhl 2012).

The General Motors (GM) Cobalt is an example in which the black swan of airbags failing to deploy lay hidden behind the interdependency of a poorly designed ignition switch and a well-designed safety mechanism. The poorly designed ignition switch caused vehicles to stall while moving. Since drivers could still maneuver their

cars, GM engineers considered the ignition switch failure to be an acceptable failure and a customer inconvenience. This "inconvenience" coupled with a design to protect passengers from injury by airbags deploying in parked cars denied drivers lifesaving protection in at least 54 crashes (Valukas 2014).

The German Electronic Health Card and the Levi Strauss & Company ERP cases illustrate how complexity and interconnectivity are perfect conditions for the black swan. As systems connect with other systems, they generate interdependencies that have never been seen before. The new interdependencies are undocumented, and emergent properties of the system that are either new to the entire body of knowledge or new to the team. The Chevy Cobalt case illustrates how a perfect design can have tragic consequences when teams do not share and communicate different perspectives of risk. The three cases mentioned provide some insight into the specific origins or sources of the black swan.

2.3 The Origins of Black Swans

Aven and Krohn identified three origins or sources of black swan risks (Figure-2). The black swan can come from outside the existing body of knowledge (Aven and Krohn 2013). These are often referred to as unknown, unknowns. In the second source, black swans can come from outside a specific perspective (Aven and Krohn 2013). In that case, the black swan is unknown to the project team but known to others. Overconfidence and time pressures may lead a team to wear metaphorical blinders. The team fails to look outside the team's experience and knowledge base to discover risks that are known to other teams or field but potential black swans to them.

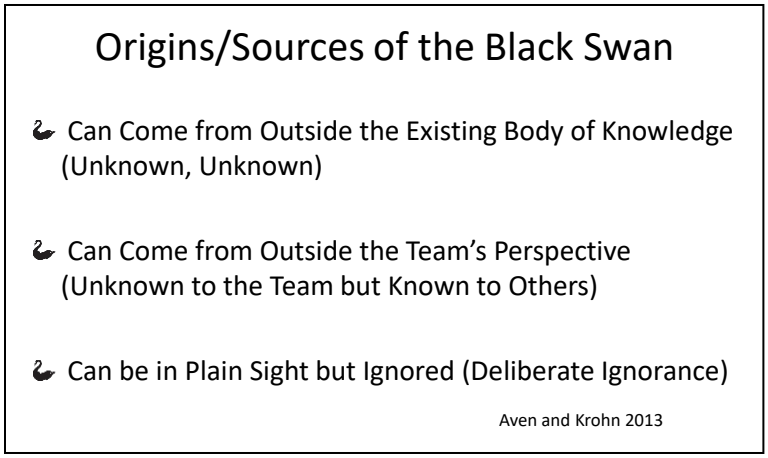


Figure 2 Where do Black Swans Come From?

In the third origin of the black swan, the risk can appear directly in the field of view of the project team, but the team judges the risks probability to be negligible (Aven and Krohn 2013). This is a common origin of black swans because group dynamics increase our vulnerability to this source of black swans. The 1986 space shuttle Challenger launch decision is a classic example in which group dynamics led the team to explain away a catastrophic risk that was in plain sight.

Morton-Thiokol was the manufacturer of the space shuttle's solid rocket booster (SRB). By all accounts, the Morton-Thiokol engineering team had identified an increased risk of hot gasses passing through joints in the SRB as outside air temperature decreased. The lower the air temperature, the more hot gas passed through the SRB joints (Everett and Mulholland 2006). Flexible O-rings and putty would normally seal the SRB joints. However, when the outside air temperatures approached freezing, the O-rings and putty lost their ability to maintain an adequate seal. The Morton-Thiokol engineers knew that the escaping hot gasses could ignite the external fuel tank, causing a catastrophic failure to the space shuttle and kill the crew. Over the course of a number of

cold weather launches, the team rationalized the problem (Vaughan 1996). The black swan was staring the Morton-Thiokol team in the face during Challenger prelaunch discussions. The Morton-Thiokol team developed what James Chiles called *fatal confidence* or the belief that one knows all that needs to be known (Chiles 2002). The team chose to ignore the black swan risk resulting in the tragic deaths of the Challenger crew when hot gasses passed through an SRB joint and ignited the external fuel tank 73-seconds into the launch (Vaughan 1996).

So far we have explored the context and origins of black swans and highlighted the nature and sources of black swan risks using several historical, systems engineering examples. Interdependencies of systems due to the Internet and globalization are increasing. We rely more and more on teams of teams to manage the increased complexity arising from systems interdependencies. Socio-technical trends guarantee that we will see more black swans in our future.

2.4 More Black Swans in Our Future

Three socio-technical trends are shaping the future. That future will have conditions that make black swans more frequent and potentially more disruptive. The three socio-technical trends are: increasing systems complexity, greater need for specialization, and reliance on multi-team systems working under time pressures (Kutsch et al. 2013; Herazo, Lizarralde, and Paquin 2012).

We can expect the number of interfaces between systems and the complexity of systems and projects to increase (Perminova, Gustafsson, and Wikström 2008; INCOSE 2014). Globalization and an ever increasing dependency on networks for functionality

add to complexity (Taleb 2009). This is fertile ground for the first source of black swans, the "unknown-unknowns." As systems get larger and more complex, we will see more problems that we have never encountered before (Chiles 2002). Even a tiny "improvement" to technology embedded in a complex world of human practices can have profound effects (Ashton 2015).

As the demand for a more productive workforce increases, we can expect that workforce to become even more specialized (Rouwette, Bastings, and Blokker 2011). As specialization increases, specialists tend to focus solely on the information available within their field of specialization (Leckie, Pettigrew, and Sylvian 1996). In an environment of rapid change, we also depend more on teams of subject matter experts to predict the future and identify risks (Farrington, Henson, and Crews 2012). These experts predict risks from their specialized perspective. A dependable framework for combining diverse expert perspectives has remained elusive (Keeney 2013). Experts become experts because they have mastered a paradigm. A paradigm is a system of beliefs, experiences, and assumptions. That paradigm focus can leave experts "black swan blind" due to *selective attention*. Selective attention is seeing what you expect to see and not necessarily what is there (Ashton 2015). Developing expertise is essential, but it can block us from seeing the unexpected. Our increasing reliance on both specialization and subject matter experts narrows our field of view. As mentioned earlier, this is the second origin of black swans: risks unknown to the team but known to others outside the team.

The conditions favorable for the third source of black swans are also increasing. We may see a greater number of teams (like the Morton-Thiokol team) choosing to

ignore the black swan appearing before them. Groups of interconnected teams (multi-team systems) are becoming more important to project success. These groups work on critical tasks under constant time pressures, so they face difficult information-sharing challenges (T. R. Chen et al. 2014). One information-sharing challenge INCOSE has identified is the communication gap between the technical and programmatic sides of project teams (INCOSE 2014). That gap makes it difficult for teams to collaborate and understand the different perspectives individual team members can have concerning the same risk. Deliberate ignorance may also lead a team dismiss the black swan directly in the team's field of view.

2.5 Deliberate Ignorance and the Black Swan

In 1989, Michael Smithson studied the roles that uncertainty and ignorance play in team decision making. He reviewed research from the fields of cognitive psychology, social psychology, organizational studies, sociology, and social anthropology to develop a conceptual framework for understanding decisions under uncertainty and ignorance. Smithson's taxonomy divides ignorance into two categories (Figure 3). The first category is ignorance due to error. We never have complete information about the future. Ignorance due to error is either a passive distortion of the available information or a state of having incomplete information (Smithson 1989).

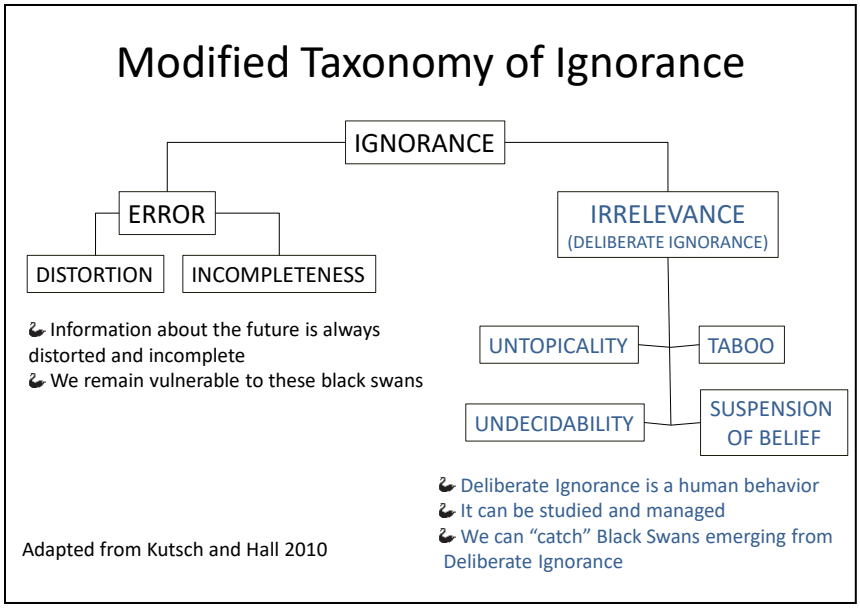


Figure 3 Modified Taxonomy of Ignorance

The second category of ignorance is irrelevance. Smithson decomposes irrelevance into three elements: untopicality, taboo topics and undecidability. Kutsch and Hall, building on Smithson's work, added *suspension of belief* as the fourth element of irrelevance. In the case of *irrelevance*, the information is available but not believed to be worthy of consideration (Kutsch and Hall 2010). Irrelevance is also referred to in the literature as *deliberate ignorance*. Since deliberate ignorance more clearly expresses the idea of a team selectively ignoring information, I will use *deliberate ignorance* instead of irrelevance. We can study deliberate ignorance further by analyzing the four elements.

Untopicality is when the team identifies information relative to the risk in a project but decides that the risk information is off-topic. A team of subject matter experts will generally discuss what is critical from their own personal experience, ignoring anything else (Margolis 2008). This results in a deliberate narrowing of the team's field of view (Kutsch and Hall 2010).

Taboo topics reflect some restriction to act placed on the team when considering information relevant to the risks in a project (Smithson 1989). Certain stakeholders in a project may not have any tolerance for some risks. In which case, the identification and management of those risks may place the existence of the project in jeopardy. The team would treat those risks as taboo and not discuss or act on them (Kutsch and Hall 2010).

Undecidability recognizes that not everyone performing risk identification will agree that a risk is real. If the team cannot agree that a risk is real, the risk may be set aside without any further discussion or action (Kutsch and Hall 2010). A common practice is to place difficult and contentious risks in the "parking lot." The intention of the parking lot is to keep the agenda moving while providing the person who raised the contentious risk some assurance that the risk will be acted upon at a later time (DeCarlo 2004). That later time often does not occur and the risk becomes ignored.

Kutsch and Hall conducted interviews of 18 information technology project managers to explore the extent of deliberate ignorance in the risk identification phase of projects. They found that project managers focused on specific information and declared other information as irrelevant, falling into the *untopicality* subcategory of deliberate ignorance (Kutsch and Hall 2010). Their findings point towards a tendency for project managers to focus on information that is easiest to process (Kutsch and Hall 2010). This leads to the team managing a stable of tame risks. Project managers prefer tame risks because those risks have many of the characteristics of a typical project activity. Tame risks are well defined, stable, have a concluding point, and are similar to previously solved problems (Kutsch and Hall 2010).

Kutsch and Hall also found it common for the 18 information technology project managers to suppress risks that may cause discomfort among stakeholders (Kutsch and Hall 2010). Those project managers had declared those risks to be *taboo*. Kutsch and Hall point to two reasons for these taboo risk topics. Project managers want their stakeholders to view them as being competent. Predicting calamity undermines that air of competence (Kutsch and Hall 2010). Another reason for taboos is that the project managers want to provide their stakeholders with information that the project manager believes the stakeholder wants to hear (Slovic et al. 2004). This reinforces the appearance that the project is safe and predictable (Kutsch and Hall 2010).

The third subcategory of deliberate ignorance is *undecidability*. Kutsch and Hall found that the lack of agreement on whether a risk was real often came down to a few reasons. First, some team members may believe that there are hidden agendas or undisclosed objectives among those identifying the risks (Kutsch and Hall 2010). Second, some team members may question the credibility of the person identifying the risk (Kutsch and Hall 2010). Finally, a general lack of trust among the team members performing the risk identification could result in undecidability (Kutsch and Hall 2010; Porath and Erez 2009).

In the course of their research, Kutsch and Hall discovered a fourth subcategory of deliberate ignorance they call *suspension of belief.* In suspension of belief, the team regards risk-related information as useless because it has no immediate impact on the project (Kutsch and Hall 2010). They found that project managers distanced themselves from the risk information with phrases like: "What can we do to prevent that?" and "If it did happen (the risk), we could get around it?" (Kutsch and Hall 2010). Why does

suspension of belief occur? It most likely comes down to the fact that managing risks consumes resources. It is more cost effective to suspend belief on a fictional risk than committing limited resources to managing the risk upfront (Kutsch et al. 2013).

Deliberate ignorance is an area that leaves our projects vulnerable to black swans. Unlike ignorance from error, we can address the deliberate ignorance problem. The premise that we can become less vulnerable to black swans by controlling deliberate ignorance is central to the experiments described later in this dissertation. At this point, we should consider the different schools of risk, risk identification techniques, and their capabilities and limitations in the ability to detect and prepare for black swans.

2.6 Risk: Objective (Risk as Fact) or Subjective (Risk as Feeling)

There are two schools of risk (Figure 4). The first school considers risks as an objective fact. The second school considers risk as a subjective construct. Let us explore both schools in light of their ability to help us catch black swans.

Proponents of risk as being an objective fact see risk as being disassociated with people's minds and feelings (Zhang 2011). This school attempts to treat risks as probability distributions (Mandelbrot and Taleb 2010). This allows the study of risk to conform to a variety of analytical methods (Starr and Whipple 1980). These methods have been fully embraced by the systems engineering community (INCOSE 2014). The objective school of risk depends upon the explicit information within the organization (El Badawy, Marwan, and Magdy 2015). Sources of explicit information relevant to risk identification include databases and status reports to enable quantitative tools and methods.

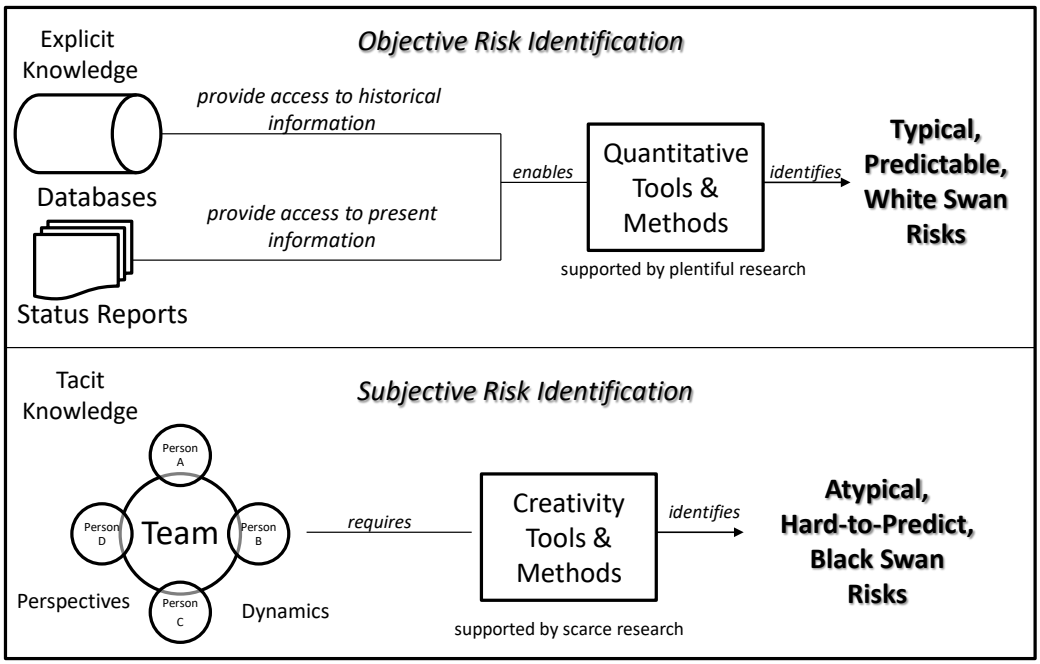


Figure 4 Two Schools of Risk (Objective and Subjective)

Risk as a feeling proposes that risks are a mental construct of people considering future developments (Zhang 2011). This implies that risk identification involves emotional, political, and situational aspects that do not fit neatly into Boolean logic. The subjective school of risk depends on tacit knowledge. Tacit knowledge exists below the organization's surface, inside a person's mind (El Badawy, Marwan, and Magdy 2015). We need to use creativity-based tools to access the tacit knowledge of the team members. A decision to ignore information concerning a risk whether it be due to untopicality, taboo, undecidability, or suspension of belief is rooted in the risk as feeling school. However, the risk as feeling has been under-represented in the project management literature, according to Zhang.

Zhang's 2011 study analyzed two project management journals, the *International Journal of Project Management* and the *Project Management Journal*, for articles

addressing risk tools and methods published in the preceding ten years. The study found 171 research articles that addressed risk and classified the subject matter as falling into one of the two schools of risk: risk as an objective fact and risk as a subjective construct. Of the 171 articles, only 12, or 7 percent, of the articles explored the subjective or feeling aspects of risk (Zhang 2011). Needless to say, more research is required to understand the tools and methods that assess risk as feeling, especially in the context of systems engineering (INCOSE 2014).

The Project Management Institute's (PMI) *Practice Standard for Project Risk Management* discusses three categories of risk identification tools and methods: historical review, focused on the past; current assessment, focused on the present; and creativity techniques, focused on the future (PMI 2009). The risk identification tools and methods based on historical information and current assessments support the risk as objective fact school. However, the techniques of reviewing lessons learned or benchmarking analogous projects may be of little help in the early identification of black swans. Objective, quantitative techniques are based on precedents and are looking for signs that the precedent is about to reappear (Taleb, Goldstein, and Spitznagel 2009). Black swans often lack precedence to the body of knowledge or the team's experience.

Since risk identification is a human information seeking behavior regarding the future of a project, we need more research on the creative tools and methods for risk identification. As systems become increasingly interconnected, interdependent, and complex, we may depend more heavily on the creative tools and techniques to avoid being blindsided by risks.

2.7 Creativity and Collaboration to Catch Black Swans

Creativity-based tools for risk identification and team collaboration may be the best ways to identify risks in complex and interdependent systems engineering projects (DeBakker, Boonstra, and Wortmann 2011). Using creativity for the early detection of black swans is important, because creativity expands our ability to innovate and anticipate the previously unforeseen (Sinfield, Gustafson, and Hindo 2013). Through the act of creation, we open and reshape our understanding and begin to notice things we had not noticed before (Ashton 2015). This makes creativity a powerful tool in catching black swans.

The 9/11 attacks are generally considered a black swan event. In the case of the 9/11 attacks, the black swan was the result of the malevolent use of creativity by the terrorists and the lack of creativity on the part of American intelligence and law enforcement agencies (James and Drown 2008). One could argue that if only the intelligence and law enforcement agencies had been more creative than the terrorists, the attacks could have been foreseen. It is difficult to make the argument that better quantitative tools could have stopped the attacks from occurring; however creativity itself is not enough. We have to combine creativity with collaboration.

Collaboration is the alignment of team members to achieve a common goal and includes the work processes to produce the desired results (F. Chen et al. 2013). Collaboration is necessary because the perceived complexity of systems overwhelms a single individual's cognitive abilities (Matta and Ashkenas 2003). The lack of collaboration results in inefficiencies and ineffective team member involvement (Konate, Sahraoui, and Kolfschoten 2013). The *Transactive Memory Model* (TMM) provides powerful insight on the collaboration process (Figure 5).

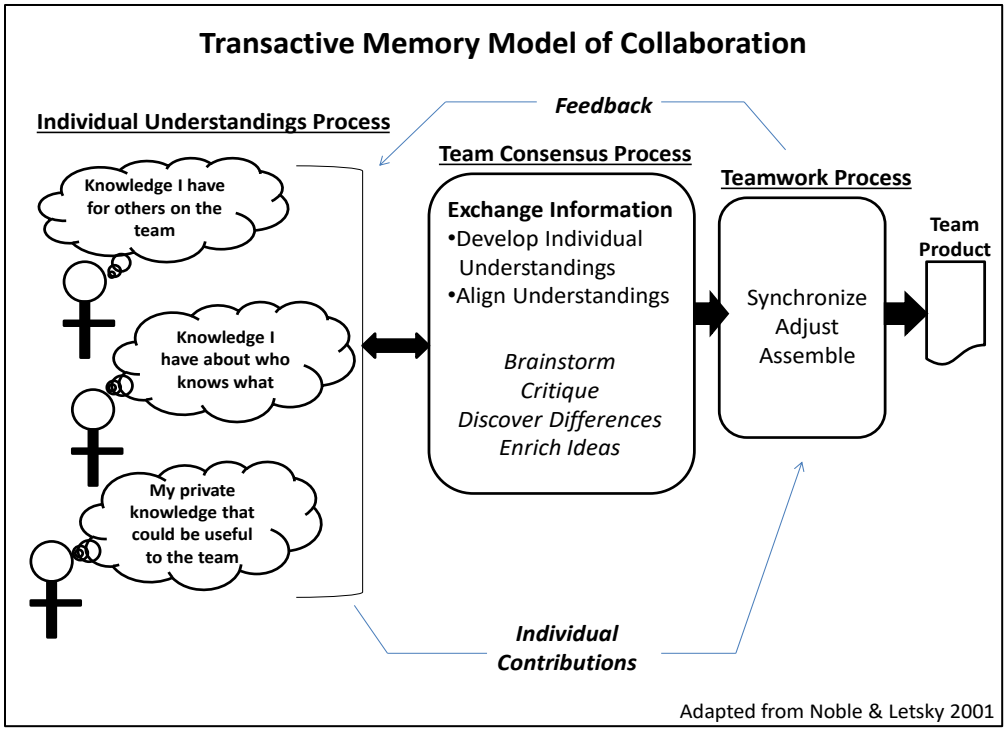


Figure 5 Transactive Memory Model of Collaboration

Three main processes compose the TMM. They are the exchange of individual understandings, forming team consensus, and conducting team work. The processes result in a team product. In the case of risk identification, the team product would be a risk list. Collaboration begins with the individual understandings. Those individual understandings include the knowledge individuals have to share with others on the team; the knowledge of who knows what on the team and how to access that other person's knowledge; private knowledge that could be useful to the team; the individual's assessment of the adequacy of their knowledge; and the individual's assessment of the direction in which the team is proceeding toward the goal (Noble and Letsky 2001). During the team consensus process, the team members exchange information to further develop their individual understandings and come into alignment. Specific techniques

that are commonly used to develop team consensus include brainstorming, critique, discovery of differences, and enrichment of others' ideas (Noble and Letsky 2001). In the third process of the TMM, the team works together to synchronize, adjust, and assemble information into the team product (Noble and Letsky 2001). As we consider creativity-based risk identification techniques, we should consider how well those techniques facilitate the collaboration processes of the TMM.

The Project Management Institute's *Practice Standard for Project Risk Management* describes creativity techniques as dependent upon the team to think creatively (PMI 2009). Those techniques include: critique, the decision balance sheet, devil's advocate, red team, scenario planning, brainstorming, and the premortem technique (Veinott, Klein, and Wiggins 2010; Klein 2007b; Foster 1978; Mateski 2009; Amer, Daim, and Jetter 2013; Wilson 2013).

2.7.1 Critique and the Decision Balance Sheet

We open our discussion with two common creativity-based approaches: the critique and the Decision Balance Sheet (DBS). All members of a team should be empowered to provide their opinion about the good and bad parts of a plan (USCG 1998). DBS is an exercise in recording the pros and cons in a plan (Prestwich, Lawton, and Conner 2003). Since critique and DBS both are a review of the good and bad parts of a plan, I discuss both here. The aim of critique is to unearth the hidden aspects of a particular course (Green 2002). Critique is an attempt to relate what exists to what one knows and does (Strausz 2011). Critique questions and contests the parameters of the

intended course (Lloyd 2012). The challenge with critique is that it does not define truth and is subject to political influences effecting those performing the critique (Green 2002).

Should we purchase a new calibration device?

	Benefits (Pros)	Costs (Cons)
Pursue the Opportunity	• Increased throughput •Higher quality	• Disposal of old device •Purchase cost •Cost of money to finance •Re-train workers
Do Not Pursue the Opportunity	• Can use capital elsewhere •No disruption to the workforce	• Live with current throughput •Customer complaints •Fall behind the Competition •Deferring the problem

Figure 6 Decision Balance Sheet

DBS can be traced as far back as Benjamin Franklin in his development of pros and cons to assess an intended course of action. Franklin described the methodology in a letter to Joseph Priestly in 1772. When confronted with a difficult decision, Franklin divided a sheet of paper in two columns. He labeled one column of the paper pros and the other column cons (Figure 6). After careful consideration and over the course of several days, he jotted down ideas in the appropriate column. When Franklin completed the list, he evaluated the balance of the two sides to come to a decision (Franklin 1772). Today DBS is used to resolve decisional conflict (Greenberg and Webster 1982). DBS has also proven to be beneficial in achieving higher degrees of motivation by its practitioners (Prochaska et al. 1994). Subjects that were considering participating in a smoking cessation program became more committed to stop smoking after performing a

DBS (Velicer et al. 1985). DBS is limited to evaluating one decision at a time and not appropriate to evaluating risks within complex system engineering or project plan. Both critique and DBS appear to be better personal decision tools than they are methods for considering risks in a plan as a collaborative team.

2.7.2 The Devil's Advocate

When considering the risks in a plan, teams tend to share information that the individuals have in common and fall victim to conformity pressures (Nemeth, Brown, and Rogers 2001). Authentic dissent stimulates a search for more information, leading to more creative solutions (Nemeth and Wachtler 1983). However, authentic dissent can have the negative consequences of marginalizing the dissenter and lowering team performance (Turner and Pratkanis 1997). The devil's advocate is an attempt to gain the benefits of information sharing from authentic dissent, while avoiding the negative consequences. The devil's advocate role plays the dissenting position, thus insulating the person from marginalization. As a risk identification technique, the devil's advocate should find all the risks in the plan (Wright and Goodwin 2009; Mateski 2009). Introducing a devil's advocate should foster constructive conflict as opposed to authentic dissent's destructive conflict (Tjosvold and Deemer 1980).

Results from Nemeth et al.'s study indicate that a team using the devil's advocate is likely to generate more original ideas compared to a control group, but those ideas tend to reinforce the original position (Nemeth, Brown, and Rogers 2001). For risk identification, the devil's advocate may have an analgesic effect, resulting in overconfidence in the plan (Schwenk 1988). That optimism would be a source of black

swan blindness. The use of a red team extends the concept of the devil's advocate from an individual antagonist to a team of antagonists.

2.7.3 Red Team

An organization may establish a red team to challenge its very own plans, programs and assumptions (Gold and Hermann 2003). Like the devil's advocate, the red team takes a contrarian position, adversarial or competitor role. Red teams have their roots in military war games but the business world frequently uses red teams to simulate competition and to test information systems' security (Mateski 2009). Red teams should identify risks in a plan by pointing out faulty assumptions or factors the enterprise has not addressed (Craig 2007). In the wake of the 9/11 attacks, Congress legislated the use of red teams in the *2004 Intelligence Reform and Terrorism Prevention Act* and the *2006 Homeland Security Authorization Ac*t to counter what they perceived to be a lack of imagination within the intelligence and law enforcement communities (Mateski 2009). Red teams do have limitations.

A red team requires training in the organization's plan and a sense of the adversary's capabilities. The organization must be open to the red team's criticism, or there will be a natural resistance to view the red team as simply picking at the plan (Malone and Schaupp 2002). A red team also requires an established set of rules and another team (often called the white team) to enforce the rules (Malone and Schaupp 2002). Red teams can be quite effective at getting an organization to think outside their traditional frames of reference but they are resource intensive (James and Drown 2008). Organizational leadership can manipulate red team findings to support a desired agenda.

For example, The Bush Administration used a red team's findings that the Hussein regime's stockpiling of aluminum tubes indicated a nuclear program as partial justification for invading Iraq in 2003 (Conway 2012). An alternative to the adversarial approaches of the devil's advocate and a red team is *scenario planning*.

2.7.4 Scenario Planning

Scenario planning explores alternative plausible futures that could jeopardize a plan (Susskind 2012). It uses current information regarding external forces and drivers to develop scenarios about the future. The organization evaluates the risks in a plan against each scenario (Ralston and Wilson 2006). The organization must then expand each scenario to account for plans and decisions already proposed or subsequently developed (Susskind 2012).

Susskind recommends an eight step process to scenario planning: (1) identify the focal issues or decisions (2) identify the driving forces impacting the focal issues or decisions (3) rank the driving forces by their level of importance and uncertainty (4) define the logic for the scenario (5) analyze the full array of scenarios (6) identify the indicators of change to monitor over time (7) assess the impacts of the focal issues or decisions under the different scenarios (8) evaluate alternative strategies (Susskind 2012). Ralston and Wilson suggest an eighteen-step process for building and using scenarios. Those eighteen steps cover four phases that focus on getting buy-in for using scenario planning; understanding the dynamics that will shape the future; creating the scenarios; and using the scenarios to aid decisions (Ralston and Wilson 2006).

SRI International (SRI) published a model for developing specific scenarios (Ralston and Wilson 2006). Each scenario begins with deciding on the key decision. The organization then considers the interaction of the driving forces of the future and the level of uncertainty about the driving forces to build the axes of the future and the scenario logic. From the axes of uncertainty and scenario logic, the organization builds the scenarios that yield decision implications. The implications then cause the organization to reconsider its decisions. The development and analysis of multiple scenarios should cover as much of the uncertainty envelope as possible (Ralston and Wilson 2006). Figure 7 combines the steps to develop and use scenarios with the SRI model and the uncertainty envelope.

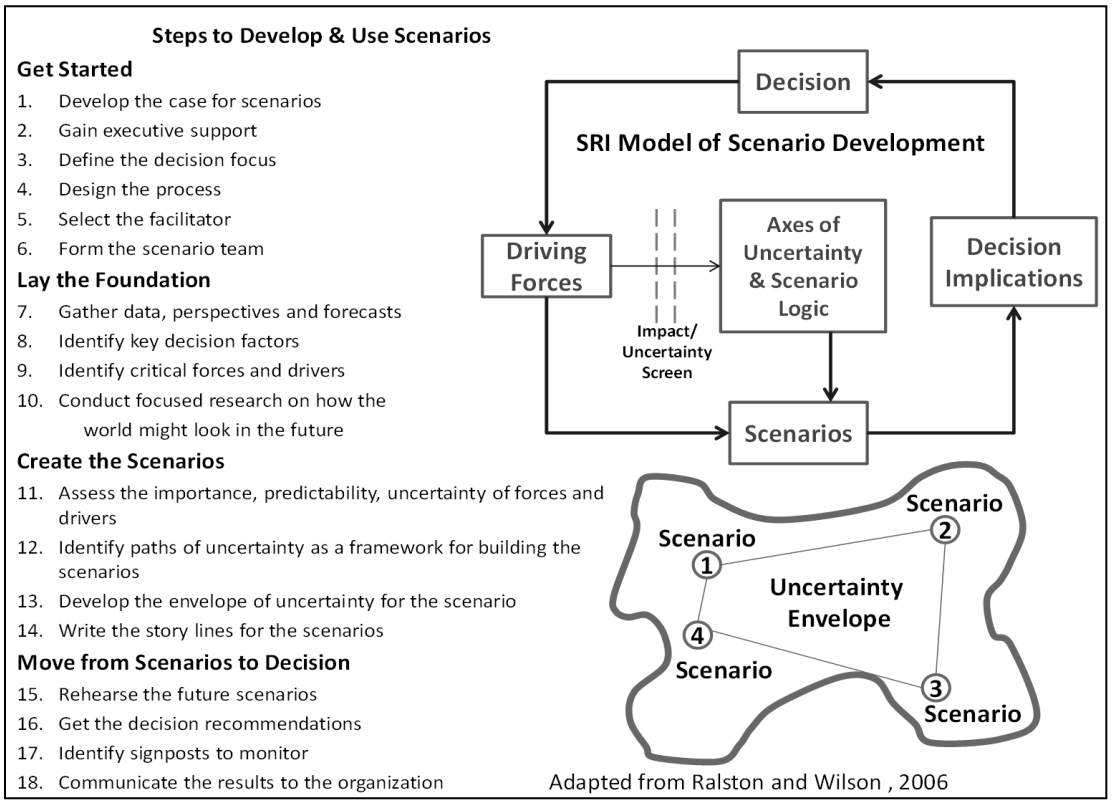


Figure 7 Scenario Planning Steps, SRI Model, and Envelope of Uncertainty

The intention of the scenario planning process is to provide some clarity on the path the organization should take relative to the plan (Konno, Nonaka, and Ogilvy 2014). Successful scenario planning depends on the organization's understanding of the current forces and drivers (Amer, Daim, and Jetter 2013). However, the organization may get so involved in the detailed analysis of the scenarios that they succumb to analysis paralysis. In practice, little effort is generally made to integrate analyses across scenarios. The integrated analyses is often left to a qualitative debate between the stakeholders (French 2015). Once the organization has evaluated the plan against future scenarios, bringing the analysis together across scenarios requires debate between stakeholders, experts and the decision maker (French 2015).

Scenario planning has been used in the technology development context. Ralston and Wilson consulted with a paper company that was considering whether or not to invest $600-million in developing a new technology (Ralston and Wilson 2006). The company used scenario planning to consider the risks in making the investment specifically because they would have to live with the new technology for at least 30 years. The scenarios included uncertainties around alternative technologies, changes in consumer values, and macroeconomic conditions. The company found that only the most optimistic scenarios provided the necessary return on investment. Therefore, the executives decided to significantly scale down the scope of the project (Ralston and Wilson 2006). In this case, scenario planning was very effective but also required significant time, training, and resources. Unlike scenario planning, brainstorming is quick and requires very little infrastructure, making brainstorming a popular and pervasive risk identification technique (Gobble 2014a).

2.7.5 Brainstorming

Brainstorming is a rapid, idea generation session. The session is focused on quantity and suspending criticism (Gobble 2014b). Creativity theorist Alex Osborn proposed the term *brainstorming* in 1953. His original concept was to gather about a dozen people in a room with differing levels of expertise and give them a single question to answer. By suspending judgment and generating a large quantity of ideas, the group can access their collective creativity (Gobble 2014a).

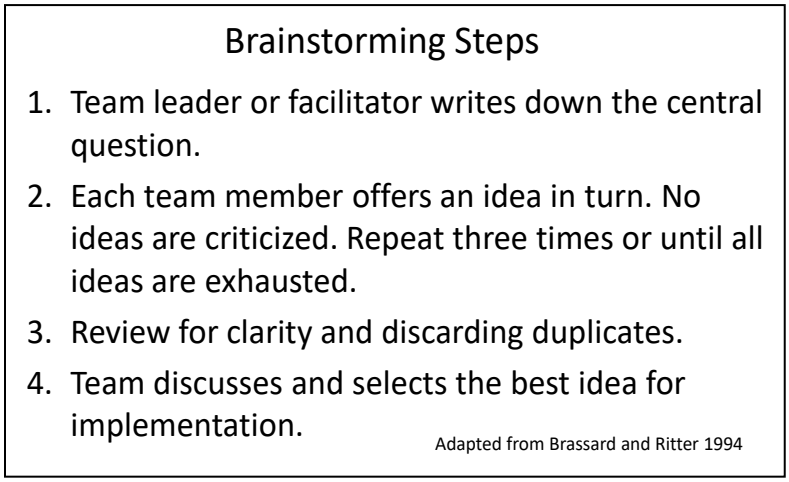


Figure 8 Brainstorming Steps

The simplicity of the method partly explains its widespread use. There are four simple steps to brainstorming listed in Figure 8 (Brassard and Ritter 1994). Wilson described brainstorming as the repeated search for visions in associative memory (Wilson 2013). Those visions create thought chains to provide information related to the question at-hand (Zhao and Hou 2010). That information is then shared and collaborated within the team and across teams. One of the world's most successful and innovative companies

uses brainstorming as an idea generation tool. Michael Lopp, senior engineering manager at Apple described how Apple's design process works to the *2008 South by Southwest Conference*. Central to Apple's process are paired design meetings. Paired design meetings, as the name suggests, occur in two parts. The first part is a brainstorming session in which designers are encouraged to think of ideas freely, without constraints. The wild ideas are passed to a second team that considers the practical implementation of the ideas generated by the first team (Walters 2008). Beyond design, brainstorming is also a popular risk identification technique.

The *Enterprise Risk Management Guide* lists brainstorming at the top of its recommended risk identification methods (Shenkir and Walker 2007). It describes the use of brainstorming sessions to enable participants to express the risks that may be known by some members of the group but not to others (Shenkir and Walker 2007). The guide encourages the use of cross-functional teams to shed light on how risks impact business units differently (Shenkir and Walker 2007). The *Department of Defense Risk Management Guide* also lists brainstorming as a primary means of identifying risks (OUSD(AT&L) 2006). Carpenter studied the effectiveness of brainstorming as risk identification technique.

Carpenter conducted an experiment involving the risks identified by audit teams. She compared the number of risk ideas and the quality of those risk ideas generated by audit teams using brainstorming against a nominal group. The experiment included 120 risk auditors. Her results indicated that while the brainstorming teams generated less risk ideas compared to the nominal group, the quality of the risk ideas was greater (Carpenter 2014). After the brainstorming session, the auditors that used brainstorming found more

fraud risks during an audit than those in the nominal group. The increased discovery of risks suggests that brainstorming heightened the situational awareness of the participants to fraud risks as they conducted the audit (Carpenter 2014).

Brainstorming as an idea generation process, including the generation of risk ideas, does have its limitations (Figure 9). Reeves et al. noted a checklist mentality as a limitation to brainstorming within the context of systems engineering at the National Aeronautics and Space Administration (NASA) (Reeves et al. 2013). In other words, even though brainstorming should be a freewheeling activity, systems engineers tended to constrain their brainstorming activities to a predefined set of categories (Reeves 2012).

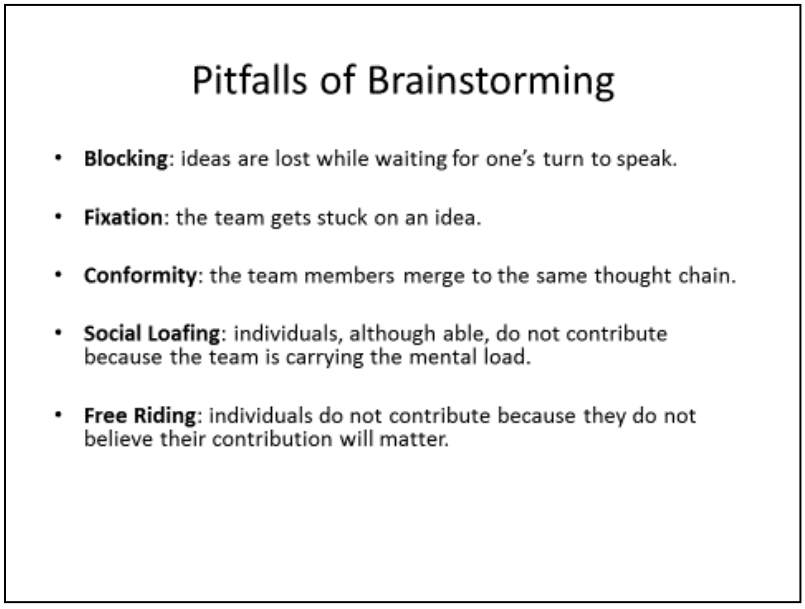


Figure 9 Pitfalls of Brainstorming

Because only one person can communicate an idea at a time, others must wait their turn and may lose ideas while waiting (Diehl and Stroebe 1987). This phenomenon is referred to as *blocking* (McGlynn et al. 2004). The brainstorming session may require a strong facilitator that is cognizant of the social and physical environment to overcome this limitation (Wilson 2013).

There is also documented evidence that brainstorming participants get stuck on each other's ideas (Paulus et al. 2015; V. Brown and Paulus 2015). This phenomenon is referred to as *fixation* or *conformity*. Kohn and Smith conducted three experiments and concluded that brainstorming teams became fixated on each others' ideas even while the number of unique ideas continued to increase (Kohn and Smith 2011). In other words, the team continued to generate unique ideas, but those ideas were very similar to ideas that had already been offered by others on the team. Fixation or conformity may be the result of participants fearing that the team will negatively evaluate their ideas (Haddou, Camilleri, and Zaraté 2014). This evaluation apprehension reduces the diversity of ideas offered. Since diversity of ideas is one of the goals of brainstorming, fixation is a significant limitation (V. Brown et al. 1998). Another problem that constrains the diversity of ideas while brainstorming is *social loafing*.

There is extensive literature on the phenomenon of social loafing that shows people tend to exert less effort while performing a task as a group than if they perform the task individually (Reinig, Briggs, and Nunamaker 2007). While brainstorming as a team, some team members, although fully qualified to contribute, will reduce their efforts (Diehl and Stroebe 1987). A variation of social loafing is *free riding* (Mullen, Johnson, and Salas 1991). In this form of social loafing, a team member believes that their

contribution is not needed to achieve team success (Haddou, Camilleri, and Zaraté 2014). Researchers have studied methods to avoid or mitigate the limitations of blocking, fixation, social loafing, and free riding in brainstorming.

Electronic brainstorming can address some of the limitations in traditional brainstorming. In electronic brainstorming the participants simultaneously type their ideas into a computer for immediate display on the other teammates' screens (Gallupe et al. 1992). In experiments of different sized groups, several studies found that electronic brainstorming reduced the effects of blocking and evaluation apprehension on group performance, particularly for large groups (Gallupe et al. 1992; Paulus et al. 2015; V. Brown et al. 1998). These electronic brainstorming tools have promise but require training and additional resources to implement. Gary Klein has offered the *premortem* as a risk identification technique that taps into a team's creativity, while avoiding the brainstorming pitfalls (Klein 2004).

2.7.6 The Premortem Technique

The *premortem* as a risk identification technique takes advantage of team collaboration while avoiding some of the drawbacks of brainstorming (Klein 2004; Klein 2009). It is the theoretical opposite of a postmortem. In a medical setting, a postmortem is an investigation into why the patient died (Klein 2007b). The information from the investigation benefits doctors, future patients, and perhaps the family of the deceased. The postmortem provides no benefit to the deceased. The *premortem*, on the other hand, makes the assumption that the patient has died and searches for answers to how the death

could have occurred (Klein 2007b). In theory, the patient should benefit from the additional insight gained from the premortem.

The premortem is a deductive method while brainstorming and the other risk identification techniques discussed earlier are inductive methods (Figure-10). In deductive logic, the claim is made that conclusions flow from the certainty of the premises (Copi and Cohen 2009). Deductive arguments make the claim of certainty and can be proven valid or invalid. Inductive logic cannot make a claim of certainty; nothing is without doubt (Copi and Cohen 2009). Induction requires the incorporation of probabilities, which may only be of moderate value (Copi and Cohen 2009). People are just not very good at assigning probabilities to events, which is a limitation of inductive techniques (Mayseless and Kruglanski 1987).

The inductive technique of brainstorming for risk identification focuses on what *could* go wrong. Recognizing what could go wrong requires the application of probabilities in inductive logic. In the premortem, the assumption is the plan failed with absolute certainty. The premortem technique's use of deductive logic removes the challenges of assigning probabilities. The focus is on what *did* go wrong with the plan (Klein 2004). Normally, people explain the past and predict the future. When a future event is uncertain, people focus their information searching and sharing on explaining the uncertainty rather than on the outcome. By making a future event certain, people become more focused on explaining the outcome (Mitchell, Edward Russo, and Pennington 1989).

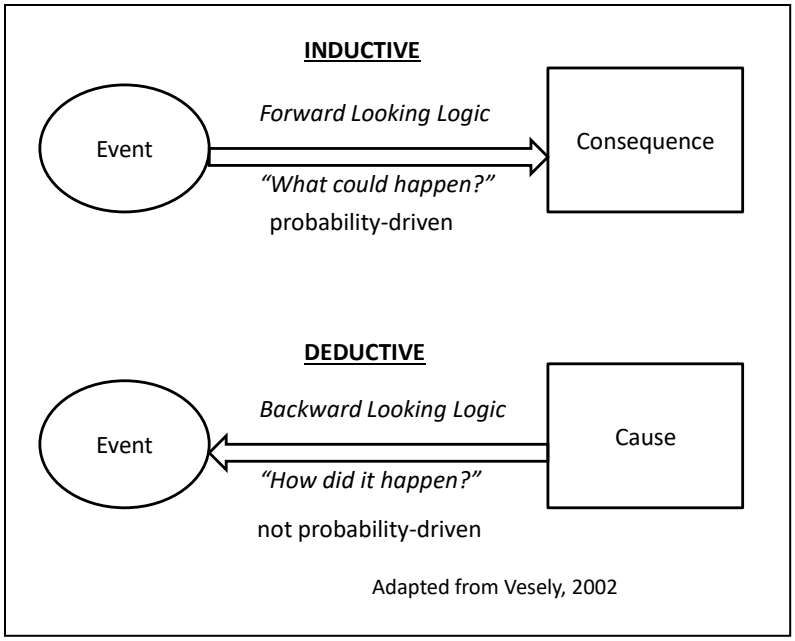


Figure 10 Inductive versus Deductive Thinking

In their 1989 study on prospective hindsight, Mitchell et al. found that by declaring the failure of an event to be certain, subjects generated more reasons for the failure of that event. They theorized that people think differently about certain events than uncertain events (Mitchell, Edward Russo, and Pennington 1989). Conversely, their subjects found it more difficult to generate explanations for uncertain events than sure ones (Mitchell, Edward Russo, and Pennington 1989). The theory is that when a future event is uncertain, which is generally the case, people will focus their mental energy on processing the uncertainty (Veinott, Klein, and Wiggins 2010). Prospective hindsight can aid risk identification by facilitating the collaborative construction of event chains for a hypothesized outcome that has been declared a certainty. The premortem technique takes advantage of prospective hindsight to identify risks in a plan.

Klein suggests that the premortem is best done at the outset of a project (Klein 2007b). He recommends the following steps in conducting a premortem:

(1) Get familiar with the plan.

(2) Imagine that the plan has failed.

(3) Individually, write down reasons why the plan failed.

(4) Present and discuss the reasons for the failure.

(5) After reviewing the list, the team determines how to improve the plan.

(Klein 2007a)

The technique of self-critiquing is not new. However, the premortem is different. When a team critiques its own plan, there is a subtle expectation that the team will not identify any showstoppers (Klein 2004). The team is emotionally invested in the plan and is not motivated to highlight the plan's fatal flaws. By declaring the plan as a certain failure, the team can achieve a higher level of candor. This increased candor allows the team to exchange perspectives and get risks on the table. The premortem technique builds on the theory of prospective hindsight and Klein's own work in the area of naturalistic decision making.

Naturalistic decision making attempts to describe how people make decisions in real world settings (Klein 2008). Studies in naturalistic decision making showed that intuition plays a primary role in making decisions as opposed to deliberate reasoning (Kahneman 2003). The intuition-based, cognitive system is a process that has the following characteristics: fast, parallel, automatic, effortless, associative, slow-learning

and emotional (Kahneman 2003). Klein concluded that naturalistic decision making depended upon an individual's experience to rapidly and effortlessly categorize situations to make effective decisions (Klein 2008). In other words, people draw from their past experiences to generate heuristics that may apply to the current situation. They then attempt to apply the best heuristic based on intuition. This is contrary to the traditional belief that people generate multiple courses of action, evaluate those courses of action and select the best possible course of action based on reasoning (Kahneman 2003). The premortem technique relies on the naturalistic decision making processes, based on expertise, heuristics, and intuition to generate risk ideas.

Klein describes a premortem session involving a military air campaign planning software development team. In the course of the premortem session, the team came to the realization that the software might take hours to run on the current hardware when the user needed information in seconds. The issue had not come-up in a series of prior meetings. After discussing the premortem findings, the team changed the plan to include a software shortcut that a member of the team had been reluctant to mention prior to the premortem session (Klein 2007b). After Klein published the premortem technique in the Harvard Business Review in 2007, several organizations and fields began adopting and adapting the method for risk identification.

Within the field of corporate finance, *TechAngel Funds* uses the premortem technique to ensure the team has conducted a full due diligence when considering a venture capital investment opportunity (Hagel 2012). In their implementation of the premortem technique, *TechAngel Funds* adds the following notes to Klein's process:

(1) Have an objective facilitator lead the premortem.

(2) Have the people who developed the plan as part of the premortem

(3) Include people from all levels of the organization.

(4) Include contrarians, those individuals that are against the plan in question. (Hagel 2012)

The *Otto Beisheim School of Management* uses the premortem technique in its Ph.D. program. The program follows the German-speaking model in which one-third of the coursework is seminars and the remaining two-thirds in developing three or four topically-related papers. The program director, Lutz Kaufmann conducts a premortem analysis with doctoral candidates early in the process to review their research plan. Unlike the typical critiquing session where the candidate is asked what might go wrong, the premortem operates on the assumption that the candidate's research project has failed (Kaufmann 2011). Kaufmann reported the key benefit of the premortem technique is a change in the doctoral candidate's mindset. The premortem forces the candidate to switch from defending every step in their research plan to a mode in which they generate a variety of reasons for the hypothetical failure. The premortem enables Kaufmann's doctoral candidates to creatively improve their research plan at the outset of the project (Kaufmann 2011).

The premortem technique has been used within the cybersecurity field. Faily et al. found that the premortem was a useful tool for managers to keep in touch with the reality of the project (Faily, Parkin, and Lyle 2014). They found that the premortem technique allowed the non-technical stakeholders to engage in the design discussion. They pointed to the fact that the premortem technique was able to capture the

stakeholder's imagination. This allowed those stakeholders to provide thought provoking feedback rather than having to rationally derive potential risks. The researchers also found that the premortem provided a framework to discuss contextual information without disclosing sensitive specifics.

Chen et al. studied the use of the premortem technique in Computer Security Incident Response Teams (CSIRTs). CSIRTs protect an organization's information technology infrastructure by working collaboratively to address cybersecurity threats. They often operate as a multi-team system, meaning two or more teams working closely together to achieve a goal. Being able to communicate and coordinate across multiple team boundaries simultaneously makes collaboration a very important skill. Systems engineering shares this multi-team characteristic with CSIRTs. Therefore, the Chen et al. study is very relevant to systems engineering and engineering management. They studied individual and team tasks in a number of CSIRTs in both the public and private sectors. Their research method involved interviews, structured observation, and surveys. One of their findings suggested that the combined individual and collaborative nature of CSIRT work often requires an individual that identifies a potential problem to make a decision on whether to involve other members of the multi-team (T. R. Chen et al. 2014). Systems engineering faces similar challenges (INCOSE 2014). As a solution to this challenge in CSIRTs, Chen et al. found the use of the premortem technique very effective.

A CSIRT premortem is a mental exercise in which the team assesses hypothetical employee mistakes that may lead to an inappropriate incident response (T. R. Chen et al. 2014). Through this mental exercise, the team gains a better understanding of the CSIRT tasks appropriate for the incident in question (T. R. Chen et al. 2014). Chen et al.'s

research went beyond the anecdotal reports of the premortem technique mentioned earlier by collecting empirical data. However, the CSIRT study did not conduct experiments of the premortem technique. Veinott et al. conducted the only known experiments of the premortem technique.

Veinott et al. studied the effectiveness of the premortem technique used within a crisis management planning scenario. They investigated the benefits of the premortem technique to counter overconfidence in a university's flu epidemic response plan (Veinott, Klein, and Wiggins 2010). Systems engineers and project managers often face the challenge of overconfidence, especially in the case of system upgrades (Kerr, Phaal, and Probert 2008). Mitigating overconfidence may allow teams to generate more realistic cost and schedule estimates and open the team's field of view to risks.

Overconfidence, which includes hubris and arrogance, often leads to ignoring the risks involved in complex systems development projects. This was the case in the Royal Australian Navy *Collins* class submarine program. Australia lacked any experience building submarines but decided that the *Collins* class submarines would be built in Australia. Overconfidence led the *Collins* development team to substantially underestimate many risks and while choosing to ignore others. Although it became operationally effective with the help of the United States Navy, the *Collins* became the largest, most expensive Australian military project primarily due to overconfidence (Yule and Woolner 2008). According to Abkowitz, overconfidence also contributed to a number of well documented black swan events such as the *Hyatt Regency* collapse, the *Union Carbide* gas leak in Bhopal, the Chernobyl reactor meltdown, the *Exxon Valdez* oil

spill, the attack on the *USS Cole*, the sinking of the *Edmund Fitzgerald*, and both the *Challenger* and *Columbia* disasters (Abkowitz 2008).

In studying the premortem's ability to counter overconfidence, Veinott et al. performed experiments to measure the confidence levels of 178 paid, undergraduate students before and after applying one of four experimental treatments (Figure 11). The experimental treatments were: critique, premortem, pro-con, and con. They found that the premortem technique reduced the subject's confidence in the plan more than the other treatments. After considering changes to the plan based on the applied treatment, the premortem technique increased confidence in the revised plan more than the other treatments. The researchers also measured each subject's understanding of the plan after applying their assigned treatment. They intended to show that the premortem technique improved the subject's understanding of the plan more than the other treatments. Their findings in this area were partially supported. An analysis of variance between the treatments and the baseline did not indicate a statistically significant advantage of using the premortem technique to increase subject understanding. In a pairwise comparison, all treatments showed an increase in subject understanding of the plan except for the critique treatment (Veinott, Klein, and Wiggins 2010). Critique was also the lowest performing treatment in reducing confidence in the plan. This led the researchers to postulate that an initial reduction in plan confidence may be necessary to achieve an improved understanding of the plan (Veinott, Klein, and Wiggins 2010).

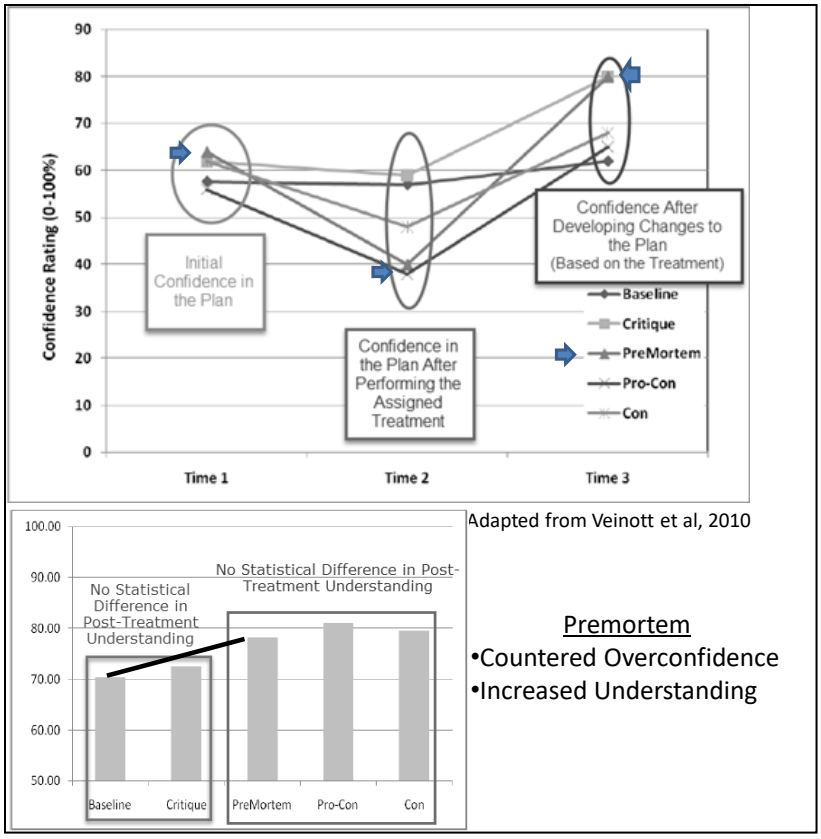


Figure 11 Summary of Veinott et al. Findings

Veinott et al. acknowledged a number of limitations in their study. The participants were undergraduate students, not experienced planners and leaders (Veinott, Klein, and Wiggins 2010). The participants reviewed a prepared plan, not of their making, therefore they likely had no emotional investment in their performance (Veinott, Klein, and Wiggins 2010).

I identified additional limitations in their study. As noted earlier, critique and DBS (i.e., pros/cons) are very similar in process and outcome. Given three of the four treatments in the Veinott experiments (critique, pros/cons, and cons) are similar

techniques, there is the possibility that those three treatments diluted each other's performance when compared to the very different premortem technique. Critique, pros/cons, and cons are good for making individual decisions. They are less suitable for team evaluation of complex plans. Finally, the results of any quality intervention (before and after study) can be influenced by the impracticality of obtaining adequate concurrent controls (ACP 2014). Despite these flaws, Veinott conducted a significant experiment-based study of the premortem technique from which to build upon.

To summarize the known literature regarding the premortem technique, I have described six cases in which the technique was applied in various fields ranging from systems engineering to crisis action planning (Table 1). Four of the six cases are anecdotal reports of the capabilities and limitations of the premortem in the fields of systems engineering, corporate finance, education, and cybersecurity. The two remaining cases are both empirical studies in the CSIRT and crisis action planning fields. However, only the Veinott study of the premortem technique (in a crisis action planning context) is an empirical study performed as a controlled experiment. One question regarding the premortem technique must be answered. Is the premortem technique nothing more than a management application of a technique widely used in systems engineering known as fault tree analysis?

Field	Application	Result
Systems Engineering	Mission Planning Software Development (anecdotal)	Premortem allowed the team to identify a previously undiscovered risk and an previously undiscussed solution.
Corporate Finance	Venture Capital Due Diligence (anecdotal)	Premortem helped the team complete a more through due diligence of investment opportunities.
Education	Doctoral Research Proposal Defense (anecdotal)	Premortem allowed candidates to switch from defending every position to generating a variety of reasons for failure and appropriate changes to the research proposal.
Cybersecurity	Project Review (anecdotal)	Premortem facilitated thought provoking feedback on risks from non-technical stakeholders
Computer Security Incident Response	Response Training and Preparation (empirical but non-experimental)	Premortem facilitated multi-team communication of risk perspectives
Crisis Action Planning	Influenza Response Plan (empirical and experimental)	Premortem reduced overconfidence in the plan.

Table 1 Summary of Known Premortem Technique Application Cases

2.8 Premortem and Fault Tree Analysis (FTA) Compared

FTA is a systematic and stylized deductive process that starts with defining an undesired event (Howat 2008). That event is further decomposed into immediate causes until the basic causes for the top-level undesired event emerge (Vesely 2002). FTA involves graphical enumeration and analysis based on known information including failure modes for all components and sets of subsystems (Tumer, Stone, and Roberts 2003). The technique depends on historical information such as maintenance manuals, component and subsystem failure mode effects analysis, and accident reports (Tumer, Stone, and Roberts 2003). These are all valid sources of risk information but do not help us with the early detection of the unprecedented black swan (Taleb 2003). In the case of

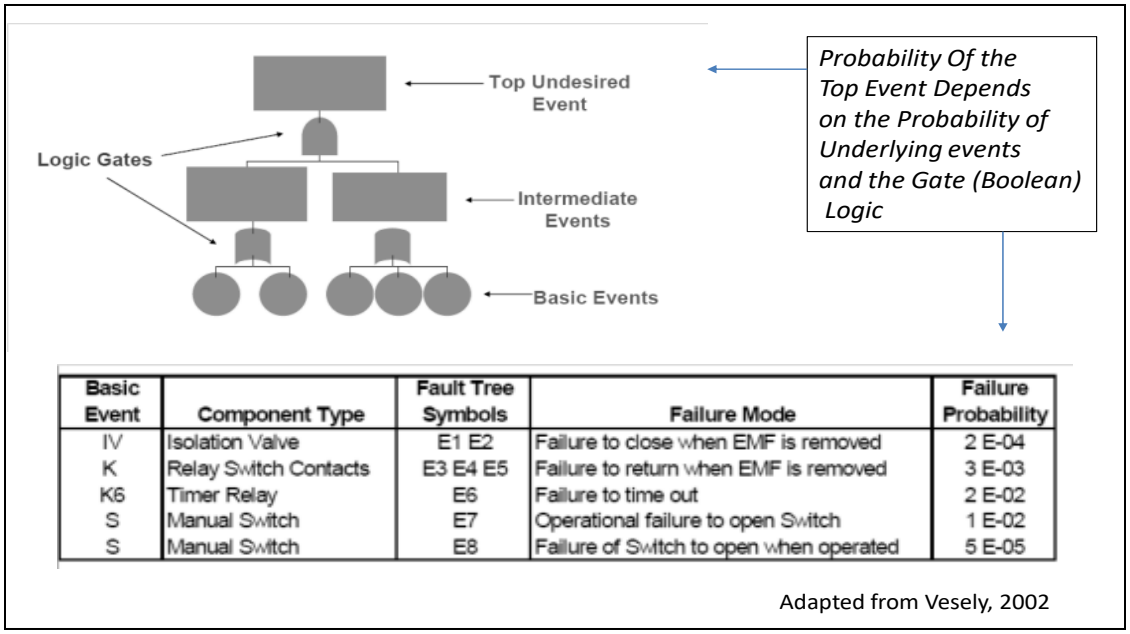
the *Columbia* disaster, the FTA considered all known wing load cases (NASA 2003). Foam insulation broke off of the external fuel tank and damaged *Columbia's* wing. This caused the actual in-flight wing loads to be outside of the safety margins derived from the FTA (NASA 2003).

The premortem and FTA are both deductive techniques, but their approach to identifying risks differ. The premortem approach relies on mental simulation and group discussion (Klein 2009). Unlike the FTA approach, the premortem does not depend on historical information or subject matter expertise. The premortem uses the team member experiences to tap into the power of the team's collective intuition (Klein 2004). The premortem approach embraces the unprecedented while the FTA approach is bound to precedent.

FTA and the premortem technique differ in purpose. A team performs an FTA to gain insight into critical aspects of a system's design with the intent of achieving risk-free operation (Tumer, Stone, and Roberts 2003). The premortem technique's purpose is to overcome team dynamics such as deliberate ignorance, group-think and overconfidence to find the key vulnerabilities in a plan (Klein 2015). The premortem technique can function in a world of "unknown, unknowns," seeking out the black swan. FTA, on the other hand, is relevant only in the world of "knowns" and "known, unknowns" in which we can apply probabilities. NASA experienced this limitation of FTA during the Apollo program. NASA had been aware of and fully analyzed the failure mode of a fire in the capsule while the capsule was in flight (Stamatelatos and Caraballo 2002). The *Apollo 1* fire is a case in which the FTA approach left NASA blind to the risk of a fire in the capsule's pure oxygen environment *while the capsule was on the launch pad* (Massie

2011). The *Apollo 1* case is an example of how an FTA can lock a team into their existing frame of reference.

The premortem and FTA differ in their artifacts as well. The FTA delivers a fault tree diagram (Figure 12), which is a graphical representation of the decomposed failure event with the failure event expressed in terms of an occurrence probability or unique paths through the system leading to the failure event called *minimum cut-sets* (Khakzad, Khan, and Amyotte 2011). To deliver this artifact, FTA assumes that all component and subsystem interactions are known and can be appropriately represented in the diagram. As systems become more complex, expansive, and interdependent, their interactions will be hidden from view and thus more difficult to represent in a fault tree diagram. An artifact from a premortem is a comprehensive list of risks that includes the unprecedented and improbable events that one will not find on a fault tree diagram. The premortem also generates actions to strengthen the plan. The FTA does not explicitly generate changes to the systems engineering plan to address the failure event. The premortem facilitates a shared understanding of failure events from each participant's perspective. FTA provides a single view of the failure relationships but does not explicitly address the contributing engineers' differences of perspectives.



Basic Event	Component Type	Fault Tree Symbols	Failure Mode	Failure Probability
IV	Isolation Valve	E1 E2	Failure to close when EMF is removed	2 E-04
K	Relay Switch Contacts	E3 E4 E5	Failure to return when EMF is removed	3 E-03
K6	Timer Relay	E6	Failure to time out	2 E-02
S	Manual Switch	E7	Operational failure to open Switch	1 E-02
S	Manual Switch	E8	Failure of Switch to open when operated	5 E-05

Adapted from Vesely, 2002

Figure 12 Example of a Fault Tree

The strength of an FTA is its systematic analysis of failures that lead to minimum cut-sets. The minimum cut-sets allow the team to focus on failure paths for additional analysis and decision-making. The strength of the premortem technique is that it has no conceptual boundaries and acknowledges that we have incomplete information about our systems, their complexities, and hidden interactions. The lack of conceptual boundaries and the acknowledgement that the team has incomplete information mitigates black swan blindness and encourages discussions about the unprecedented.

The premortem does have its limitations, in that the technique does not assign probabilities to the failures it generates. Without probabilities, it is difficult to discuss changes to the plan in both engineering and economic terms. An economic discussion requires risk management actions to be monetized. Monetization typically involves

estimating the cost of the change to the plan by the probability of the related risk occurring. The premortem proposes actions to strengthen the plan without probabilities, making it difficult to perform a cost-benefit analysis for the changes proposed to strengthen the plan. Another limitation of the premortem compared to FTA is that the premortem should have a trained facilitator (Bartram 2015). This is a significant limitation if the quality of a premortem session is directly correlated with the capabilities of a facilitator. FTA does not require a facilitator, but it must be a planned and coordinated activity. A final limitation of the premortem is that it can cause an organization to become too risk averse. This could be detrimental in organizations that already have a culture of risk avoidance (Eriksen 2015).

In summary, FTA requires a complete understanding of the system of interest (Howat 2008), but complete understanding of complex, interconnected systems may not be possible and may lead to overconfidence (Walker, Shenkir, and Barton 2014). That, in turn, leaves us vulnerable to black swan risks (Aven 2015). FTA's utility depends upon an accurate assignment of probabilities from component reliability history or subject matter expertise (Massie 2011). Relying on the historical measurements is less helpful as systems become more complex and interconnected. The use of expert opinion is a valid method for obtaining probabilities, but the results will depend upon getting the right mix of experts (Merrick 2008). If you change the mix of experts, you are likely to get a different set of probabilities. FTA requires that the boundaries of the system be clearly defined (Stamatelatos and Caraballo 2002). With globalization and highly networked systems of systems, it is increasingly difficult to define where one system ends and another system begins (Taleb, Goldstein, and Spitznagel 2009; INCOSE 2014).

FTA shares the deductive characteristic of the premortem technique, but the premortem is not merely a project management application of FTA. The premortem has its limitations but is better suited than FTA to catch black swans in a complex environment.

2.9 Other Techniques for Creativity- Based Risk Identification

There are many other techniques for creativity-based risk identification. In the following section, I will describe a number of these techniques based on their similarity to brainstorming, premortem, and FTA.

Independent scholar and futurist Joel Barker developed the *implications wheel* as a tool to examine the primary, secondary, and tertiary consequences of actions (Barker 2000; Farrington, Henson, and Crews 2012). In Figure 13, the initiating event is displayed in the center of the implications wheel as the introduction of a *Capstone Exercise* into a program management training course. Both the positive and negative implications of the Capstone Exercise are displayed. The tertiary-level consequences are also part of the implications wheel diagram. The technique has become popular with futurists examining the positive and negative implications of technology (Wagner 2005). The implications wheel is an inductive, forward looking application of brainstorming (Farrington, Crews, and Blenkle 2013). As a risk identification tool, the implications wheel attempts to discover the unintended consequences of an action (Barker 2000). The implications wheel appears to be a special application of the group brainstorming technique (Farrington, Henson, and Crews 2012).

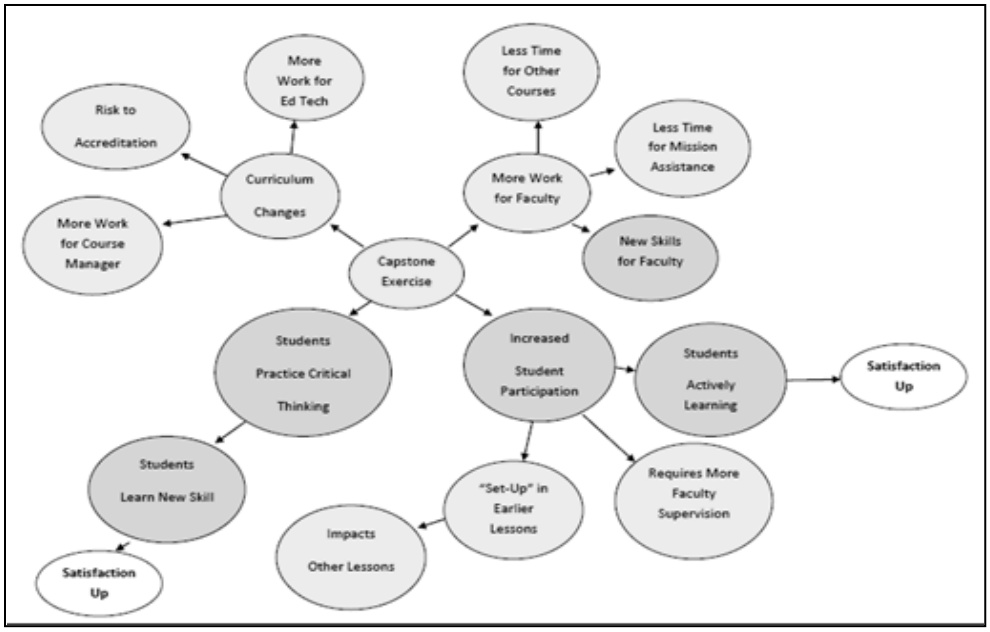


Figure 13 Implications Wheel

Teams could also use *The Six Thinking Hats* technique (Figure 14) as part of creativity-based risk identification. Edward DeBono designed the method as an alternative to argument-based thinking (DeBono 1999). He describes the *Six Thinking Hats* as *parallel thinking*. Parallel thinking acknowledges that the brain partitions thinking into analytical, critical, constructive, emotional, divergent, and transversal modes (Varvoglis 2003). DeBono advocates using these natural partitions to facilitate creative problem solving. Using *The Six Thinking Hats* method, a team simultaneously activates the same partition of thinking to analyze a problem using six metaphorical thinking hats. If the team wanted to exercise the analytical thinking partition, all team members wear the metaphorical "white hat" that calls for an analysis of facts, assumptions, information needed, and information sources. When that analysis is

complete, the team might choose to activate the emotional partition. In that case, all members wear the metaphorical "red hat" and provide their gut reaction based on intuition. This would continue until the team had analyzed the problem from all six partitions/thinking hats. The technique overcomes common group information pathologies and provides a thorough, orthogonal analysis (Puccio et al. 2010). The "black hat" is the thinking mode that would be most helpful in creativity-based risk identification. The black hat calls for the team to think in terms of what can go wrong or the weaknesses in a plan (DeBono 1999). DeBono's technique is similar to the inductive, brainstorming method. *The Six Thinking Hats* is a framework for creative thinking with an element of risk analysis when the team uses the black hat (Sheth 2012). It is not a creativity-based tool for risk identification.

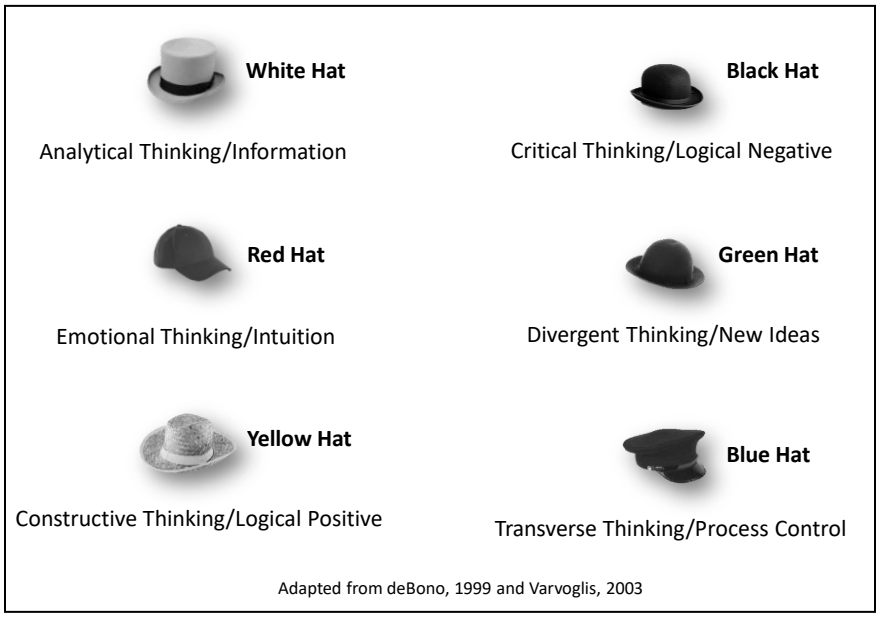


Figure 14 Six Thinking Hats

Strengths-Weaknesses-Opportunities-Threats (*SWOT*) analysis is another technique that could be applied to creativity-based risk identification. SWOT analysis came from a team of researchers at the Stanford Research Institute (SRI) between 1960 and 1970 that were looking into what had gone wrong with corporate planning (Humphrey 2005). After interviewing over 5,000 executives, the team found that part of the problem with corporate planning is the demand placed on the executive's attention. Their work led to a four part approach to analyzing the firm's overall strategy in terms of strengths, weaknesses, opportunities and threats (Osita, Onyebuchi, and Nzekwe 2014). The four parts are often displayed in a two-by-two grid (Figure 15) in which strengths and weaknesses are paired in one half of the grid because they are both views of the firm's internal situation. Opportunities and threats are paired in the other half of the grid because they reflect the firm's external environment. The weaknesses and threats quadrants of the SWOT analysis may be applicable to risk identification because both weaknesses and threats negatively impact the ability of the firm to achieve its goals (Osita, Onyebuchi, and Nzekwe 2014). Additionally, a SWOT analysis is based firmly on the "knowns" (IBM 2006). It does not push a team to consider the unexpected or unprecedented. SWOT is a very good tool for achieving better situational awareness but is not a tool for risk identification.

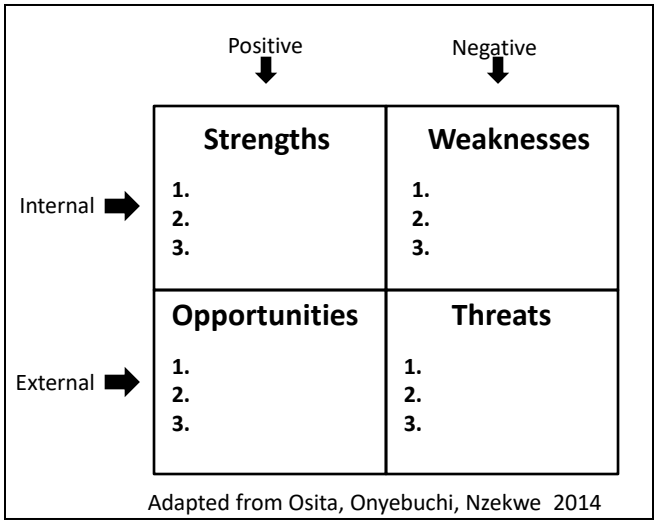


Figure 15 SWOT Analysis Grid

The previous examples of the implications wheel, Six Thinking Hats, and SWOT are creativity-based risk identification with roots in neuroscience and management. The next examples have origins in engineering. Level of Protection Analysis (LOPA) is a technique that allows a team to assess the risk associated with hazardous events by applying cause-consequence scenarios (Altabbakh et al. 2013; Howat 2008). Using the organization's risk standards, the team assesses the likelihood that the initiating event will get through the independent layers of protection to cause a catastrophe. The LOPA forces the team to consider safeguards to manage threats to the system. Like FTA, LOPA is limited to analyzing one scenario at a time. Figure 16 depicts the LOPA model for the Challenger disaster. If LOPA can model only one scenario at a time, it is difficult to apply the technique to complex and interconnected systems (Altabbakh et al. 2013).

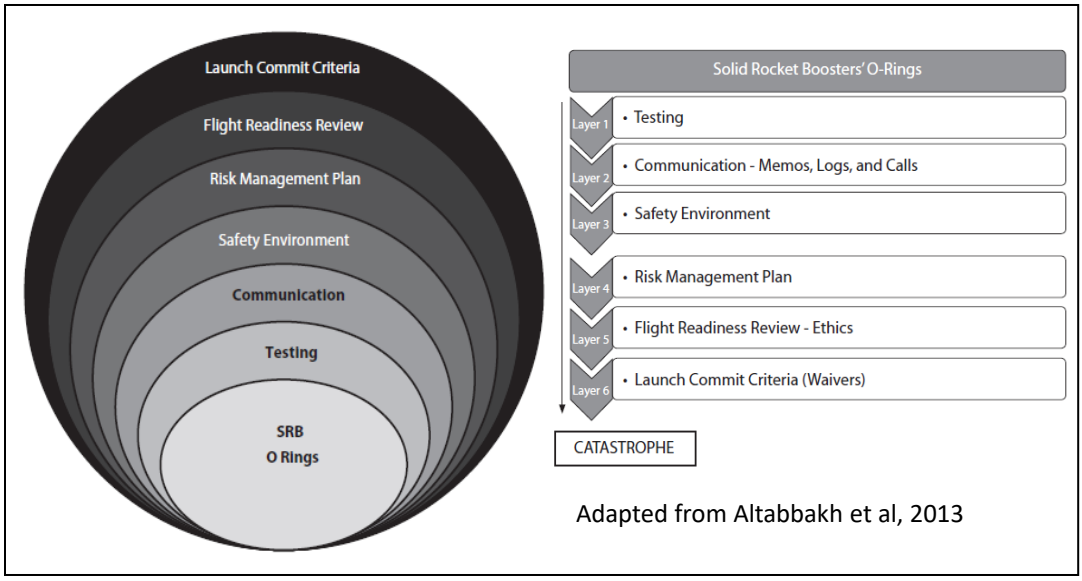


Figure 16 Layer of Protection Analysis Model of the Challenger Disaster

James Reason developed the *Swiss Cheese Model* to address the risks of accidents in complex systems (Reason 1997). He represents the scenario under consideration as stacks of Swiss cheese in which the cheese represents a layer of defense and the holes represent either latent conditions or active failures (Altabbakh et al. 2013). Latent conditions are states of the organization, such as the level of training and are difficult to detect. Active failures are unsafe acts caused by human interaction with the system (Altabbakh et al. 2013). The goal of the team is to ensure that none of the holes in the defense layers align, allowing an adverse event through the system. Figure 17 is a *Swiss Cheese Model* for a motorcycle safety program. In that model, the system has four layers of defense beginning with driver's education and ending with motorcycle skills testing. As long as none of the latent conditions such as poor quality of training or variations in compliance line-up for all four layers, the system will be robust against failures. The

Swiss Cheese Model is a useful method of visualizing risk vulnerability. However, in the real world the holes are constantly in motion. Another limitation of the *Swiss Cheese Model*, like LOPA and FTA, is that the model is based on what is known and visible to the team.

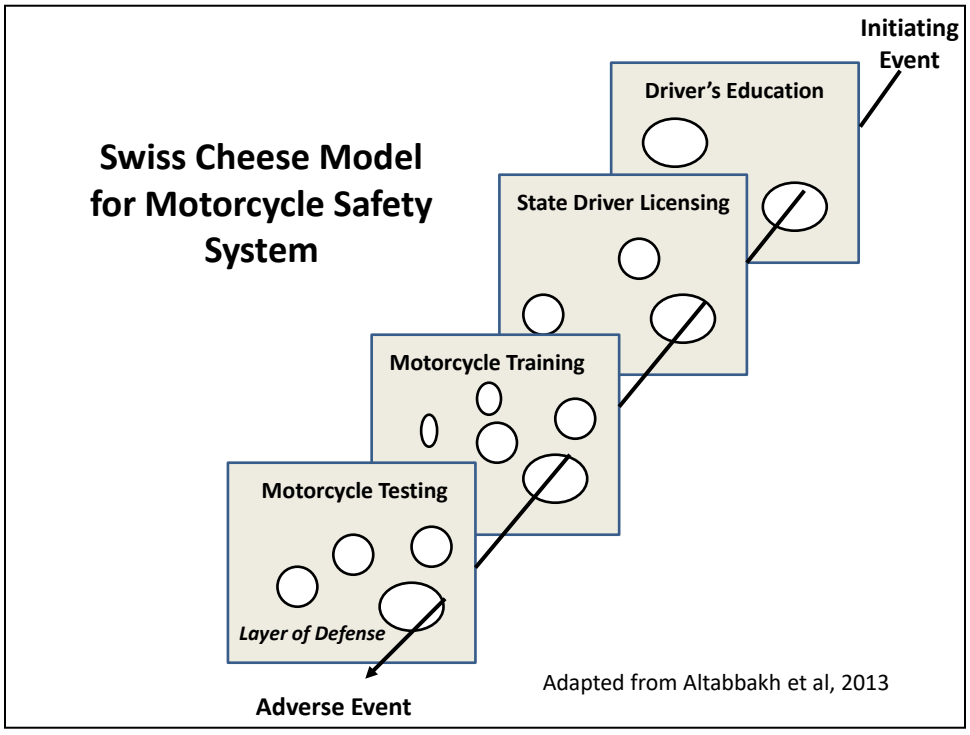


Figure 17 Swiss Cheese Model

The Failure Modes Effects Analysis (FMEA) is another possible creativity-based risk identification technique. FMEA is an inductive technique that assesses all the possible effects from a single cause (Vesely 2002; Altabbakh et al. 2013). In other words, FMEA has the team brainstorm all the possible downstream effects of a failure. The FMEA characterizes failures in three ways: the likelihood of occurrence, the severity of the consequence, and the ability to detect the failure (Shirouyehzad, Dabestani, and

Badakhshian 2011). The objective of a FMEA is not necessarily to identify risks but to prevent unacceptable failures from reaching the end user (Martínez León et al. 2013). The focus on preventing unacceptable failures implies that the team will accept certain failures. By accepting some failures, FMEA becomes a proactive tool for management to apply scarce resources (Olson 2012). FMEA is an important and useful tool in which risk identification could be a byproduct of brainstorming failures.

Six additional creativity-based tools were discussed in this section: *Implications Wheel, Six Thinking Hats, SWOT, LOPA, Swiss Cheese Model,* and *FMEA*. All have proven to be fine visualization and communication techniques that rely on brainstorming for their creativity. Several of these methods depend upon historical information or existing organizational standards. None of these techniques are superior tools for the early detection of black swans as brainstorming and the premortem technique.

2.10 Summary of the Literature Review

Risk identification is often taken for granted but is critical to risk management because every risk not identified is tacitly accepted. Black swans arise from our inability to identify the atypical and unprecedented risks and make appropriate changes. Black swans can come from outside the current body of knowledge or outside the team's perspective. Very often, the black swan appears in front of the team, but the team can fall victim to deliberate ignorance. Deliberate ignorance can be studied, managed, and controlled to catch black swans.

Of the two schools of risk, the subjective school of risk has been least studied but may be a better approach to the early detection of black swans than the objective school.

The reason for this is two-fold: the objective school relies on precedence and the black swan is unprecedented, and the subjective school relies on creativity and collaboration, which have proven to expand our ability to anticipate the unforeseen.

I reviewed the literature on the creativity-based risk identification tools listed in Table 2. Of all the tools, brainstorming and the premortem technique appear to be worthy of further examination and experimentation because they both meet the portability, scalability, low-cost, creativity-based, and collaborative characteristics. Those characteristics are important practical requirements for teams that must perform risk identification under time pressure. Brainstorming and premortem are risk identification tools. They do not require adaptation to become risk identification techniques like the *Six Thinking Hats* for example.

	Portable	Scalable	Low Cost	Creative	Collaborative	Used as a Risk ID Tool
Critique	●	●	●	●		
DBS	●	●	●	●		
Devil's Advocate	●	●	●	●		●
Red Team		●		●	●	●
Scenario Planning		●		●	●	●
Brainstorming	●	●	●	●	●	●
Premortem	●	●	●	●	●	●
FTA						
Implications Wheel	●	●	●	●	●	
Six Thinking Hats	●	●	●	●	●	
SWOT	●	●	●	●	●	
LOPA	●		●	●	●	
Swiss Cheese Model	●	●	●	●	●	
FMEA		●				

Table 2 Comparison of Creativity-Based Risk Identification Techniques

Brainstorming continues to be a popular tool for idea generation, including the generation of risk ideas. With only one known experimental study, the premortem deserves additional scrutiny. Measuring the benefits of the premortem technique's ability to activate creativity and collaboration may provide teams with a tool for the early detection and planning for the black swan.

Chapter 3 - Contributions

3.1 Great Expectations

This research will contribute to the body of risk management practices. I expect this research will be of considerable benefit to both the systems engineering and project management professions by adding a new tool to their risk management approaches. Both the *Systems Engineering Handbook* and the *Project Management Body of Knowledge* have risk management as a common and critical process. By focusing on the common process of risk management, I expect my research will bring the two professions closer together.

I will be providing empirical, experiment-based results on the benefits of creativity-based techniques for risk identification. These techniques complement the quantitative techniques for risk identification. Quantitative techniques are essential to consider risks in complex systems from the perspective of history, lessons learned, and the existing body of knowledge. The creativity-based techniques for risk identification, such as brainstorming and the premortem technique, provide a method for considering risks that may be unprecedented or suppressed by the information pathologies of teams. Those unprecedented and suppressed risks are the home of the black swan. I expect this research to encourage the widespread adoption of the premortem technique in engineering management and systems engineering. Widespread use of the premortem technique for risk identification will contribute to the following:

- A safer tomorrow
- Robust systems and projects ready for wildness, swings between extremes

- Greater synergy between systems engineering and project management

3.2 A Safer Tomorrow

Disasters, both natural and man-made, appear to becoming more frequent and severe (Hossain and Uddin 2012; Patil, Grantham, and Steele 2012; Aven and Krohn 2013; Tarrant 2010). Whether it is an inept response to a coastal hurricane, a terrorist attack, a mine collapse, a ship sinking, or housing market collapse there is a common refrain in hindsight: *"We should have seen it coming."* If only we could get to the hindsight part upfront. The premortem technique's use of prospective hindsight provides teams with a mechanism to see some of the unprecedented and suppressed risks in order to catch those black swans early.

Man-made disasters are the result of human action or inaction that sets in motion a chain of events leading to the black swan event (Abkowitz 2008). From the vantage point of prospective hindsight, using the premortem technique, if we can imagine the disaster, we should be able to change the plan and break the chain of events before they occur. Natural disasters are out of our control. However, our level of preparedness and response has a direct impact on the consequences (J Brooks 2011). The premortem technique can aid our critical infrastructure systems engineers and planners by imagining the worst and considering preemptive measures.

Lack of planning and preparedness is a common theme in nearly every recorded catastrophe (Reeves et al. 2013). People object to committing time and resources to prepare for a disaster that they have never experienced. When planning and preparation does occur, teams fail to consider the variety of possible scenarios. The scenarios they

consider underestimate the impact of the disaster and overestimate the efficiency and effectiveness of the response (Masys 2012; Abkowitz 2008; Chiles 2002). The premortem technique encourages teams to imagine the unimaginable and expect the unexpected.

Too many disasters are the result of overconfidence. In the most extreme cases, overconfidence can manifest itself as arrogance or hubris. This appears either in the person in charge being reckless for selfish purposes or an expert who has become overconfident in their ability to handle any situation (Menkhoff, Schmeling, and Schmidt 2013; Taleb, Goldstein, and Spitznagel 2009; Abkowitz 2008). In the first case, the team's concerns fall on deaf ears or are stifled due to fear of reprisal from their leadership (Walker, Shenkir, and Barton 2014; Abkowitz 2008). In the second case, the expert is blind to the possibilities and anchored in his or her frame of reference. When information counter to the expert's frame of reference (i.e., the black swan) begins to appear, the experts redouble their efforts, vowing that the situation will not get the better of them (Ayton and McClelland 1997). The premortem technique makes it safe for team members to challenge both the team leader and the subject matter expert (Klein 2009).

Communication failures are present in nearly every disaster as either a contributing cause of the event or exacerbating the severity of the consequences after the event (Hossain and Uddin 2012; Abkowitz 2008). A common form of communication failure is when one member of the team does not share critical information with other members of the team. Referring back to the *Transactional Memory Model* of collaboration, an individual either believes the knowledge is already known by others or the individual believes the information is not important (Noble and Letsky 2001). The

premortem facilitates communication by encouraging all team members to contribute ideas without presumption of prior knowledge. We saw this in the cybersecurity application of the premortem technique where the non-technical team members were able to visualize the failures and contribute as well as the team's technical experts (Faily, Parkin, and Lyle 2014). In that case, previously unmentioned risk of delayed processing times came to the surface as well as a previously unmentioned solution (Klein 2004). The premortem facilitates TMM's individual understanding process allowing teams to communicate perspectives to avoid disaster.

We can control the common contributors to disasters: lack of planning and preparedness; overconfidence; and communication failures. We either influence the cause of disasters or the impact the severity of all disasters. Disasters are therefore failures of human systems. Let us consider one human systems failure, the grounding of the *Exxon Valdez* and ecological disaster that followed.

Ever since the Alaska Pipeline opened in 1977, oil tankers have traveled the treacherous and environmentally fragile Prince William Sound. When the *Exxon Valdez*, avoiding ice in the shipping channel, came to rest on Bligh Reef and began spilling over 11-million gallons of crude oil, at least six emergency response plans were in place (Abkowitz 2008). Everything proved to be ineffective. The risks of an oil spill of that magnitude had been underestimated and the response capabilities overestimated. For example, the response plans called for 4,000 gallons of oil dispersant to be on hand in the town closest to the spill and only another 16,800 gallons available across the region. Hundreds of thousands of gallons of dispersant would have been needed to break-up the *Exxon Valdez* spill. There is additional evidence of underestimation of the risk and

overestimation of the response in the shortage of containment booms (8,000-feet), collection bladders (1,000-gallons), and oil collection skimmers (two) (Abkowitz 2008). Clearly, there was either overconfidence or a failure of imagination or communication among the response planning team. If the response planning team had performed a premortem, the environmental response system could have been more prepared. It could have mitigated the $900-million economic and environmental costs (Abkowitz 2008). The premortem technique allows us to imagine a future disaster today for a safer tomorrow.

3.3 Ready for Wildness

The nature of our world, technology, and risk are changing rapidly. Many senior executives believe the unpredictability and wildness and swings between extremes that breed black swans are on the rise. Of the 1,093 senior executives polled in an enterprise risk management survey, 65 percent admitted that they were caught off guard by an operational surprise "somewhat" to "extensively" in the preceding five years. Fifty-nine percent of the 1,093 stated that the volume and complexity of risks has also increased in the same time period (Beasley, Branson, and Hancock 2015). It appears that discontinuities, irregularities, and volatilities are proliferating (Bernstein 1998). American physicist and author Mark Buchanan referred to this trend towards wild swings between extremes as the *upheavability* of systems (Buchanan 2000).

The risk management tools of defining probabilities, measuring the mean, and minimizing risk through diversification do not account for subtle or hidden interdependencies and the influence of human behaviors on systems. This has resulted in

systems that are subject to the effects of nonlinearities where the results are disproportional to the cause (Mlodinow 2008). This new world order, or more accurately new world disorder, gave rise to chaos theory in the 1990s. Chaos theory views the world as being in constant state of turbulence and volatility (Chorafas 1994). According to chaos theory, deviations no longer cluster conveniently around the mean because the mean is always in a constant state of change (Bernstein 1998). Therefore, we should ready our systems for wildness.

The premortem technique for risk identification embraces human behavior as a source of nonlinearities. It also encourages teams to put less weight on past experiences and imagine the unintended consequences. The past will not reveal how the black swan will emerge (Taleb, Goldstein, and Spitznagel 2009). However, the results of a premortem session can provide managers and systems engineers with initiatives that the team can apply to achieve system robustness.

A premortem session should lead to a disrupter analysis, which is a study of how the improbable risk may impact the organization (LeMerle 2011). A team should periodically assess the system's ability to withstand black swans. Disrupter analysis is beneficial in all phases of the system lifecycle but most beneficial in early design. The team can map the theoretical impacts of black swans through the breadth and depth of the system. This should lead to the development of a more robust system design. A natural extension of the disrupter analysis is a disrupter list.

From a premortem session, the team can classify potential black swans by categories such as technological, environmental, economic, political, and societal. After reviewing the items on the disrupter list and their impacts on the system, the team can

synthesize the results to obtain a short list of black swan events that could threaten the system, its performance, and stakeholders (LeMerle 2011). By using prospective hindsight, the premortem technique enables engineering managers and systems engineers to see the sources of wildness and ready their systems for what may lay in wait.

3.4 Close the Gap Between the Technical and Programmatic

According to the International Council on Systems Engineering (INCOSE) the field of systems engineering is currently focused on practice and not on underlying theory. INCOSE has called for the development of a systems engineering foundational theory that encompasses the social and human sciences (INCOSE 2015a). This research contributes to the development of that foundational theory in three areas:

- Risk Communication
- Collaborative Engineering
- Serendipitous Innovation

My research into creativity-based risk identification, specifically the benefit of using the premortem technique within a systems engineering context, addresses a key challenge facing systems engineering, that the "technical and programmatic sides of projects are poorly coupled … hampering effective project risk-based decision making" (INCOSE 2014). This research introduces the systems engineering field to tools that process the knowledge from experts and stakeholders into a useable form for senior decision makers.

As we have seen, the characteristic that separates 21st-century systems engineering from 20th-century systems engineering is the fact that the most dangerous risks are often

hidden from our direct observation. They are hidden in the complex set of interdependencies and volatility of the modern world. The most visible risks receive a disproportionate amount of attention from our risk management processes than less visible risks (Ross 1999). Since black swans are less visible to traditional risk identification techniques, unless we use techniques like the premortem, many dangerous risks will go unidentified and unmanaged. The challenge in catching the black swan comes down to risk communication.

This research contributes to the field of risk communication. Risk communication is founded on the assumption that the utility of risk management depends upon our abilities to perceive and explain risks to other members of the team and stakeholders (Donnelly et al. 2012). It acknowledges that risk identification occurs in an environment of informal rules that color our perceptions of risk. Gary Klein found that intuition is the basis of those perceptions (Klein 2004). He found that the premortem technique provides an environment in which invisible, unspoken risks can come to the surface. It allows the team to communicate their personal visions and perceptions of risks that would otherwise go unspoken, unidentified, and unmanaged. I expect a byproduct of the premortem technique to be improved collaborative engineering.

Collaborative engineering is the integration of teams and organizations across all boundaries. The boundaries include the borders between functional silos, the borders between the technical and programmatic, the borders between supplier and customer, the borders between system and subsystem, the borders between system and the worldwide web, and the borders between system and system. INCOSE expects the systems engineering profession to play a major role in collaborative engineering in the next

decade (INCOSE 2014). This research contributes to collaborative engineering by providing a tool for the systems engineer to analyze the unthinkable. In a traditional setting for risk identification, it would be reasonable for a team member to become defensive at the theoretical prospect of their part of the system failing or being vulnerable to a black swan risk. The common retorts are, "that can't happen" or "if it does happen, there is nothing we can do." The premortem technique, which starts out pessimistically by imagining a fiasco, ends optimistically with the team collaboratively developing changes to the plan to avoid the fiasco.

I expect the improved collaboration from the premortem technique to create an environment for serendipitous innovation. Serendipity or happy accidents can come out of a premortem in two ways. First, black swans can be either detrimental or beneficial. Although the former gets most of the attention, the latter can be the outcome of a premortem as well. For every cloud, there is a silver lining. Second, happy accidents can occur as the team develops changes to the plan to address the detrimental black swan. The premortem technique enables the team to purposefully analyze failure while remaining perceptive and sensitive to a wider context and range of innovative changes to the plan. The team that has performed a premortem is in a frame of mind that is open to possibilities. Ultimately, a proactive attitude toward chance events places the team in a position to follow up on serendipity (El Badawy, Marwan, and Magdy 2015).

3.5 The Challenge

To deliver on the contributions described in this chapter, my challenge was to design a set of experiments that compare the premortem technique to a traditional and

common risk identification technique. The experiments must generate data on the level of collaboration within teams, the number and quality of the risks identified, and the ability of the two techniques to catch black swans.

Chapter 4 - Research Methodology

4.1 Research Goals

My research goal was to measure the benefits of the premortem technique for risk identification within a systems engineering context. I used two studies mentioned in Chapter-2 as conceptual building blocks for my research methodology. The first study is the Chen et al. study on Computer Security Incident Response Teams (CSIRTs). In that study, the researchers concluded that CSIRTs could improve multi-team situational awareness by using the premortem technique (T. R. Chen et al. 2014). The researchers reached their conclusion by basing their empirical study that included structured observation of job performance, supervisor interviews, and participant surveys. The study although empirical, lacked the characteristics of experimentation. The key features of the study that would be helpful in achieving my research goal are the use of structured observation and participant surveys. The fact that the participants were experienced professionals, in a work environment is also applicable. I believe using experienced professionals is an important aspect of my research. However, trying to perform experimentation within the participant's actual work environment is less practical.

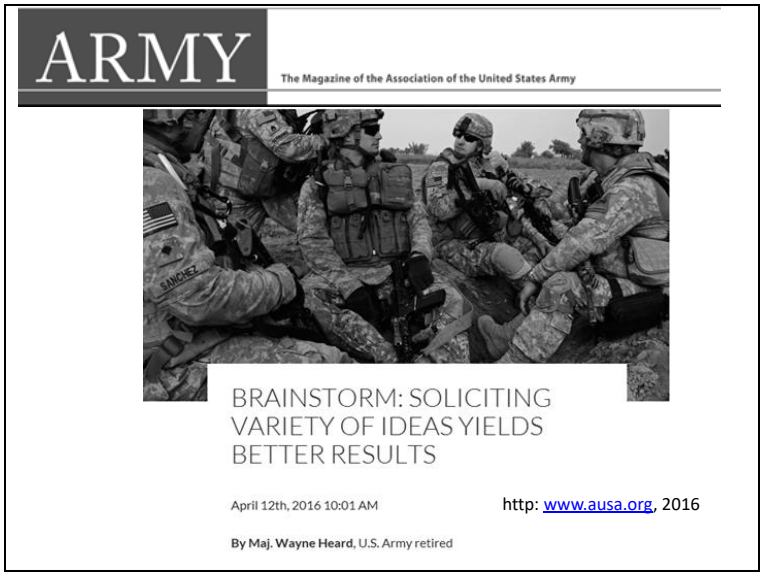


Figure 18 An Example of the Persistence and Popularity of Brainstorming

The Veinott et al. study is the other body of research I built upon. Unlike the Chen et al. study, the Veinott et al. study was experimental based. It measured post-treatment changes in confidence of the participants (Veinott, Klein, and Wiggins 2010). As I mentioned in Chapter 2, the researchers measured the before and after confidence levels on four treatments by means of a survey. The participants were paid undergraduate students working individually. My goal was to use unpaid, experienced leaders and professionals, collaborating as a team. My approach also improved on the Veinott et al. approach by performing side by side "Treatment-A versus Treatment-B" experiments. In doing so, I avoided the pitfalls of researching the premortem technique as an intervention study (ACP 2014). Since I am attempting to quantify the benefit of the premortem technique, a series of side-by-side experiments measuring the premortem

technique against a single appropriate benchmark treatment was more convincing and relevant.

My research goal was to demonstrate that the premortem technique is better than a traditional, creativity-based technique for risk identification. Of all the creativity-based techniques discussed in Chapter 2, including devil's advocate, red team, and scenario planning, brainstorming is the best technique to compete with the premortem technique. There are a number of reasons why brainstorming should be the benchmark. First, brainstorming in general, even with its known pitfalls remains the "go to" technique for teams to generate ideas (Gobble 2014b; Maier 1972). It is a common risk identification tool as well. The May 2016 edition of *ARMY Magazine* (Figure 18) touts the power of brainstorming to identify risks in military plans and make changes to the plan based on those risks (Heard 2016). Second, brainstorming as an inductive technique is the logical antithesis of the premortem (Figure 19). Deductive logic is less common than inductive logic in engineering fields. In engineering, we generally collect evidence to draw conclusions (Vesely 2002; Altabbakh et al. 2013). Using brainstorming as a benchmark may indirectly measure the suitability of deductive logic over inductive logic in a systems engineering context. Third, because of its widespread use, brainstorming required very little training for my test subjects. The Defense Acquisition University, the primary source of experiment participants, has trained teams to perform brainstorming for decades. This reduced the logistics burden of the experiments, although it could provide brainstorming with an intrinsic advantage over the more novel premortem technique.

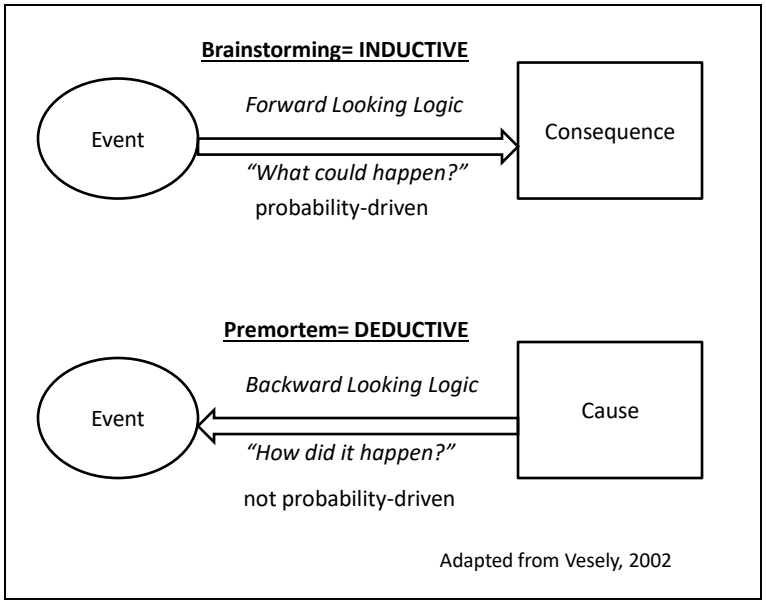


Figure 19 Brainstorming versus Premortem

Another goal of my research was to measure the increase the premortem technique provides to team collaboration over the benchmark (brainstorming). On the surface, measuring the level of team collaboration may appear to be inconsequential. After all, collaboration is an activity, not an outcome. I believe the level of collaboration is likely to impact the outcome of catching the black swan in a number of ways. Going back to the *Transactive Memory Model* discussed in Chapter 2, increased collaboration brings tacit knowledge to the surface for individuals to share with their teammates. In the process of collaborating, the team can share perspectives and experiences relevant to the risks (Figure 20). The hidden and unspoken risks are common sources of black swans.

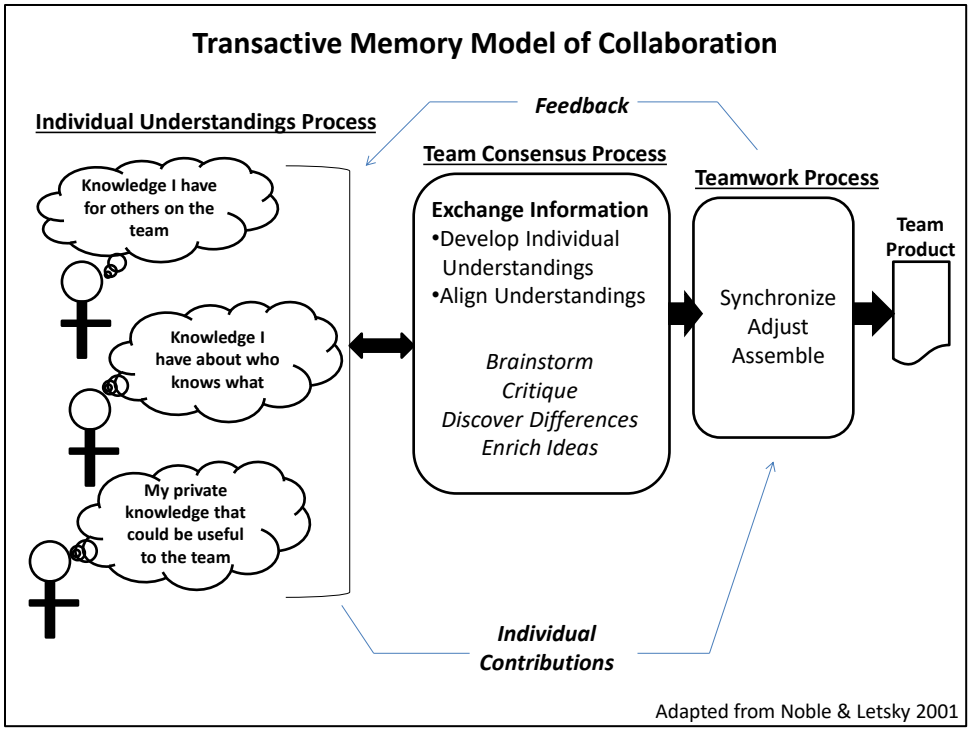


Figure 20 Increasing Collaboration Expected to Aid in Catching Black Swans

Collaboration also enables serendipity, helping teams identify *silver linings* and innovative changes to the plan. With increased collaboration, we can expect to close the gap between the technical and programmatic sides of a project. Collaboration allows the team to look across boundaries to identify black swans lurking in the shadows.

To demonstrate the benefits of the premortem technique to the engineering management and systems engineering community, I needed to test the premortem in a systems engineering context. One of the trends in systems engineering that INCOSE believes will be dominant in the next decade is *brownfield* development (INCOSE 2014). *Greenfield* development is new development. *Brownfield* development improves and existing system. *Brownfield* development often involves technology insertion. There are

several definitions of technology insertion. I use the following: "The positioning of a new or current technology assemblage into an existing system to improve overall system's capabilities" (Hobbs 1994). Testing the premortem (and brainstorming) in a technology insertion scenario is relevant to my research goal in two ways.

Technology insertion in support of incremental or iterative development is the dominant method of systems development (INCOSE 2015b). Technology insertion can be a source of overconfidence in a team (Patil, Grantham, and Steele 2012). The Government Accountability Office cites overconfidence in technology maturity in system upgrades as a challenge in defense acquisition programs (GAO 2015).

Given a technology insertion scenario, I can address my research goal regarding the risks that the team develops. Measuring the quantity of unique risk ideas is of secondary importance to the quality of the risks. My research challenges the underlying premise of brainstorming, that quantity yields quality (Reinig and Briggs 2013). My research goal is to show that the premortem technique generates more quality risks irrespective of the number of risks. By treating risks as *ideas*, I can measure the quality of risks (and proposed changes to the plan) in a similar manner as Briggs and Reinig did in testing their *Bounded Ideation Theory* (Briggs and Reinig 2010).

The most important and difficult research goal was to demonstrate that the premortem technique is better than brainstorming at catching black swans. To do this, I need to return to Chapter 1's discussion of the origins or sources of the black swan. Proving one technique is better than another at catching the black swan that comes from outside the existing body of knowledge, the unknown-unknown may be impossible. However, it is possible to demonstrate one technique is better than another at catching

black swans that are unknown to the team but known to others and black swans resulting from deliberate ignorance (Figure 21). This can be achieved by simulating a historical event to determine if brainstorming or premortem identifies risks that the historical team missed due to a narrow field of view or deliberate ignorance.

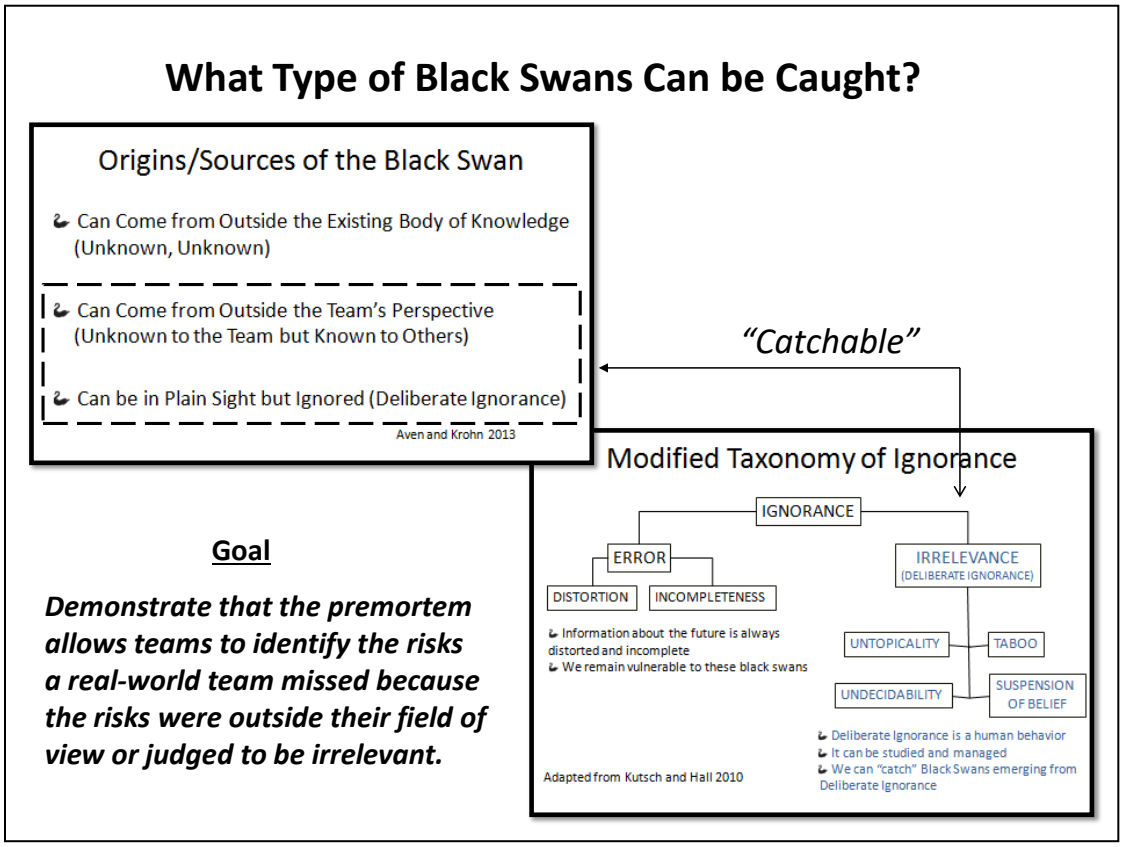


Figure 21 "Catchable" Black Swans

To summarize my research goals, I improved upon the two most relevant studies of the premortem technique by using subjects that are experienced professionals and leaders working as a team. I measured the premortem in a series of "Treatment-A versus Treatment-B" experiments against brainstorming in terms of increased collaboration, the

quality of risks identified and changes to the plan proposed. Most importantly, I intended to test if the premortem technique catches more black swans that originate from outside a team's field of view or deliberate ignorance than brainstorming.

4.2 Research Methods

My research methods included structured observation and two forms of expert elicitation. Those expert elicitation methods are participant surveys and a modified Delphi panel. Structured observation is a research method most often found in the field of education and training. A search of the *ProQuest* database found 97 dissertations or theses that used or referenced structured observation. Of those 97, all but three addressed education and training subjects. None of the 97 dissertations or theses was from an engineering field. Expert elicitation, on the other hand, is a common research method in the George Washington University School of Engineering and Applied Sciences.

Structured observation is a quantitative research method often linked to fixed design experiments in which the researcher takes a detached, pure observer position (Robson 2011). It has the advantage of being a direct method. In other words, instead of asking subjects questions in an interview or on a survey, the researcher records what the subjects do and say directly. Since I wanted to measure collaboration in the real world, structured observation seemed to be best for the task at hand. Data from structured observation may prove to be an interesting contrast or complement to data collected from my other research methods. Structured observation has some limitations. From a practical perspective, it can be very time consuming. From a data integrity perspective, there is no doubt that the observer affects the situation (Robson 2011). There are also a

number of potential biases the researcher must proactively mitigate. Critical to the effectiveness of the structured observation method is the use of a coding scheme. This involves the use of predetermined categories for recording what the researcher observes. There are a number of bases for coding schemes such as non-verbal behavior (body language), spatial behaviors (movement), extra-linguistic behaviors (loudness, interrupting), and linguistic behaviors (Smith 1991). My research method involved solely the latter. I coded the actual content of the team talking and its structural characteristics (what is said over a period of time). The data from observing the subjects can be confirmed or contradicted by the data the subjects report on a survey.

Expert elicitation via survey is my second research method. Since I targeted experienced professionals as participants, a survey of the participants fits the definition of expert elicitation (Meyer and Booker 1991). A research survey is a set of questions, statements, or scales usually asked the same way of all participants (Stake 2010). I used a *Likert* scale to quantify the responses. A *Likert* scale enables a number of parametric data analysis techniques (Carifio and Perla 2007). The participant survey responses can be used to correlate against the data from structured observation. For example, if the structured observation shows a high level of team collaboration, the survey responses can assess if the participants also felt they were very collaborative. One of the most important pieces of information for the participant surveys was demographic information. This was necessary to show that there was not statistical difference (for the dimensions measured) between the subjects using brainstorming and the subjects using premortem. Participant surveys have the advantage of generating a large amount of data in a relatively short period of time. There can be challenges with survey design and low

response rates. To mitigate these challenges, I ran trials to validate the survey and distribute paper instruments to the participants directly. My intention was to design a survey that the participants completed themselves in less than 15 minutes. Those preventions improved the validity of the data as well as the response rate.

My third research method was a modified Delphi panel. In a Delphi panel, the experts give their judgments to a moderator. The experts never interact directly with one another. The Rand Corporation developed this form of expert elicitation to avoid biases that come with experts interacting, such as a dominant expert causing other experts to agree to a judgment that they do not hold (Meyer and Booker 1991). In the classic Delphi, the moderator distributes the judgments to the panel anonymously to allow the experts to revise their initial judgments. The process repeats until the experts achieve consensus. Delphi has been used in a number of risk identification scenarios including software project risks, security projects risks, and public transportation system risks (Kutsch et al. 2013; J Brooks 2011; Sarkar and Dutta 2012). I intended to have the Delphi panel judge the quality of each team's risks and proposed changes to the plan and to decide if any of the team's risks matched the black swans. I call these matches "black swan hits."

For the purpose of my research goals, I was not interested in having the experts achieve consensus. First, I preferred to have the experts unfiltered, first impressions of the risk and changes to the plan quality and the black swan hits. This allowed me to measure the divergence of the two treatments. Second, conducting multiple rounds to reach consensus for the quantity of data points expected was impractical. If I were to conduct 10 experiments for each treatment, and each experiment generates four risks,

four changes to the plan, and four potential black swan hits, there was a total of 240 data points. Driving a panel of experts to consensus on all 240 data points would be time consuming and pointless for my purpose. Therefore, I modified the classic Delphi to exclude multiple rounds to consensus. This form of modified Delphi is consistent with the two-step approach of first soliciting and then combining judgments from a diverse group of experts in order to quantify variables that are intangible or vague (Tseng 2008).

To summarize my research methods, I used three approaches to achieve my research goals. Structured observation allowed a direct coding of the linguistics behaviors as teams use one of the two techniques (brainstorming or premortem). A survey of participants collected demographic information and either confirmed or contradicted the data from structured observation. A modified Delphi panel provided judgment on the quality of each team's work and allowed me to combine and quantify intangible or vague variables. Given my research goals and the methods I selected to achieve those goals, I outlined my research plan.

4.3 Research Plan

As systems get bigger and more complex, we will face problems that have never been encountered before. Our traditional tools for risk identification, based on historical data and Boolean logic, may not be robust enough in the complex and interconnected environment of 21st century systems engineering. We may need to rely more heavily on creativity-based risk identification techniques to catch atypical risks. My research plan was to determine if the premortem technique works better than brainstorming for risk identification in a technology insertion (systems upgrade) scenario. I expected to find

that risks resulting from brainstorming would be typical, predictable risks while risks resulting from the premortem technique would be more creative and provide better changes to the technology insertion plan. My research question is: *Given a technology insertion scenario, does the premortem technique improve team collaboration resulting in better risk identification and changes to the plan as compared to a team performing the brainstorming technique for risk identification?*

A set of experiments can answer the research question. Twenty teams of similar size and demographics were given the same technology insertion plan. Ten of the teams used the premortem technique, and ten teams used brainstorming to identify risks in the plan and propose changes to the plan in consideration of those risks. The teams were instructed to generate a deliverable set, a list of risks that are creative yet plausible, and a list of changes to the plan that are creative and useful. I compared the level of team collaboration, the team's risk list and changes to the technology insertion plan using structured observation, participant questionnaires, and a modified Delphi (expert panel) review of each team's two deliverables. I tested four major research hypotheses in the experiments, and they are described below.

- Null Hypothesis #1: Team collaboration using the premortem technique is equal to team collaboration using brainstorming. The alternative hypothesis, *H1a: Team collaboration using the premortem technique is greater than team collaboration using brainstorming.* Null Hypothesis #1 can be decomposed into four subordinate null hypotheses.

- Null Hypothesis #11: The number of individual differences in perspective discovered using the premortem technique is equal to the number of differences in perspective discovered using brainstorming.

  The alternative hypothesis, H11a: The number of individual differences in perspective discovered using the premortem technique is greater than the number of differences in perspective discovered using brainstorming.
- Null Hypothesis #12: The number of ideas enriched using the premortem technique is equal to the number of ideas enriched using brainstorming.

  The alternative hypothesis, H12a: The number of ideas enriched using the premortem technique is greater than the number of ideas enriched using brainstorming.
- Null Hypothesis #13: The number of critiques applied using the premortem technique is equal to the number of critiques applied using brainstorming.

  The alternative hypothesis, H13a: The number of critiques applied using the premortem technique is greater than the number of critiques applied using brainstorming.
- Null Hypothesis #14: The number of original risk ideas offered using the premortem technique is equal to the number of original risk ideas offered using brainstorming.

  The alternative hypothesis, H14a: The number of original risk ideas offered using the premortem technique is greater than the number of original risk ideas offered using brainstorming.

- Null Hypothesis #2: The number of quality risks identified using the premortem technique is equal to the number of quality risks identified using brainstorming.

  The alternative hypothesis, H2a: The number of quality risks identified using the premortem technique is greater than the number of quality risks identified using brainstorming.

- Null Hypothesis #3: The number of quality proposed changes to the plan using the premortem technique is equal to the number of quality proposed changes to the plan using brainstorming.

  The alternative hypothesis, H3a: The number of quality proposed changes to the plan using the premortem technique is greater than the number of quality proposed changes to the plan using brainstorming.

- Null Hypothesis #4: The number of historical risks identified using the premortem technique is equal to the number of historical risks identified using brainstorming.

  A historical risk is a risk that the historical team should have identified but did not. I refer to matching the historical risk as a *black swan hit*.

  The alternative hypothesis, H4a: The number of historical risks identified using the premortem technique is greater than the number of historical risks identified using brainstorming.

4.4 Data Collection Plan

I used structured observation to collect data to determine Null Hypothesis #1. I expected the subject survey results to correlate with the structured observation data. The expert panel judgment results determined Null Hypotheses #2, #3, and #4. In the

following sections, I will discuss the nature of the subjects, expert panel members, and the specific experimental steps and data collection instruments.

4.4.1 Subjects and Expert Panel Members

Unlike many studies that use undergraduate students, my experiments involved experienced professionals and leaders. I first describe the subjects involved in the experiments and then describe the composition of the expert panel. For the experiments, my target audience consisted of students in the Defense Systems Management College (DSMC) executive programs courses, the Eisenhower School's Senior Acquisition Course, United States Coast Guard's (USCG) program office personnel, United States Navy (USN) Future Frigate program personnel, and Textron Corporation program personnel. All subjects were unpaid volunteers.

DSMC is part of the Defense Acquisition University and is chartered to provide executive-level acquisition management, international acquisition management, and requirements management training ("Defense Systems Management College" 2016). DSMC also provides training to government agencies and the defense industry. DSMC provides a set of executive-level resident courses for skills improvement, critical thinking and decision making, and training for leaders of the Department of Defense's Major Defense Acquisition Programs. The DSMC subjects came from the three executive resident courses and participated outside of scheduled class time on a completely voluntary basis (Figure 22).

DSMC EXECUTIVE RESIDENCE COURSES
Program Manager's Skills Course
This course provides O-5/GS-14, Level-III Program Management Career Field acquisition professionals with the latest acquisition policies and proven practices in the areas of Requirements, Acquisition, Finance, and Technical Management. Additionally, students will have an opportunity to examine and discuss key program manager skills and lessons learned and develop a plan for strengthening their skills for program success.
Program Manager's Course
This course is designed to improve Department of Defense acquisition outcomes by strengthening the analytical, critical thinking and decision-making skills of potential leaders of major defense acquisition programs and program support organizations. Applying the proven doctrine of “train as you fight” participants analyze acquisition case studies representing contemporary acquisition program challenges and dilemmas; apply a broad cross section of knowledge of the acquisition environment and experience; and deepen their understanding of acquisition principles and practices through peer and instructor mentoring and coaching. Speakers, team projects, media training, and management simulations round out and enrich the course.
Executive Program Manager's Course
Assignment specific course designed to meet the learning and performance needs of newly selected Program Executive Officers, Deputy Program Executive Officers, and Major Defense Acquisition Program Managers and Deputy Program Managers. Skills and behaviors are developed through a concentrated four week period preceded by approximately 8-16 hours of advanced pre-course assignment.

Source http://www.dau.mil

Figure 22 DSMC Executive Resident Courses

The Eisenhower School (formerly known as the Industrial College of the Armed Forces) is part of the National Defense University (NDU). Graduates receive a Master of Science degree in National Resource Strategy. The mission of the Eisenhower School is to:

"Prepare selected military officers and civilians for strategic leadership and success in developing our national security strategy and in evaluating and marshalling, and managing resources in the execution of that strategy." (NDU 2016)

Subjects from the United States Coast Guard came from two areas. The Director of Acquisition Programs (CG-93) oversees the acquisition, integration and delivery of the Coast Guard's Major Acquisition Systems (USCG 2016). The second area of the Coast Guard that provided subjects was the Engineering and Logistics Directorate (CG-4). CG-4 provides technical, logistics and engineering support to all Coast Guard operating programs (USCG 2016).

Subjects from the USN Future Frigate program office (PMS-515) are part of the Program Executive Office for Littoral Combat Ships. The PMS-515 personnel manage the multi-mission upgrades for the existing frigate fleet and mission packages for the new Littoral Combat Ships (Antonio 2015). These subjects had both programmatic and technical experience.

Figure 23 Textron Systems Overview

Subjects from Textron came from the Textron Systems business unit. That business unit develops and integrates products, services, and support to aerospace and defense customers (Textron 2012). Subjects participating in this study work on

unmanned aircraft, marine, and land systems as engineers and program managers (Figure 23).

The expert panel members came exclusively from the DSMC faculty. I selectively recruited the expert panel members to provide a set of diverse perspectives when evaluating the products from the experiments. The panel members had a wide range of subject matter expertise to include operational requirements, engineering, finance, contract management, test and evaluation, and program management. Expert panel member biographies are in Appendix A.

4.4.2 Scenario Selection

The technology insertion scenario for the subjects to apply during the experiments was critical to the data collection. The criteria for the selection of the scenario were:

- The scenario can be generalized to other technology insertion, product upgrade engineering efforts.
- The subjects can readily visualize the operational concept of the product receiving the upgrade.
- The subjects do not require unique engineering or management knowledge, skills, or abilities to understand the scenario.
- The scenario is based on a documented project failure that resulted from the team not looking outside their field of view or falling victim to the elements of deliberate ignorance. These are the "catchable" black swans mentioned earlier.

- The documented failure cannot be widely known (subjects did not know the black swans beforehand).

Although I considered a number of scenarios, I decided the Mark-14 torpedo technology scenario best fit the criteria. James Chiles describes the Mark-14 torpedo project in his book, *Inviting Disaster*. The United States Navy formed a top secret project to upgrade the reliable Mark-14 torpedo, which depended on a contact fuse for detonation, with a proximity fuse. The proximity fuse was a mature technology from an anti-ship mine. The Navy's intention was to increase the capability of the Mark-14 by integrating the proven proximity fuse, while keeping the contact fuse capability for the expanding submarine warfare in World War II Pacific Theater of operations.

The project was a disaster, resulting in the fielding of unreliable and dangerous torpedoes. Chiles indicated the main reason for the project failure as overconfidence and the team explaining away problems (Chiles 2002). I inferred four black swans from the Mark-14 case:

- The development team tested torpedoes without explosives to conserve limited test articles.
- The development team tested the torpedoes without hitting objects so that they could be recovered and reused for more testing.
- The development team did not consider that the environmental conditions in operation would be different than test conditions, resulting in torpedoes travelling too deep for the proximity fuse to sense the magnetic field of the enemy ships.

- The integration of the magnetic fuse caused interference with the contact fuse, resulting in the torpedoes bouncing off the enemy ships when fired in the contact fuse mode.

In reviewing the above list, a reasonable person might say, "these look like white swans to me." One of the characteristics of the black swan I mentioned in Chapter 1 was that black swans appear to be obvious in the perspective of hindsight. The important point for my experiments was that that the premortem technique would help the subject teams see what the original team failed to see, better than brainstorming.

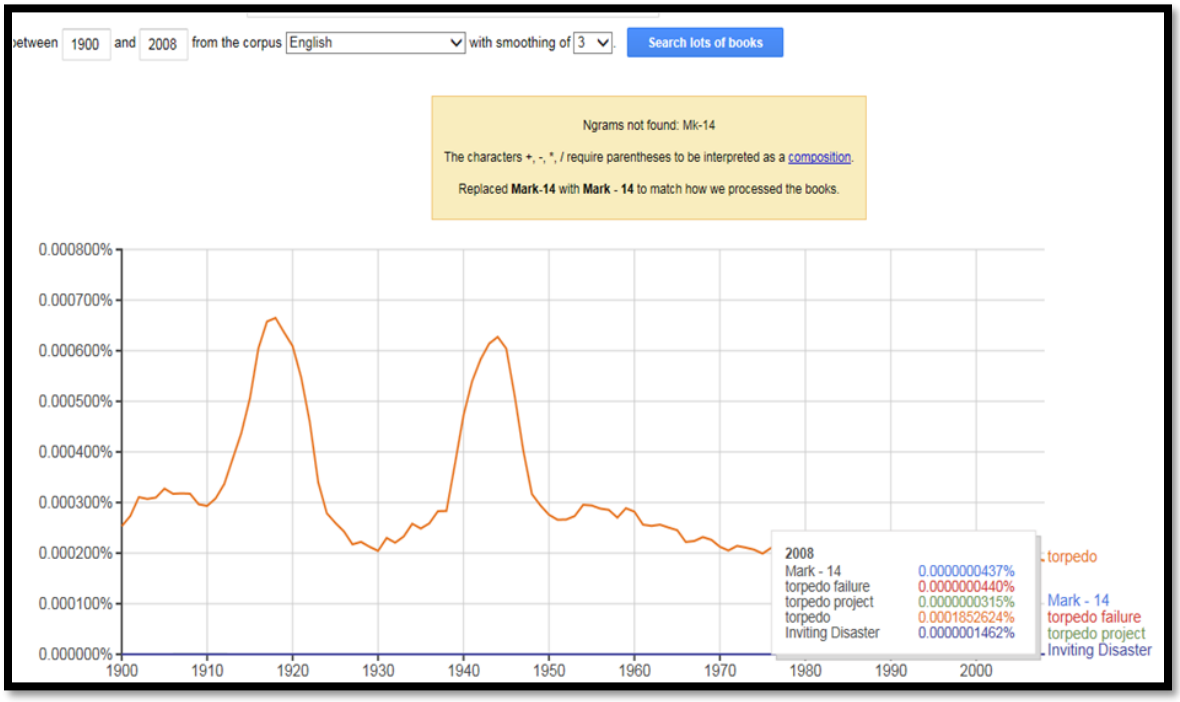


Figure 24 N-Gram Viewer Results for Words and Phrases Related to the Mark-14 Case

The scenario satisfied the selection criteria. Although it was a naval weapon project, the scenario can be generalized to any project that seeks to improve the performance of a fielded capability by applying a mature technical capability. A software upgrade or a facilities improvement would be appropriate analogies to the Mark-14 project. With some very simple visual aids, the subjects can visualize the operational concept of the Mark-14 regardless of their technical, operational, or programmatic background. The Mark-14 scenario does not involve unique engineering or management concepts. The black swans in the Mark-14 are of the "catchable" variety, resulting from the team's narrow field of view and deliberate ignorance.

The final criterion of not being widely known is difficult to test in advance. The *Google Books N-Gram Corpus* enables the quantitative analysis of linguistic and cultural trends recorded in millions of books (Lin et al. 2012). For example, an N-Gram search for key words and phrases from the Mark-14 scenario yields a very low (0.0000000437-percent) occurrence in the books corpus between 1900 and 2008 for the word "Mark-14" (Figure 24). This is an indicator that the Mark-14 case is not widely published. Unfortunately, this criterion can only be validated, on a subject-by-subject basis, after the experiment has been administered.

4.4.3 Design of Experiment

I designed a classic *Treatment-A versus Treatment-B* experiment. This type of randomized, controlled experiment is often referred to as *A/B testing*. As the name suggests, two versions are compared, A and B. Each test is identical except for one

variation that may impact the subject's behavior. The resulting data supports two-sample hypothesis statistical analysis (Kohavi et al. 2009). Statistical tools such as the *Mann-Whitney* test can then be applied to test the hypotheses. A/B testing has become a common technique in the fields of marketing and business intelligence. Google uses A/B testing in software development, measuring the change in human behavior resulting from a change to the user interface (Kohavi and Longbotham 2015).

The first step in designing an A/B test is to define the Overall Evaluation Criteria (OEC). The OEC is the quantitative measurement of the experiment's objective, sometimes referred to as the response or dependent variable (Kohavi et al. 2009). As I discussed in the Research Goals section, the OEC for my experiment are:

- Increased team collaboration (with subordinate measures)
- Increased quality of risks identified
- Increased quality of the changes to the plan offered based on the risks identified
- Increased number of "black swan hits"

The factor thought to influence the OEC in my experiment is the type of risk identification technique, either brainstorming or premortem, the team uses to assess the technology insertion plan. This leads me to the two variants or treatments, *brainstorming and premortem*. In A/B testing, the desired power of the test should be between 80 percent and 95 percent (Kohavi and Longbotham 2015; Kohavi et al. 2009). In testing web design, the number of individual users exposed to the treatments determines the power of the test. Since my OEC depend upon a difference in team behavior and not

individual behavior, the common formula for determining sample size, $n = \frac{16\sigma^2}{\Delta^2}$ where n is the required observations, σ^2 is the variance in the OEC, Δ^2 is the amount of change you want to detect, and the coefficient *16* delivers the 80-percent probability of rejecting the null hypothesis that there is no difference between A and B treatments (Kohavi et al. 2009) was not practical for my application. This impracticality was due to my lack of prior knowledge of the population variance and inability to sample hundreds or thousands of teams.

To establish a powerful test that is also economically reasonable, I turned to the laboratory sciences. The *Mead* method for determining the sample size is based on the law of diminishing returns for laboratory animal experiments. Although the subjects for my experiments were not laboratory animals, I considered the number of subjects in my experiments to be a valuable and limited resource. I believed the Mead method is appropriate for determining the number of experiments to run in this study.

Mead's method is based on the errors in the degrees of freedom or *E*.

$$E = (Total\ Number\ of\ Animals) - (Total\ Number\ of\ Treatment\ Combinations)$$

The experiment is appropriately sized when $10 < E < 20$ (Robert Hubrecht and Kirkwood 2010). With two treatments, assuming the minimum number of subjects for each of my experiment teams is four, according to the Mead method the minimum number of teams per treatment would also be four, as calculated below.

$$E = (4\ teams * 4\ persons * 2\ treatments) - (8\ team_treatment\ combinations),$$

$$Therefore\ E = 16 \quad and \quad 10 < E < 20$$

If the objective is to compare Treatment-A to Treatment-B, the number of teams per treatment must be the same (Kohavi et al. 2009). Because I was also trying to use survey information, my goal was to have ten experiments per treatment. According to Mead's method, this resulted in $E = 60$. This might appear excessive, but I expected ten experiments per treatment and at least four subjects per experiment would provide more powerful survey data.

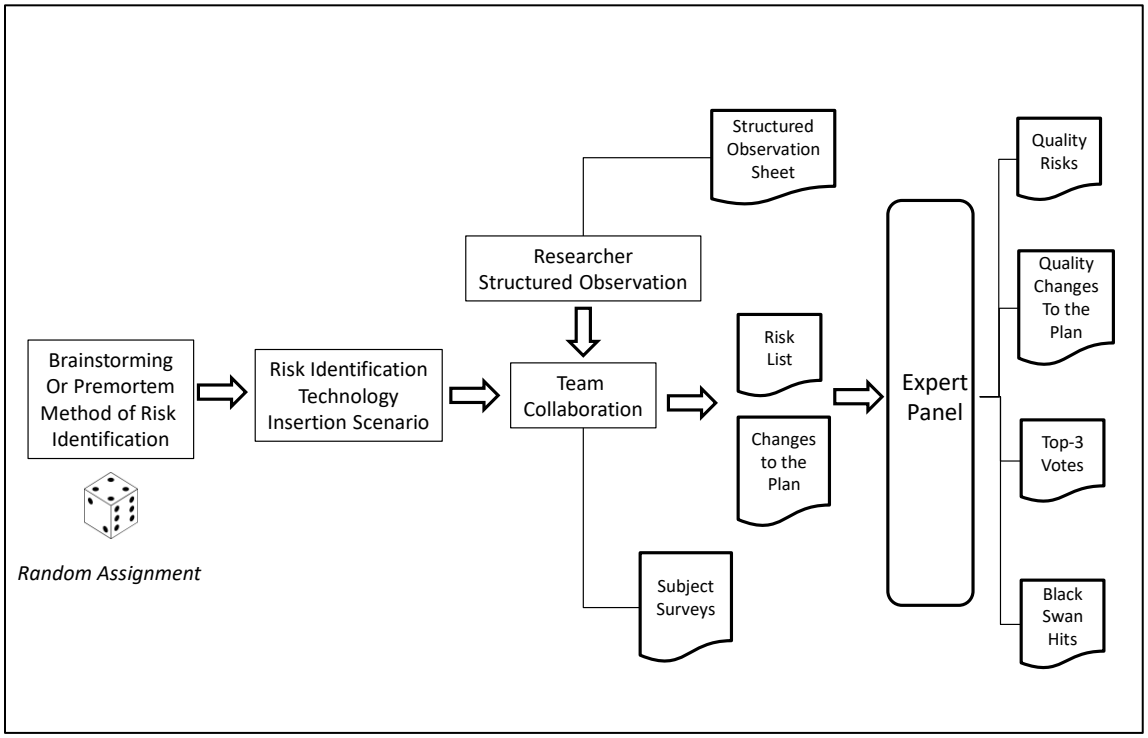


Figure 25 Data Collection Diagram

I will now describe the flow of the data collection supporting the experiments (Figure-25). The data collection flow began with random assignment of the subject team to a treatment. I used a die roll for random assignment. Prior to knowing the number of

subjects or composition of the team, I rolled a die. If the roll resulted in an odd number, I assigned the team the brainstorming technique for risk identification. For an even roll, the team used the premortem technique for risk identification.

My protocol established the minimum team size at four and the maximum team size at six. I set the protocol as a range because trying to get an exact number for every experiment would not be practical given my dependence on volunteers. I set the minimum as four because I needed the team to self-facilitate to avoid any bias an outside facilitator may introduce. In the case of three team members, given one team member was participating and facilitating, the remaining two team members may have been at a disadvantage for collaboration. I set the maximum number of team members at six because I believed this was the maximum number of participants that I could accurately record using paper and pencil during structured observation.

Structured observation was the method I used to record collaboration during the experiment. I considered using video or audio recording to capture collaboration during the experiment. This would definitely allow a more accurate recording of collaborative exchanges during the experiment. However, I believed video and audio recording would cause privacy concerns that would deter volunteers from participating in this study. I was also concerned that the use of video and audio devices would cause a *Hawthorne* effect. The subjects, knowing that their voice and image were being recorded, would modify their behaviors and change experiment outcomes due to the increased attention (Stake 2010). I believed the observer coding by means of paper and pencil would influence subject behavior much less than audio or video recording.

By the end of the collaboration, the team produced their top risks and changes to the plan based on those risks. The experiment session concluded with a participant survey. Participation in all aspects of the experiment, including providing survey responses, was completely voluntary. I allowed several points in the process for subjects to opt-out of the data collection. After I conducted ten experiments per treatment, I turned over the team products (risk lists and changes to the plan) to the expert panel for scoring.

The construct of the expert panel was discussed earlier in this chapter. I directly recruited thirty DSMC faculty members to serve on the expert panel. As discussed earlier, the panel members graded all twenty sets of team products independently. The twenty sets represented ten sets for each treatment. The panel members scored the team products based on specific scoring criteria. The expert panel also provided me with their top three sets of deliverable based on their personal assessment criteria, role playing as senior decision makers.

I will now describe the specific session activities I planned for in and around the experiment itself. The experiment was given as part of an elective workshop or mission assistance to organizations. DSMC has found that the length of an elective will have a direct impact on the number of participants that will volunteer to participate in an elective workshop. Experience has shown that the time for an elective workshop should not exceed two hours. Therefore, I used 120 minutes to size the entire session that includes instruction, administration, and experimentation (Figure 26).

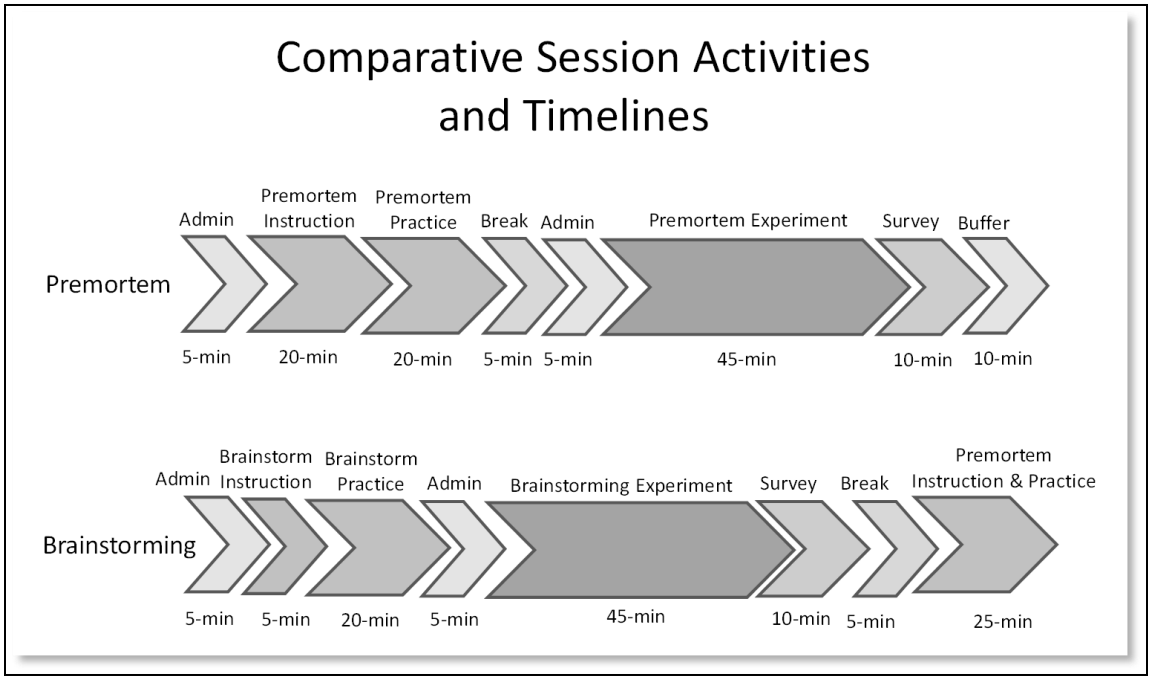


Figure 26 Treatment Activity Timelines Compared

Both treatment sessions began with five minutes of administration for introductions and to cover the information in the human research consent form. This was followed by a block of instruction on the technique the team applied in the experiment. Because brainstorming is so common and simple, that block of instruction was expected to take less than five minutes. Since there are many ways to perform brainstorming, the focus was on calibrating the brainstorming teams on the specific method all brainstorming teams applied. The specific method of brainstorming can be found in the *Memory Jogger II* and has been taught at the Defense Acquisition University since the 1990s (Brassard and Ritter 1994). The premortem, as a new technique, required more instruction time. I expected the premortem would require 20 minutes of instruction.

Teams for both treatments then had 20 minutes to practice their assigned technique in which I served as the facilitator. The brainstorming team practiced brainstorming the risks in a plan to commit a crime and brainstorm changes to the plan based on those risks. The premortem team practiced performing a premortem on a plan to add wind turbines to the Eiffel Tower and also create changes to the plan to address the risks they identify. The practice scenarios were designed to be easy to visualize and understand.

At the conclusion of the practice exercise, teams performing the premortem took a five-minute break before receiving the administrative set-up for the experiment. Brainstorming teams went right into the administrative set-up for the experiment. During the administrative set-up, I read to the subjects (as they followed along with their personal copy) the experiment instructions. The subjects then were given three minutes to read the one-page project charter that outlined the technology insertion plan. I then answered the team's questions before beginning the experiment. Once the experiment began, I did not interact with the subjects other than to provide them with the time remaining in five-minute intervals after the first 15 minutes had expired.

I intended to make the following variables fixed between all experiments and across both treatments:

- All experiments received the same technology insertion scenario. The scenario was the "Marlin Torpedo Upgrade," which mirrored the Mark-14 case study discussed earlier.
- All experiments received exactly 45 minutes. A large sweeping-hand timer was visible to all subjects.

- Teams had four, five, or six team members.

There are some variables that I acknowledged beforehand that would be impractical for me to control. This is a list of those variables I did not control:

- The mix of experience levels within teams
- The mix of education levels within teams
- The mix of professional backgrounds within teams
- The level of familiarity of the team members to one another
- The environmental conditions, such as size and location of the classroom, level of classroom lighting and acoustics
- My skill level in providing the instruction, conducting the experiments, and collecting the data
- My rapport with the subjects and their interest and commitment to participating

The experiment concluded when the team submitted its deliverables or 45 minutes have expired. Immediately upon concluding the experiment, I administered the 15-statement survey. Subjects received a paper survey form and a blank envelope. I read each statement and possible response so that I could answer questions and clarify statements once for the whole team. It also prevented me from unintentionally influencing subject scoring by watching them complete the survey.

At the conclusion of the survey, teams performing the premortem treatment were finished. There was a ten-minute buffer in the premortem treatment timeline that I applied to the premortem technique instruction if necessary. After the survey, the

brainstorming teams took a break and then received premortem technique instruction and a practice exercise. Their session concluded with a brief discussion comparing and contrasting brainstorming and premortem. Only the teams performing the brainstorming treatment had the opportunity to experience both techniques. That provided the opportunity to capture anecdotal comments from participants regarding the similarities and differences of the two techniques. In the final moments of the session for both treatments, I asked the team if they were familiar with any cases similar to the Marlin Torpedo Upgrade to determine if they had prior knowledge of the Mark-14 case.

4.4.4 Data Collection Instruments

I used four data collection instruments: the structured observation sheet, the team deliverables, the subject survey, and the expert panel grade and feedback sheet. In this section, I will describe the purpose of each data collection instrument, the method by which it was used, and the outcome of each data instrument.

The structured observation sheet allowed me to record and measure the level of collaboration among the team members while conducting the experiment. As mentioned earlier, this technique provides the advantage of recording behavior directly. I based my structured observation data collection instrument on Robson's example (Figure 27). He establishes rows to record the occurrence of the behaviors he expects to observe and columns representing time intervals (Robson 2011). By using tick marks, the observer can record the occurrence of the behaviors over time. My structured observation sheet numbered the subjects one through six in the rows. The columns represent nine, five-minute time intervals covering the entire forty-five-minute experiment. I expected to

record four relevant behaviors, so a single tick mark system was not sufficient. I therefore established a simple coding system:

| = the subject offers an original risk idea to the team for consideration

Ƥ = a subject offers a different perspective on a risk idea proposed by another

E = a subject enriches (or builds upon) a risk idea proposed by another

C = a subject critiques a risk idea proposed by another

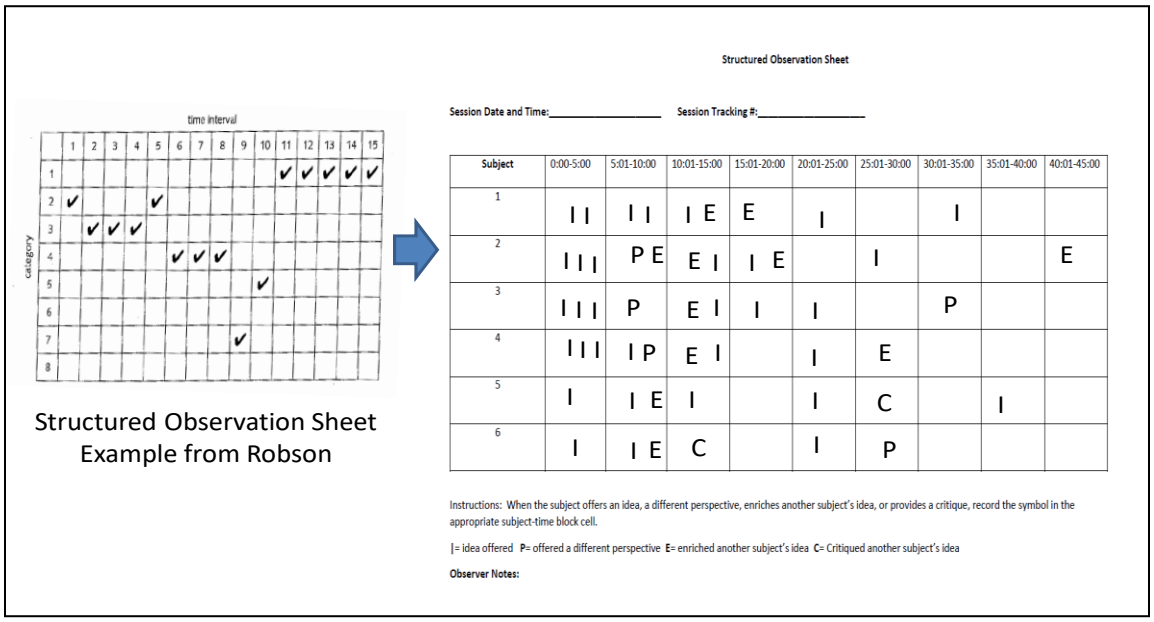


Figure 27 Structured Observation Data Collection Instrument

Below is a notional set of collaborative exchanges to illustrate the use of the structured observation system:

Subject-1: "The Marlin torpedo sinks its own ship." Observer records: |

Subject-2: "Perhaps the torpedo sinks its own ship because it explodes in the launch tube. Observer records: E

Subject-3: "I don't see the torpedo sinking its own ship by exploding in the launch tube. I see the torpedo leaving the launch tube and homing back in on its own ship." Observer records: P

Subject-4: "The Marlin torpedo cannot sink its own ship. We have safeguards in place." Observer records: C

From the structured observation sheet, I expected to map the level of collaboration of each experiment. I then compared the collaboration levels between the two treatments. This can be done for the aggregate of the collaborative exchanges or each type of exchange independently. This allowed me to perform hypothesis testing to address Null Hypothesis #1: *Team collaboration using the premortem technique is equal to team collaboration using brainstorming*. It also may allow future examination of other interesting questions such as:

- Is the premortem technique less likely to experience social loafing than the brainstorming technique?"
- Does the team leader/facilitator contribute more to collaboration using the premortem technique than the brainstorming technique?
- Does the premortem technique have less periods of time when no collaboration is occurring than the brainstorming technique?

As discussed earlier in this chapter, there are limitations and pitfalls to the structured observation method of data collection in general and the data collection instrument specifically. I will discuss the limitations and mitigations in the *Pilot Testing* section.

The team deliverables data collection instrument had two parts: the *risk list* and *the changes to the plan* based on those risks. The purpose of this instrument was to capture the output of the team's collaboration process. I established a level of consistency in the format of the team deliverables to facilitate the expert panel scoring of the deliverables. Therefore, I needed to eliminate issues that might come from the team-to-team differences in the number of *risks* and *changes to the plan* identified as well as differences in legibility. All teams delivered four *risks* and four *changes to the plan* (Figure-28). I explain how I arrived at that number in the *Pilot Testing* section. I transcribed handwritten submissions into a typed format to eliminate difference in legibility. However, I did not correct the team's grammar. I also informed the team that the team did not need to list the *risks* and *changes to the plan* in priority order. This was to preclude the team from unnecessarily devoting time to debate rank ordering.

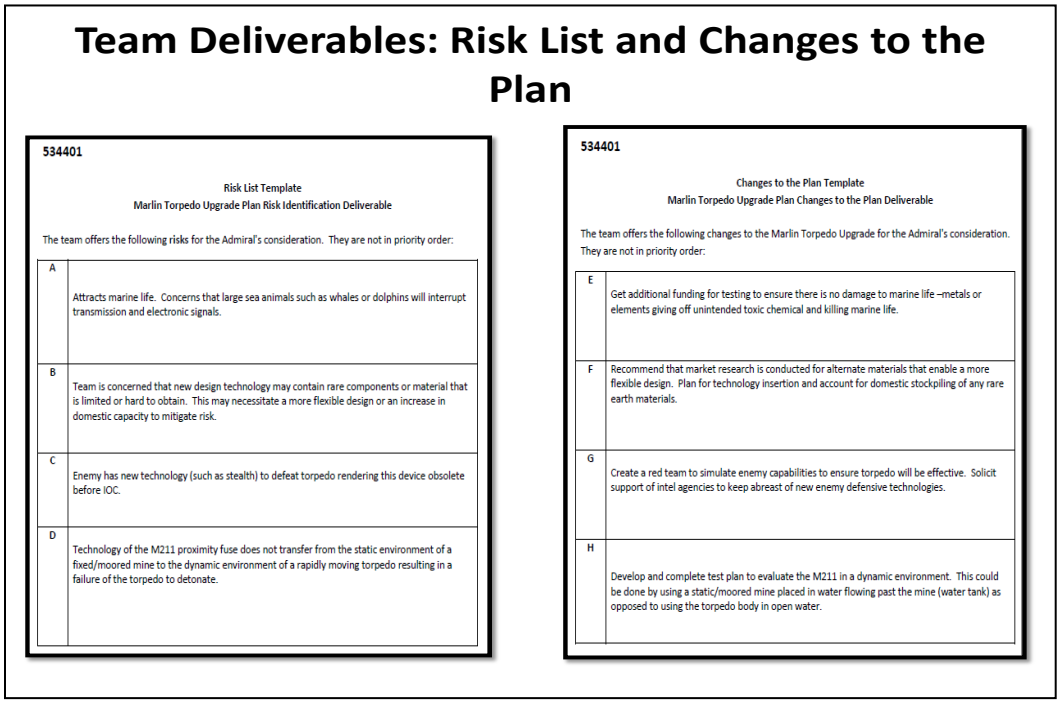


Figure 28 Team Deliverables Data Collection Instrument

The expert panel judged the quality of the *risks* and *changes to the plan* without consideration of the team's prioritization. The expert panel also reviewed the team's risk list for *black swan hits*. The expected outcome of the deliverables data collection instrument is an expert panel grade for each risk item (A through D) and each change to the plan (E through H) for every experiment. This facilitated the testing of the following null hypotheses:

- Null Hypothesis #2: *The number of quality risks identified using the premortem technique is equal to the number of quality risks identified using brainstorming.*
- Null Hypothesis #3: *The number of quality proposed changes to the plan using the premortem technique is equal to the number of quality proposed changes to the plan using brainstorming.*
- Null Hypothesis #4: *The number of "historical risks" (black swan hits) identified using the premortem technique is equal to the number of "historical risks" identified using brainstorming.*

The deliverables data collection instrument served as the bridge between the team collaboration experiment and the expert panel review of the outputs of collaboration.

The purpose of the subject survey/questionnaire (Figure-29) was to do the following:

- Collect demographic information.

- Measure the subject's perspective on the level of team collaboration they experienced in the experiment, measure of the subject's perspective the treatment used against other personal benchmarks.
- Measure the subject's perspectives on risk attitudes in the real world.

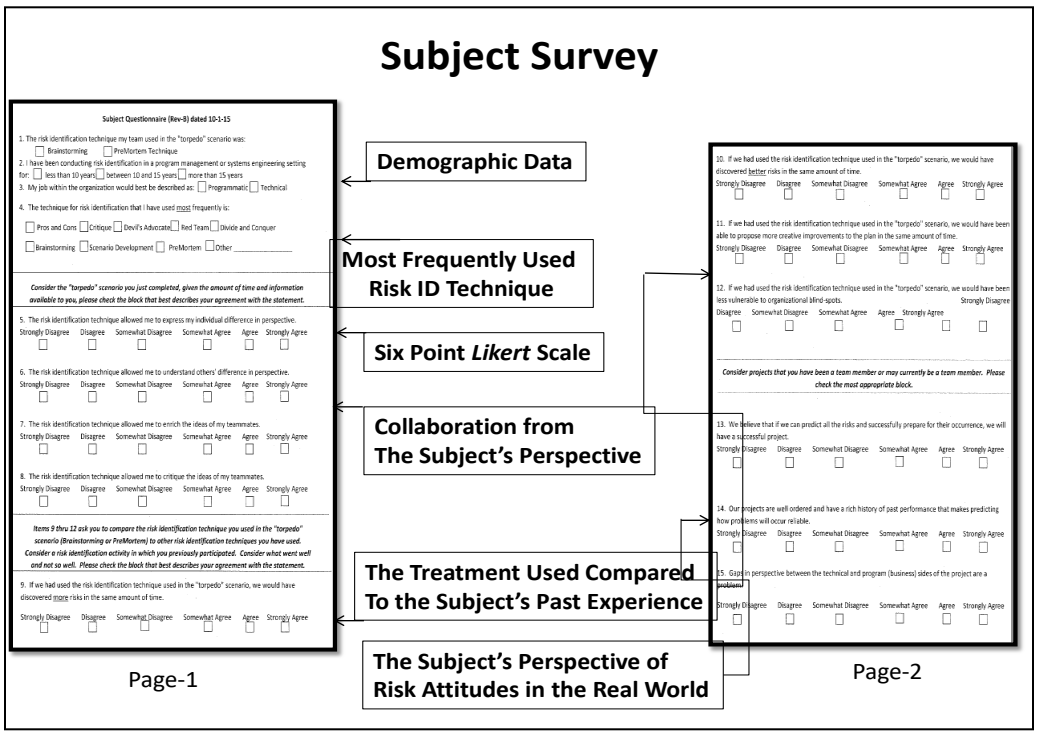


Figure 29 Components of the Subject Survey/Questionnaire

The demographic data was intended to confirm that there was no statistical difference between the subjects using the two treatments. There were two demographic measures, experience performing risk identification and career background. I considered collecting other demographics such as highest education level, age, sex, and career specialty area, but they did not appear to be relevant to the research question. None of those measures were likely to influence the outcome A/B testing. However, if one treatment had a statistically significant advantage over the other in risk identification

experience or career background, it could bias the results. In the demographic data section of the survey, I asked the subject to identify the risk identification method they have most frequently used. I expected this data to confirm that brainstorming was indeed the most used risk identification technique. If this was true, we can infer that brainstorming had an inherent advantage of familiarity over the premortem technique. I also used this information to determine if a team that has been tasked with performing the premortem technique has any members that have used the premortem technique. This could provide an advantage to the premortem in the A/B testing.

The remaining sections of the survey/questionnaire use a six point *Likert* scale. I asked the subject to rate their agreement with a statement by checking one of the following: *Strongly Disagree, Disagree, Somewhat Disagree, Somewhat Agree, Agree,* and *Strongly Agree*. I did not offer the subject a neutral position. Studies have shown that given a neutral option, subjects will select the neutral by default (Fowler 1995). The first section using the *Likert* responses was my attempt to either confirm or contradict the collaboration reported in structured observation. Since I provided respondents with anonymity, I was not able to correlate an individual subject's observed level of collaboration with their self-reported collaboration level. However, I can correlate a team's observed level of collaboration with the team's survey results.

The second section, also using the six-point *Likert* scale, measured the subject's perspective of the treatment they used in the experiment against another technique that they have experienced in a real world setting. I expected this data to show that collectively, the subjects that used the premortem would report that the premortem was better than their personal benchmark. The final section of the survey/questionnaire was

intended to confirm or contradict some underlying principles of my research goals. These included the level the subject's belief in the correlation between the ability to predict risks and successful project outcomes; the predictability of real world projects; and the gap between the programmatic and technical sides of real world projects.

The first and last sections of the survey were aggregated without regard to treatment. I expected the number of aggregated surveys to exceed n = 80 based on the minimum number of team participants of four and twenty total experiments. The second section was used to compare the two treatments. Each treatment would have at least n = 40 surveys.

The final data collection instrument is the expert panel score sheet. When all the experiments were complete, I compiled the twenty sets of deliverables for expert panel grading. As a modified Delphi technique, the expert panel was graded individually and independently. I made no attempt to drive the panel to consensus. Each panel member graded the twenty sets of deliverables with no identification as to the technique the team used to create the deliverables. Each deliverable had two parts, the *risk list* and the *changes to the plan*. The risk list represents what the team believes are four of their most creative yet plausible risks. The changes to the plan represent four creative and useful changes to the plan based on the four risks identified. Therefore, each panel member had eight components per team to score. The panel member must assign a creativity score and a plausibility score for each risk. Each change to the plan also received a two-part score, creativity and usefulness. The scoring scale for all four elements is 0 to 100. I instructed expert panel members to use the grade of 70 as the low end of a quality score.

Any score lower than a 70 would be interpreted as not making the quality standard. Figure 30 is an example of the expert panel scoring sheet.

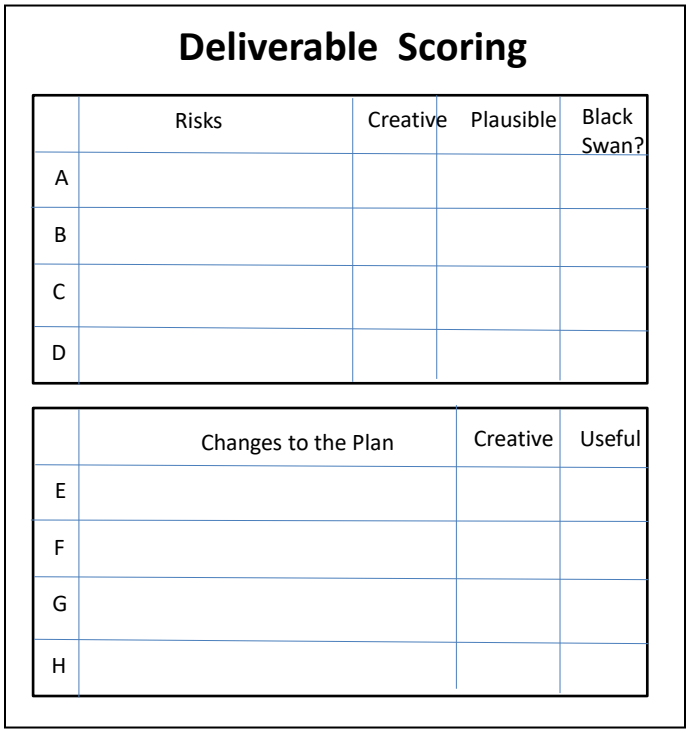


Figure 30 Scoring for Each Deliverable Element

Borrowing from Reinig and Briggs' *quality in ideation* research (Reinig and Briggs 2013), I defined a *quality* risk or *quality* change to the plan score as one that receives a 70 or greater score from the expert panel member for both elements. For example, a risk that received a grade of 90 for creativity and a 60 for plausibility would not be considered a *quality* risk. A change to the plan that received a 70 for creativity and a 70 for usefulness would be considered a *quality* change to the plan. I expected this data to support hypothesis testing and may be useful in understanding different levels of creativity between brainstorming and the premortem techniques for risk identification.

The second area that each expert panel member must consider was whether the risks identified by the team on its risk list match the four black swans from the Mark-14 case. I worded my description of the four black swans so that there was room for the expert panel members to exercise their judgment in determining if a risk matches one of the black swans. I did not use the expected risk statement "if, then" formula to describe the black swans. That would limit the expert panel member's flexibility.

The last task for the expert panel member was, after grading all twenty sets of deliverables, to place themselves in the senior decision maker's role. They identified their top three deliverable sets. I encouraged the expert panel member to abandon the creativity-plausibility and creativity-usefulness construct and apply their own criteria for quality based on their experience. This data provided another dimension to compare brainstorming and premortem.

My challenge was to design an expert panel scoring instrument that avoided having the expert panel member search through stacks of deliverables to perform the scoring and "top-3" rankings. To achieve this aim, I created a score sheet that allowed the expert panel member to see all twenty deliverables and all their scores in one glance. The grading sheet was printed on engineering design paper which is 36-inches by 24-inches in size. Figure 31 shows the layout of the Expert Panel Scoring Sheet.

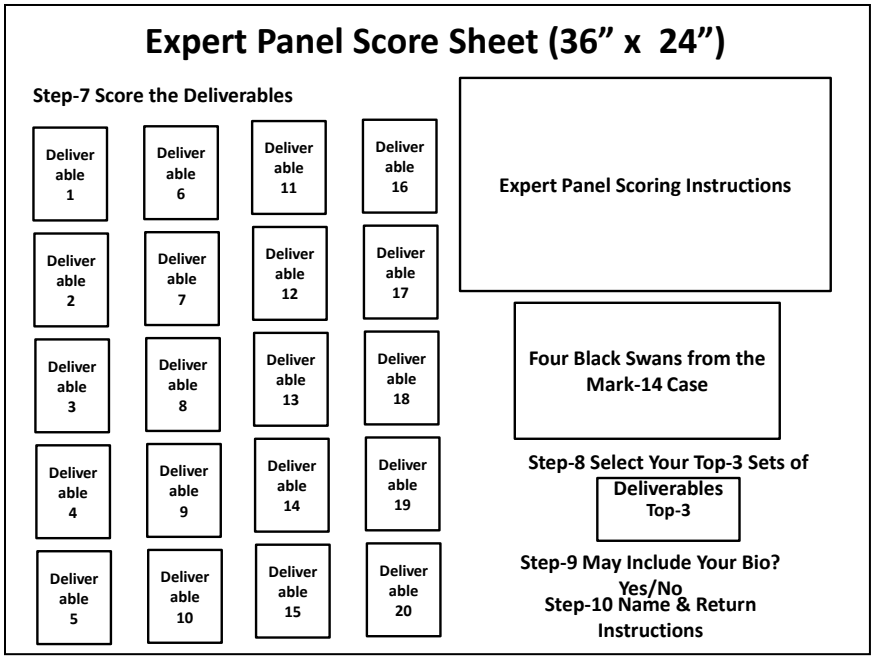


Figure 31 Layout for the Expert Panel Scoring Sheet

Accompanying the grading sheet was an instruction sheet (Figure 32), also printed on engineering design paper. The instruction sheet had all the information the subjects received in the experiment. This included the instructions to the subjects and the Marlin Upgrade Project Charter, which describes the technology insertion plan. The instruction sheet also had the description of the four black swans from the Mark-14 case and a calibrating example. The calibrating example showed a notional set of deliverables and grading. The grading has explanations for the grade assigned. Both the instructions sheet and grading sheet were arranged in a step-wise fashion. The steps begin on the instruction sheet with *Step 1: Read the Consent Form* and end on the grading sheet with *Step 10: Print Your Name.* My intention with this design was to facilitate consistency in an expert panel member's scoring from the first deliverable set to the last deliverable set.

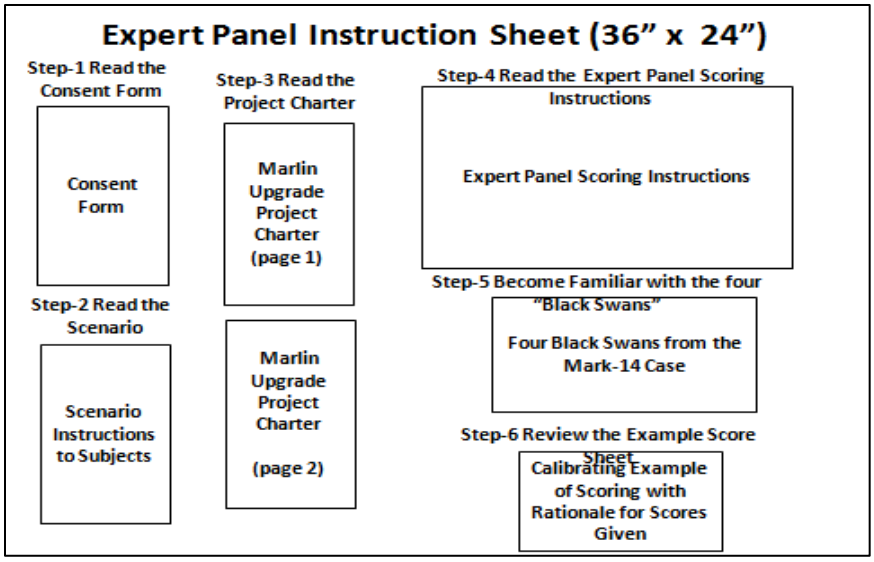


Figure 32 Layout for the Expert Panel Instructions Sheet

4.4.5 The Instructional System Design

The experimentation part of my research took place inside a Defense Acquisition University (DAU), Defense Systems Management College (DSMC) training workshop. I believe it is relevant to discuss the instructional system design used to construct the training portion of the workshop because the quality of the training may directly influence the quality and consistency of the experiments.

DSMC follows the ADDIE Model (Figure 33) for curriculum and learning asset development. ADDIE is an acronym for the framework of generic processes training developers use to build lessons and lesson content (Forest 2016): Analysis, Design, Development, Implementation and Evaluation.

Since the experiments occur inside a training workshop, DSMC developed the learning content using the ADDIE framework. This disciplined approach is an essential requirement for DAU and DSMC maintaining its Council on Occupational Education and International Association of Continuing Education and Training accreditations (DAU 2014). The following describes the ADDIE elements as they apply to the workshop curriculum.

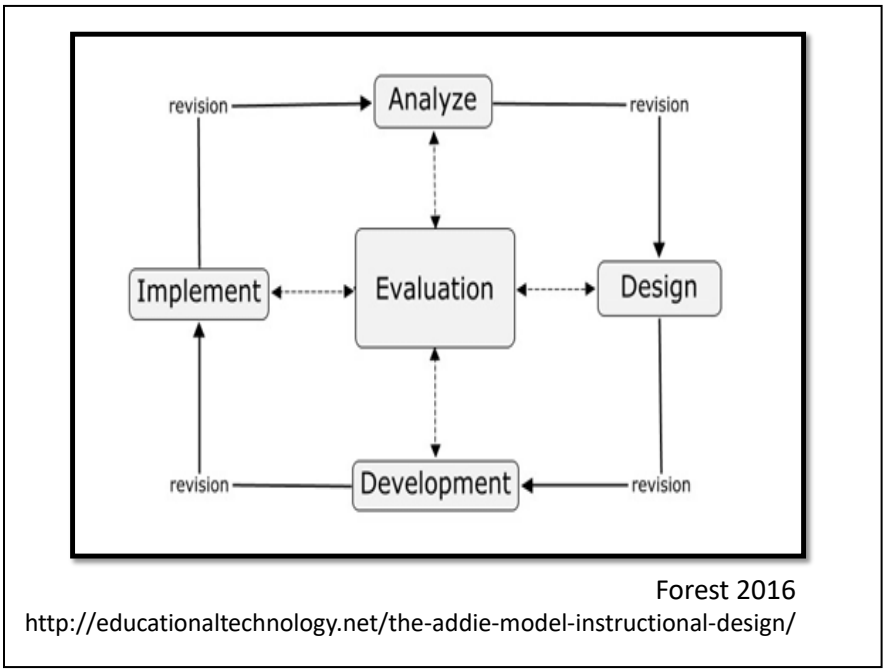


Figure 33 The ADDIE Framework for Instructional Design

The analysis phase clarified the instructional objectives, the learning environment, and the learners' existing skills. In the case of our workshop, I expected the learners to be experienced professionals, knowledgeable in both risk terminology and the brainstorming technique. However, I expected them to be unfamiliar with and untrained in the premortem technique. Therefore, the premortem technique required additional

learning content when compared to the brainstorming technique. The learning must occur within a two-hour block and advertised as a voluntary elective learning opportunity, outside scheduled class time. DSMC developed the workshop based on adult learning principles.

The adult learning principles in the design phase included motivational design and the use of process worksheets. Motivational design has four components: attention, relevance, confidence, and satisfaction (Merrill et al. 1996). In order to gain and retain the learner's attention, the workshop incorporated a range of media to include slides, whiteboard notes, handouts, and a video. The video is an excerpt from the National Geographic Channel program, *Challenger: The Untold Story*. The video highlighted the elements of deliberate ignorance that can make teams blind to the black swan (Everett and Mulholland 2006). The learners were also introduced to questions and problems regarding risk identification techniques to stimulate their curiosity. To make the learning content relevant, the learners were able to practice the technique in a technology insertion scenario that they can relate to their own experiences. The lesson developed confidence in the learner to apply the assigned risk identification technique in the experiment. That occurred through the use of a practice exercise that includes self-critique and instructor feedback. In order to provide the learner with a sense of achievement, the learners applied their newly learned skill immediately. Learners received a process worksheet that describes the phases of their assigned technique. The process worksheet was included with the experiment instructions. The use of process worksheets has shown to improve learner performance (Kirschner, Sweller, and Clark 2006).

The DSMC peer review process applies to the development phase of the workshop curriculum. The purpose of the peer review is to generate real-time feedback of the delivery expertise and the quality of the learning products (slides, handouts, exercises, and videos) (DSMC 2012). The peer review is also a forum for faculty to share best practices. DSMC policy requires a peer review for all new lessons and learning materials. The peer review process may occur over several sessions and varying levels of detail. For example, it is common practice for the lesson to receive a peer review at the *storyboard* level. After corrective actions are made, the lesson may be peer reviewed as a *walkthrough*. The *walkthrough* may be followed by additional corrective actions and a *full lesson* peer review. At the conclusion of the full lesson peer review, the instructor self-critiques followed by peer critiques and suggestions. In the implementation phase of the ADDIE framework, instructors receive training, and lesson materials are prepared. In the case of our workshop, instructor training and materials refinement occurred as part of *pilot testing* discussed later in this chapter. Evaluation of the workshop curriculum continued as part of DSMC's peer review process for existing learning assets. That process requires twice annual peer reviews during the actual teaching event (DSMC 2012).

4.4.6 Peer Reviews/Pilot Tests and Mock Expert Panel

In this section, I will describe the processes I followed to validate the instructional design and the data collection instruments. The process included one *storyboard* peer review (Peer Review #0); four full lesson peer reviews, which also served as my pilot tests (Peer Review #1, #2, #3, and #4); and a mock expert panel review.

I held Peer Review #0 on 6 August 2015 with the Director of DSMC Executive Courses and two other faculty members. The purpose of that peer review was to storyboard the lesson concept and the viability of using two variants of the lesson. One variant of the lesson would be focused on the premortem with no instruction and practical exercise regarding brainstorming. I would present that lesson content to teams performing the premortem technique in the experiment portion of the lesson. The second variant would provide the teams with instruction on the brainstorming technique and a brainstorming practical exercise before the teams perform brainstorming in the experiment. The second lesson variant would conclude with premortem instruction, a premortem practical exercise and a wrap-up discussion comparing and contrasting brainstorming and premortem for risk identification. The peer review members agreed that both variants were feasible, acceptable, and suitable approaches for the purposes of enabling my A/B experimentation method. The peer review participants were concerned that the lesson would require more than the allotted two hours. My action from Peer Review #0 was to practice the delivery of both lesson variants while being timed and make the necessary modifications to the lesson content to fit within a two-hour limit prior to Peer Review #1.

I held Peer Review #1 on 3 September 2015 with five DSMC faculty members. The purpose of the peer review was to present the premortem variant lesson with the five faculty members playing the role of learners/test subjects. The peer review also served as a pilot test for the data collection methods and instruments. At this point in the lesson and experimental design development, I provided the learner/test subjects with a detailed breakdown of the forty-five-minute experiment. That breakdown included three separate

fifteen-minute timed blocks. The first fifteen minutes were intended to provide the team with time to plan and organize to execute their assigned technique. The second 15-minute block was to collaborate while executing the technique. It was during this block that I performed structured observation. The final fifteen-minute timed block directed the team to develop the risk list and changes to the plan.

I identified a number of shortcomings in my employment of the structured observation method and the structured observation data collection instrument (Figure 34). First, I found that collaboration occurred during the entire forty-five-minute experiment. My data collection instrument only covered fifteen minutes. Second, the structured observation data collection instrument divided the fifteen minutes into one-minute increments. It was too difficult for me to keep up with the collaborative behaviors for five subjects over the four dimensions (risk idea offered, difference in perspective offered, an enrichment offered, and a critique given) while also keeping track of the time in one-minute intervals.

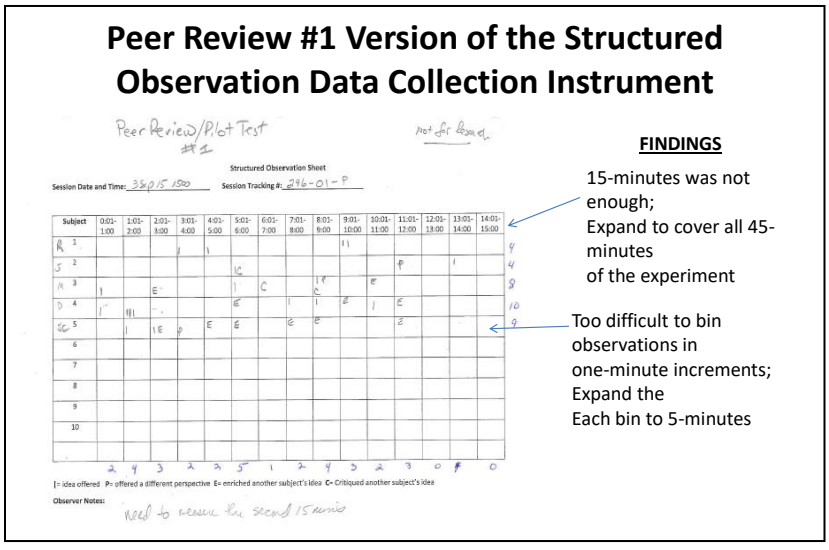


Figure 34 Peer Review #1 Structured Observation Findings

For Peer Review #1, the team's butcher paper notes served as the data collection instruments for the *risk list* and the *changes to the plan* deliverable. At the conclusion of the experiment, it was not clear to the team nor to me which of the many bullets and notes were the risks and changes to the plan should go to the expert panel for review. The subject survey contained ten statements for *Likert* response. At the conclusion of the survey, the subjects placed their survey in a plain white envelope. The envelope was placed on a separate table. This concluded the experiment and pilot testing. The total time for the lesson, experiment, and survey was one hour and fifty-two minutes. The peer review panel approved the lesson and experiment content. They suggested providing more flexibility for the team to manage their time during the experiment. Three separate fifteen-minute blocks were too rigid. However, they believed the forty-five-minute time to perform the technique and create the two deliverables was achievable and applied adequate time pressure. The peer review panel recommended providing the team with a standard deliverable template that they should complete by hand or electronically. This would force the team to clearly form their *risk list* and *changes to the plan* for expert panel scoring. They also suggested having the team provide their top four risks and changes to the plan in order to give the team time to prepare the deliverable template. With regard to the subject survey, the peer review panel believed the selection options and *Likert* scale were straightforward and that my reading through of the survey questions and responses facilitated the response selection process. The peer review panel suggested expanding the survey to allow the subjects to measure the technique they

applied in the experiment with another technique they had used in their professional experience.

The peer review panel agreed that *Challenger: The Untold Story* video captured and retained their attention. The peer review panel members also said the exercises and experiment scenario were relevant and timed appropriately. Two of the five panel members suggested I consider expanding the workshop by an additional thirty-minutes to allow more time for the learner/subjects to practice and receive feedback on their assigned technique. I asked the peer review panel if my structured observation data collection distracted them from performing the experiment or influenced their performance in any way. They responded that they had not noticed my data collection. I asked the peer review panel members if audio or video recording the experiment would improve the accuracy of my structured observation data collection. All five said that recording the experiment may improve accuracy but would definitely influence team behavior. Three of the peer review panel members provided anecdotal examples of how audio and video recording influenced student behavior in other lessons. I reviewed the completed structured observation sheet and asked the peer review panel members if they agreed with the recorded level and dimensions of their collaboration. All believed that the structured observation sheet was reasonably accurate.

I then asked the peer review panel members if I had given them adequate opportunity to opt out of taking the survey and anonymity when they completed the survey. They responded positively to both. One peer review panel member commented that even though she was in a small group, she could have easily submitted a blank survey if she did not want to participate. She said that this was possible because I was

occupied reading the survey options and responses. Another peer review panel member commented that my reading the survey encouraged him not to race through the survey. I asked the peer review panel members if they knew the real world case from which the technology scenario experiment was derived. They did not.

Peer Review #1 concluded with a walk through of the brainstorming lesson variant. The peer review panel believed that the brainstorming lesson variant was adequate and would require no further review. Based on my observations and feedback from the Peer Review #1 panel, I validated the following:

- The lesson materials adequately prepared the subjects to exercise their assigned risk identification technique.
- The technology insertion scenario in the experiment was interesting and relevant to the target audience.
- The forty-five-minute allocated time for the experiment was sufficient.
- The structured observation coding method was reasonably accurate in recording the level and nature of the subject's collaboration.
- The structured observation method did not influence subject behavior.
- The subject survey and *Likert* scale is appropriate.
- The subjects can opt out of the survey and the process allows for subject anonymity.

Based on my observations and feedback from the Peer Review #1 panel, the following items were changed for Peer Review #2:

- Adjusted the timed portion of the experiment to be more flexible (not divided into three timed increments) while maintaining the 45-minute total time for the experiment.
- Planned to perform structured observation over the entire 45-minute experiment.
- Changed the structured observation time intervals from one-minute to five-minutes.
- Developed templates for the deliverables that allow the team to focus on their top four risks and changes to the plan.
- Added additional response items to the survey that allowed the subjects to measure their opinion of the technique applied in the experiment against another technique they have used in their experience.
- Divided the survey form into three parts with specific instructions to the subject regarding the change in direction of the statements for each part.

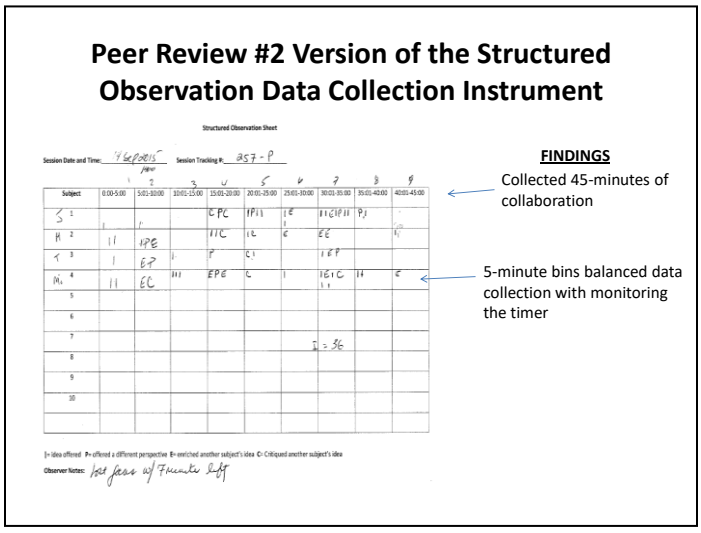


Figure 35 Peer Review #2 Validated Changes to the Structured Observation Data Collection Instrument

I held Peer Review #2 on 14 September 2015 with four DSMC faculty members. The lesson, experiment, and survey were completed in two hours and three minutes. The focus of Peer Review #2 was to validate the changes made from Peer Review #1. This included pilot testing of the revised structured observation data collection instrument, the deliverables template, and the survey data collection instrument with the additional response section. The peer review panel members for Peer Review #2 had not participated in Peer Review #1. They independently confirmed the Peer Review #1 comments regarding the relevance and appropriateness of the lesson, the practical exercise, and the technology insertion scenario for the experiment. I noted that coding the observed behavior during structured observation was easier to do in five-minute increments even though I was now coding for the entire forty-five minutes of the experiment (Figure 35). At the conclusion of the experiment and survey collection, I asked the peer review panel about the deliverables template. They indicated that four risks and four changes to the plan are achievable in the allotted time and that the template (Figure 36) enabled the team to focus. They suggested that I provide a paper copy of the deliverables template to each of the team members/subjects. This could allow the team the option to divide the task of preparing the *risk list* and *changes to the plan* deliverable. Peer Review #2 panel members confirmed the Peer Review #1 belief that my method of structured observation reasonably recorded their individual collaborations and did not influence their performance during the experiment. I did note that having performed structured observation for a team size of five in Peer Review #1 and a team size of four in Peer Review #2, the largest size team that I could accurately perform structured

observation would be a team size of six. Peer Review #2 panel members suggested adding a process worksheet to the bottom of the *Instructions for Subjects* form. The process worksheet would simply list the steps the team would need to perform to complete their assigned risk identification technique.

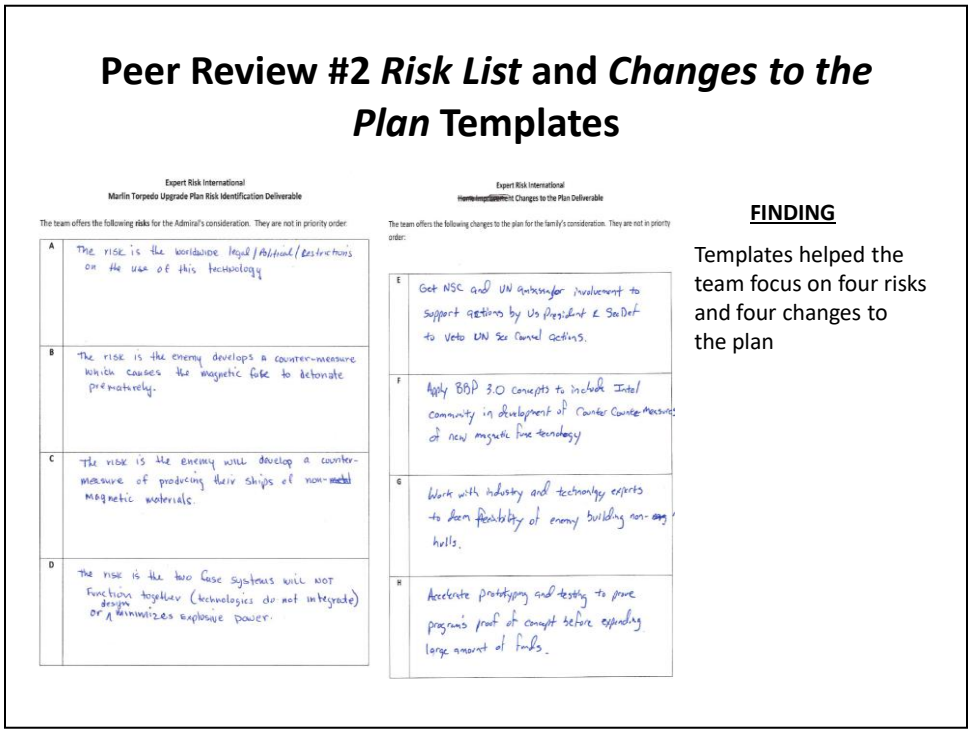


Figure 36 Peer Review #2 Validated the Benefit of Providing Templates for the Team Deliverables

Peer Review #2 validated the following:

- Structured observation is a reasonably accurate method to record collaborative behaviors over the entire forty-five-minute experiment.

- Dividing the structured observation data collection instrument into five-minute increments significantly reduced the attention required to monitor the timer.
- The additional data collection area on the survey was acceptable to the subjects.
- The deliverables template enabled the team to focus on preparing the four risks and four changes to the plan.

With the data collection instruments in stable form, the focus for Peer Review #3 was to improve my personal delivery of the lesson and to continue practicing the structured observation method. I took the following actions from Peer Review #2 to prepare for Peer Review #3:

- Provide a set of deliverable templates for each team member/subject.
- Add the process worksheet (execution steps) to the bottom the *Instructions to Subjects* form. Only the process for the assigned risk identification technique should be visible. The process for the other technique was masked.

Peer Review #3 was held on 23 September 2015 with three faculty members. Those faculty members had not participated in either of the previous peer reviews. The purpose of Peer Review #3 was to validate the benefit of having a set of deliverables for each team member/subject and the benefit of the process worksheet at the bottom of the *Instructions to Subjects* form (Figure 37). Peer Review #3 would provide additional practice in the data collection techniques and offer new insights not gained in the

previous peer reviews. The lesson, experiment, and survey were completed in one hour and fifty-three minutes. Peer Review #3 provided insight into running the experiment with three team members/subjects. Because the team is self-facilitating through its assigned risk identification technique, the peer review panel members noted that with three team members, the collaboration could be too dependent on the skill of the team leader in facilitation. Therefore, the peer review panel members recommended the team should not be smaller than four persons. The team did use their individual deliverable templates to divide the task of preparing the risk list and changes to the plan. This involved transcribing and elaborating on notes the teams had made collectively on butcher paper and whiteboard. The team collectively reviewed and adjusted their individual write-ups before submitting the deliverable and ending the experiment. When I asked the team if they found the process worksheet at the bottom of the *Instructions to Subjects* form helpful, all three members agreed that it was a beneficial resource to stay on task.

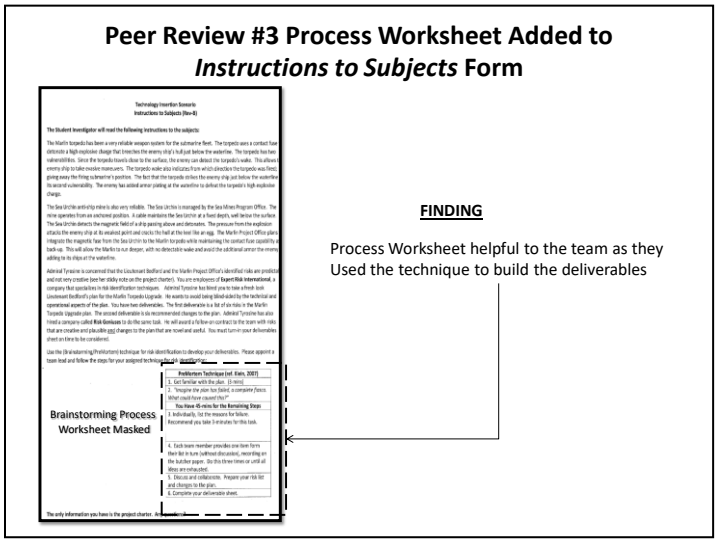


Figure 37 Peer Review #3 Validated the Benefit of Providing a Process Worksheet at the Bottom of the Instructions to Subjects

Peer Review #3 validated the following:

- Minimum team size should be four persons.
- Individual deliverable sheets helped the team collaborate and submit the deliverables.
- Insertion of the process worksheet at the bottom of the *Instructions to Subjects* form helped the team stay on task.

No changes were made to the lesson content or data collection instruments between Peer Review #3 and Peer Review #4. The purpose of Peer Review #4 was to gain additional practice in the data collection methods and uncover insights not revealed in the previous peer reviews.

Peer Review #4 was held on 9 October 2015 and included five faculty members that had not been involved in the previous peer reviews. The lesson, experiment, and survey were completed in one hour and fifty-seven minutes. The peer review panel confirmed the findings from Peer Review #3. One panel member suggested modifying the practice exercise. The practice exercise required the team members to advise a family on the risks in a home improvement project. The peer review panel member suggested that the scenario could be viewed as reinforcing gender stereotypes. I agreed to modify the practice exercise. The peer review panel suggested that the change to the practice exercise scenario could be made without further peer review. The peer review panel members unanimously agreed that the lesson, experiment, and data collection methods and instruments were ready to be presented to learners/subjects.

The peer review/pilot testing process allowed me to incrementally improve the workshop and experiment delivery. It allowed me to address several deficiencies in the data collection methods and instruments. Although all of the above could be further improved over time, the methods for eliciting information relative to Null Hypothesis #1: *Team collaboration using the premortem technique is equal to team collaboration using brainstorming*, proved to be satisfactory. I then turned my attention to validating the methods for eliciting information relative to the remaining three null hypotheses, the expert panel.

In order to test the modified Delphi technique or expert panel data collection method and data collection instruments, I held a *mock* expert panel. The panel consisted of four faculty members that had not been part of the peer review process. I was limited to four expert panel members due to faculty availability. I also wanted to avoid inducing survey fatigue among the pool of available experts (Meyer and Booker 1991). Recall that I modified the classic Delphi method by performing a single round of evaluations without the intention of driving the panel members to consensus.

For the mock expert panel, I prepared three sets of mock deliverables. The mock deliverables contained *risks* and *changes to the plan* I created with varying levels of creativity, plausibility, and usefulness. I had designed some of the risks to be *black swan hits*, risks that matched the risks the Mark-14 team missed. I used the deliverables template validated during the peer review/pilot testing process with one exception. The deliverables for the mock expert panel contained six *risks* and six *changes to the plan* instead of four each. The additional *risks* and *changes to the plan* provided 50 percent more data points for analysis. The *risks* and *changes to the plan* were handwritten on the

deliverable templates. I also provided the expert panel members direction to identify their top set of deliverables of the three sets.

The expert panel members worked independently and apart from each other and were directed to grade all elements on a scale of 0 to 100. Risks were scored on two dimensions: creativity and plausibility. Risks were to also be considered for a *black swan hit*. *Changes to the plan* were to be scored on two dimensions, creativity and usefulness. The expert panel members received the scoring package electronically. They received three *Adobe Acrobat* files containing the mock deliverables. They also received an *Excel* workbook with a set of instructions, definitions, and the four black swans from the Mark-14 case on one sheet and a scoring form for the three sets of deliverables on another sheet. The mock expert panel members were given five days to return the Excel workbook via e-mail with the deliverables scored.

I expected the mock expert panel to prove the following:

- Exchange of materials, ungraded packages going out and scored packages returning, can be done electronically.
- The mock expert panel member scores for the risks and changes to the plan would closely match my expected scores.
- The mock expert panel black swan hits would closely match my expected black swan hits.
- The mock expert panel selection of the top set of deliverables would match my selection of the top set of deliverables.

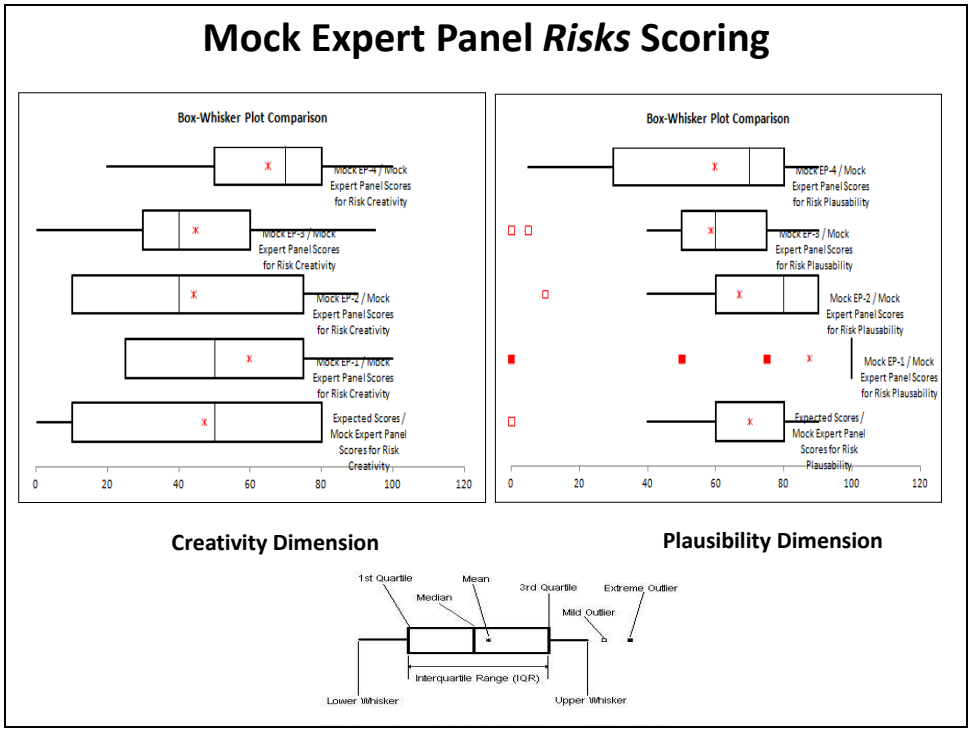


Figure 38 Mock Expert Panel Member Scoring for Risk Dimensions

The four mock expert panel members reviewed the three mock deliverables sets from 26 through 30 October 2015. The box-whisker plots in Figure 38 indicate that the mock expert panel members' scores interquartile ranges range (IQR) for scoring the Risk Creativity dimension closely matches the IQR for the expected scores. A one-way Analysis of Variation (ANOVA) measured the degree of variation in scoring between the mock expert panel members and the expected scores for Risk Creativity dimension as $F = 2.13$, $p = 0.0839$. On average, the mock expert panel members scored the Risk Creativity dimension 6.1 points higher than the expected scores.

The Risk Plausibility dimension is much less calibrated. Considering the box-whisker plot for this dimension, clearly Mock Expert Panel Member #1 is not calibrated. A one-way Analysis of Variation (ANOVA) measured the degree of variation in scoring between the mock expert panel members and the expected scores for Risk Plausibility dimension as $F = 3.43$, $p = 0.119$. On average, however, the mock expert panel members scored the Risk Plausibility dimension just 1.8 points lower than the expected scores.

Evaluating the Changes to the Plan in the Creativity dimension, the box-whisker plot in Figure 39 shows some calibration between the mock expert panel members and the expected scores. A one-way Analysis of Variation (ANOVA) measured the degree of variation in scoring between the mock expert panel members and the expected scores for Changes to the Plan Creativity dimension as $F = 1.72$, $p = 0.1525$. On average, the mock expert panel members scored the Changes to the Plan Creativity dimension 3.4 points higher than the expected scores.

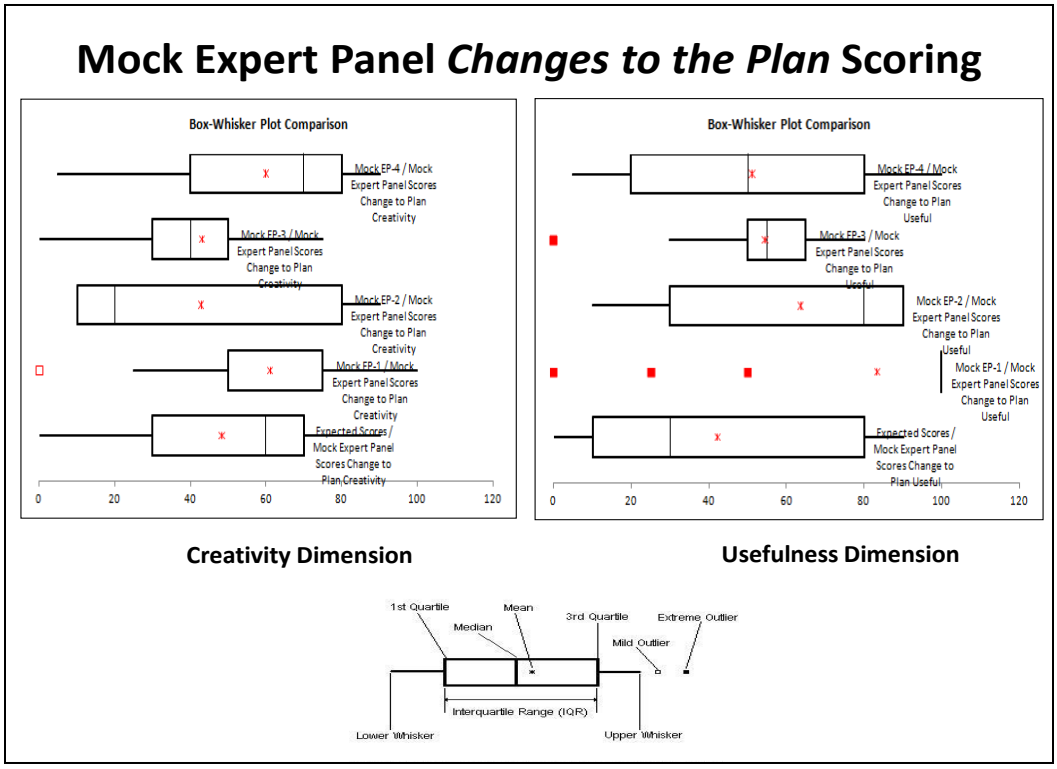


Figure 39 Mock Expert Panel Scoring for Changes to the Plan Dimensions

The box-whisker plot for the Usefulness dimension of the Changes to the Plan once again shows Mock Expert Panel Member #1 as being not calibrated in relation to the expected scores and the other mock expert panel members. A one-way Analysis of Variation (ANOVA) measured the degree of variation in scoring between the mock expert panel members and the expected scores for Changes to the Plan Usefulness dimension as $F = 4.70$, $p = 0.0018$. On average, the mock expert panel members scored the Changes to the Plan Usefulness dimension 20.9 points higher than the expected scores. I attribute the skewed results to Mock Expert Panel Member #1's scoring.

I expected four black swan hits in the 18 risks identified in the three Risk Lists. Mock Expert Panel Member #1 matched three of the four expected black swan hits, with no risks misidentified as black swans. Mock Expert Panel Member #2 correctly matched two of the four expected black swan hits but misidentified one risk as a black swan hit. Mock Expert Panel Member #3 did not correctly match any of the expected black swan hits and misidentified one risk as a black swan hit. Mock Expert Panel Member #4 matched two of the four expected black swan hits, with no risks misidentified as black swans.

Finally, I expected the mock expert panel members to select Deliverable Set #259-01 as the top deliverable set for overall quality. Mock Expert Panel Member #2 met this expectation. The remaining three mock expert panel members did not provide a response. I interviewed each of the mock expert panel members for their perspectives, insights, and recommendations.

Mock Expert Panel Member #1 admitted to waiting until the last minute to score the deliverables. He also commented that he felt confused and unclear by the definition of the Plausibility dimension. Mock Expert Panel Member #2 reported that having to transition from reading the deliverables in paper format (he printed the Adobe Acrobat documents) and inputting scores in an Excel worksheet was clumsy and may have impacted his scoring. The mock expert panel member recommended printing the deliverables and score sheet in a binder for each panel member. Mock Expert Panel Member #3 also reported being fatigued by having to switch back and forth from files. Mock Expert Panel Member #4 reported that he found it difficult to score consistently between the deliverable sets. He also reported that he found it difficult to read the

handwriting on the deliverables sheets. Mock Expert Panel Members #1, #3, and #4 reported that they missed the requirement to identify their top deliverable set. All of the mock expert panel members suggested that I provide the context related to the scenario and not just the deliverables.

Here are my findings from the mock expert panel:

- The deliverables and score sheets cannot be exchanged or graded electronically.
- The expert panel members can experience fatigue while scoring, resulting in a deterioration of consistency in scoring between deliverables.
- An expert panel member can skew the results if they are not calibrated.
- The instructions to score the deliverables must be very clear.

The following are the actions I made to the expert panel materials and process based on the mock expert panel:

- I recruited 30 expert panel members to score the deliverables. The large number was expected to mitigate the effects of an expert panel member that is not calibrated.
- I provided a calibrating example in the scoring instructions. I provided rationale for the scores provided in the calibrating example. I also stated in the instructions that the expert panel member should set 70 as the cut-off for quality on the 0 to 100 scale. I also ensured that the definitions for the dimensions (creative, plausible, and useful) were clearly stated on the instructions sheet. All of this was provided to the expert panel members on a single engineering sheet with dimensions 36-inches by 24-inches.

- I provided all of the expert panel members with the entire context materials that I had provided to the subjects during the experiment.
- To minimize expert panel member fatigue, I typed all of the handwritten deliverables and placed all the deliverable sets on a single engineering sheet with dimensions 36-inches by 24-inches. This format also addressed the challenge of the expert panel member providing consistent scoring across all of the deliverable sets. The expert panel member had all of the deliverables and their scores in their field of view.
- I made it clear that I am available to clarify instructions or definitions but cannot provide any insight into the risks or changes to the plan in the deliverable sets.
- I provided step-wise instructions so that expert panel members would not miss any items.

The peer review/pilot testing process allowed me to refine the workshop content as well as the materials supporting the experiment. During the course of four peer reviews/pilot tests, I received feedback from the subjects (experienced faculty members), validating my structured observation technique for data collection. I found, and peer review feedback confirmed, that limiting the size of the teams to four, five, or six persons would allow me to evenly distribute my attention across the subjects while I recorded the behaviors of interest during the experiment. Recording the observed behaviors immediately using paper and pencil mitigated selective memory bias while avoiding the *Hawthorne effect* of audio or video recording (Robson 2011). Reviewing the results of the structured observation that I had coded on the structured observation data collection

sheet with the subjects after the experiments confirmed that the structured observation method adequately recorded the team's collaboration.

The peer reviews/pilot tests refined the subject survey data collection instrument. The process validated the wording of the questions as well and the use of the six-point *Likert* scale. The peer review feedback made me aware of the missed opportunity to ask the subjects to measure their experience using their assigned technique in the experiment against another technique they had previously used. This resulted in a revision of the survey data collection instrument.

The mock expert panel enabled a number of refinements in how the expert panel would receive the team deliverables for scoring. The mock expert panel made me aware of potential pitfalls, such as one expert panel member skewing results. This led me to implement mitigation steps such as recruiting a large pool of expert panel members and providing a calibrating example. The peer review/pilot testing and mock expert panel processes validated my data collection methods and provided valuable practice in employing those methods. The data from the peer review/pilot tests and mock expert panel also allowed me to practice the data analysis methods described in the next section.

4.5 Data Analysis Methods

Three primary sources of data support my research plan. The structured observation sheet measured the dependent variables regarding the level of team collaboration while performing the assigned treatment (brainstorming or premortem). The subject surveys provided demographic information and data that may confirm or contradict data obtained through structured observation. The expert panel scoring

measured the dependent variable related to the quality of the *risks* and *changes to the plan* the teams produced. The expert panel also measured the dependent variable regarding the number of black swan hits the teams report in the risk list. Finally, the expert panel provides a "top three" ranking of deliverable sets. All of the data were analyzed in a typical Treatment-A versus Treatment-B fashion. However, the data can be further analyzed comparing team size, experience level, and organization type (government or industry). The following provides an overview of the data analysis methods I plan to use and their expected application. Except where noted, I performed all of the data analysis using *Stat Tools-7* from *Pallisades*.

4.5.1 Lilliefors Test for Normality

Parametric testing assumes that the data under test follows a normal distribution. The most common parametric tests are the z-test, t-test, and the Analysis of Variation (ANOVA). The distribution of data measuring team collaboration using brainstorming and the premortem techniques is unknown. Although the brainstorming of ideas has been observed to follow an ogive or bullet-shaped distribution (Briggs and Reinig 2010), normality for both techniques cannot be assumed. Prior to applying a parametric test, I first checked the distribution of the data for normality. The chi-squared goodness-of-fit fit test is a common test for normality. However, the Lilliefors test is more likely to detect non-normality if it exists (Palisade 2015) and is appropriate for small sample sizes (Lilliefors 1967). When using the Lilliefors test for normality, if the Lilliefors test statistic is greater than the Lilliefors critical value, we reject the normality assumption. In that case a nonparametric method can be applied to test the null hypothesis.

4.5.2 Mann-Whitney Test

Nonparametric tests make no assumption about the type of underlying distribution. However, some nonparametric tests do make assumptions about the general shape of distribution. For example, the Wilcoxon signed-rank test assumes the distributions are symmetric. The Mann-Whitney test does not make any assumption about the shape of the distribution (Palisade 2015). There are two forms of the Mann-Whitney test. One form, the Median form assumes the two distributions under test have the same shape. The General form does not make that assumption. The General form denies that either probability distribution tends to generate smaller values than the other. I used the General form of the Mann-Whitney. As a precedent, Reeves, Eveleigh, Holzer, and Sarkani used the Mann-Whitney test in their study of risk biases at the National Aeronautics and Space Agency (Reeves et al. 2013). To add additional validity to nonparametric test results, the observed data can be transformed into a normal distribution to facilitate a parametric test.

4.5.3 Box-Cox Transformation

A parametric hypothesis test can be applied to empirical data that is not normally distributed if it is transformed to a normal distribution. The Box-Cox transformation is a widely used method for this purpose (Sakia 1992). Box-Cox is a logarithmic transformation so "1" should be added to each observed value to eliminate any "0" values in the data set. This is necessary because the log of "0" does not exist (Wessa 2015). The Box-Cox transformation of the variable x is indexed by λ, and defined as

$x'_\lambda = \frac{e^{\lambda \log(x)} - 1}{\lambda}$ (Scott 2016). Several computer programs can perform the Box-Cox transformation. I used an online tool produced by Patrick Wessa PhD, Professor of Statistics and Information Science at the University of Basel in Switzerland. The tool can be accessed at http://www.wessa.net/rwasp_boxcoxnorm.wasp/. The Box-Cox transformation has shown to be robust to non-normality as long as the data has reasonable symmetry (Sakia 1992). However, it could be limited in situations where extreme outliers exist. Parametric tests can be applied once the observed data has been transformed to a normal distribution. This can be a confirmation of the hypothesis test results using the nonparametric Mann-Whitney test.

4.5.4 Parametric Tests

I expected to draw inferences from two independent samples, brainstorming and premortem. In the cases where I compared aggregated results from the twenty expected experiments (10 samples from the brainstorming and 10 samples from premortem) and an unknown population standard deviation, I applied the student's t-distribution for the hypothesis testing. In the cases in which I am comparing the individual subject data for each treatment, I can assume a large sample size. Applying my rule of at least four subjects per team and my goal of ten experiments per treatment, I achieved a sample size of at least 40 per treatment. A z-test may apply in this case. In the cases in which I want to compare variations between the different team sizes (teams with four, five, or six members), I used a One-Way ANOVA to determine the cause of the variation in the data. The ANOVA can also be used when the sample sizes are not equal (Ott and Longnecker 2010).

4.5.5 Goodman and Kruskal's Gamma

I expected there would be cases in which I want express the strength of association between two ordinal variables. For example, is there a strong association between teams that have high levels of collaboration and those teams that received high scores for the quality of their deliverables? Goodman and Kruskal's Gamma expresses the strength of that association between two variables (Elm and Goldenson 2012). Gamma values near "1" indicate a very strong supporting relationship between the variables. A Gamma value near "-1" indicates a very strong opposing relationship. A Gamma of "0" would indicate that a relationship between the variables does not exist.

Gamma can be computed in *Excel* by first developing a contingence table of the two variables being compared by the desired ordinal relationship (i.e. high, medium, low). The contingency table compares the number of observations that are in agreement and the number of observations that are in disagreement (or inversion). An example of an agreement is *Team Collaboration* was *high* and *Quality of the Deliverable* was also *high*. An example of an inversion is *Team Collaboration* was *high* but *Quality of the Deliverable* was *low*. Gamma is expressed as $Gamma = \frac{Na-Ni}{Na+Ni}$ where *Na* is the number of agreements and *Ni* is the number of inversions. Gamma can be tested for statistical significance using a t-test with the formula: $t \approx \sqrt{\frac{Na+Ni}{N(1-Gamma^2)}}$ where *N* is the number of observations (Goodman and Kruskal 1963). If the value of *t* is greater than the critical value of *t*, the result is statistically significant.

4.6 Summary of the Research Methodology

In this chapter, I described the methods I used to provide evidence of the benefits of using the premortem technique for risk identification in a systems engineering context, specifically a brownfield/technology insertion scenario. Using the classic Treatment-A versus Treatment-B approach, I measured the performance of the premortem technique against the performance of the brainstorming technique. The brainstorming technique is a worthy comparison because it is a popular and simple technique that happens to be the logical opposite (an inductive technique) of the premortem (a deductive technique). My research method used expert elicitation from two sources. The subjects of the experiments were experienced professionals from government and industry. The expert panel members were faculty from DSMC with a wide range of subject matter expertise and professional backgrounds.

Data collection occurred through the use of structured observation of the team's level of collaboration during the experiments and a survey from the subjects participating in the experiment. The subject teams produced a risk list and a set of changes to the plan that the modified Delphi (expert panel) scored for quality and black swan hits. In order to score black swan hits, the scenario in the experiment was based on a historical technology insertion project in which the team missed identifying risks due to blind spots or deliberate ignorance. The Mark-14 torpedo upgrade case fits the scenario requirements and is not widely known.

The design of experiment is based on A/B test design. The minimum and desired number of experiments per treatment (four and ten respectively) exceeded the guidelines in Mead's method for experiment size. Since the experiment occurred within the context of DSMC workshop, the quality of the experiment may be directly influenced by the

quality of the workshop lesson materials. To ensure quality, the workshop lesson materials and instructor certification followed the DAU/DSMC instructional systems design process. That process also served as a method to pilot test the data collection methods and instruments to be used on and by the subjects. A mock expert panel validates the modified Delphi technique.

Five peer review/pilot tests and one mock expert panel were held over the course of three months. Those activities improved the workshop lesson materials, data collection methods, and instruments. The peer review/pilot tests and mock expert panel allowed me to practice the data analysis techniques I used on the real experiment and expert panel results. Those techniques included the Lilliefors test for normality, Mann-Whitney nonparametric hypothesis tests of non-normal data, the Box-Cox transformation of non-normal data to normality, parametric hypothesis tests of normal data (t-test, z-test, ANOVA), and Goodman and Kruskal's Gamma test for the strength of association between two ordinal variables. With the validation of all elements of the research plan and approval from the George Washington University, Office of Human Research to proceed with the study (#111508), I conducted experiments and held the expert panel from November 2015 through March 2016.

Chapter 5 - Results

As systems get larger and more complex, we will face problems that that we have never encountered before. Traditional risk identification techniques may not be robust enough to deal with the growing complexity of 21st-century systems engineering. This chapter presents the results of experiments designed to find out if the premortem technique is better than brainstorming at identifying the atypical risks. Given a technology insertion scenario, does the premortem technique improve team collaboration, resulting in better risk identification and changes to the plan as compared to a team performing brainstorming?

5.1 Demographic Data and Response Rates

I performed 21 experiments. Ten of those experiments involved teams using the premortem technique and 11 teams using the brainstorming technique for risk identification. One of the teams using the brainstorming technique violated protocol when one member of the four-person team had to leave mid-experiment. This brought the team below the minimum number of four-persons required for an experiment. Although I continued the experiment and collected data from the experiment, that data has not been included in the analysis. That team's deliverables were not presented to the expert panel for scoring. I met the desired goal of performing ten experiments for each of the treatments.

Seven of the twenty experiments were performed in an academic environment with five of those experiments performed at DSMC and two of the experiments performed at the Eisenhower School. Thirteen of the twenty experiments were

performed as part of DAU's mission assistance and outreach to organizations. Seven of the 13 experiments involved government organizations (two from the US Navy's PMS-515 Future Frigate program and five from the US Coast Guard). The remaining experiments involved organizations from the Textron Corporation.

Top-Level Demographics

Organization	# of Experiments Conducted	# of Experiments Using Premortem	# of Experiments Using Brainstorming		# of 4-Person Teams	# of 5-Person Teams	# of 6-Person Teams
DSMC Students	5	2	3		0	2	3
Eisenhower School Students	2	1	1		2	0	0
PMS-515 Future Frigate Program	2	1	1		2	0	0
US Coast Guard Programs	5	3	2		2	0	3
Textron Corporation	6	3	3		2	1	3
TOTAL	**20**	**10**	**10**		**8**	**3**	**9**

Table 3 Subject Teams by Organization, Treatment, and Team Size

Team sizes varied between four, five, and six persons per team. Teams of six persons accounted for nine of the twenty experiments. Teams of four persons accounted for eight of the twenty experiments. Teams of five persons accounted for the remaining three experiments. Random assignment was used to assign teams to their respective treatments. No consideration to organization type or team size was given in the process of assigning treatments. A compilation of the top-level demographics is in Table 3.

A total of 101 subjects participated in the experiments. Fifty-one of the subjects used the premortem technique, and fifty subjects used the brainstorming technique. The detailed subject demographics are based on the responses the subjects gave to the demographics section of the subject's survey. One subject from the premortem treatment did not complete a survey. Therefore, the survey data is based on fifty surveys from each of the two treatments for a total of 100 surveys (a 99 percent survey response rate).

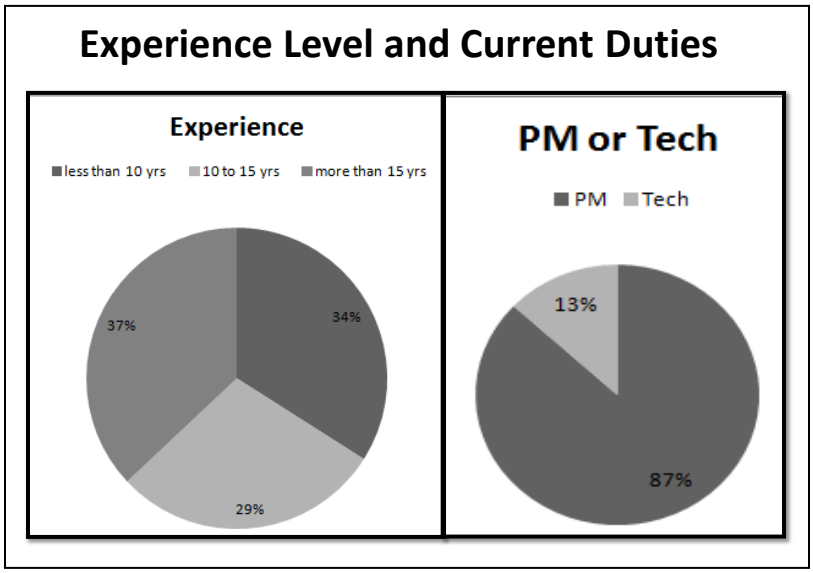


Figure 40 Subject Experience Level and Duties

Within the demographics portion of the survey, the subjects were asked to rate their level of experience at performing risk identification in a program management or systems engineering setting and whether they would describe their current job as either programmatic or technical (Figure 40). The subjects were also asked to select the risk identification technique they have used most frequently. The subjects were given nine alternatives to select from including brainstorming, premortem, and "other."

The survey data showed that 67 percent of respondents had ten or more years of experience in performing risk identification in a program management or systems engineering setting. The majority of the respondents (87 percent) considered their current job to be programmatic. The respondents reported that their most frequently used technique is brainstorming (45 percent) followed by Pros/Cons (24 percent) (Figure 41). Only one respondent reported that the premortem was his/her most frequently used risk identification technique.

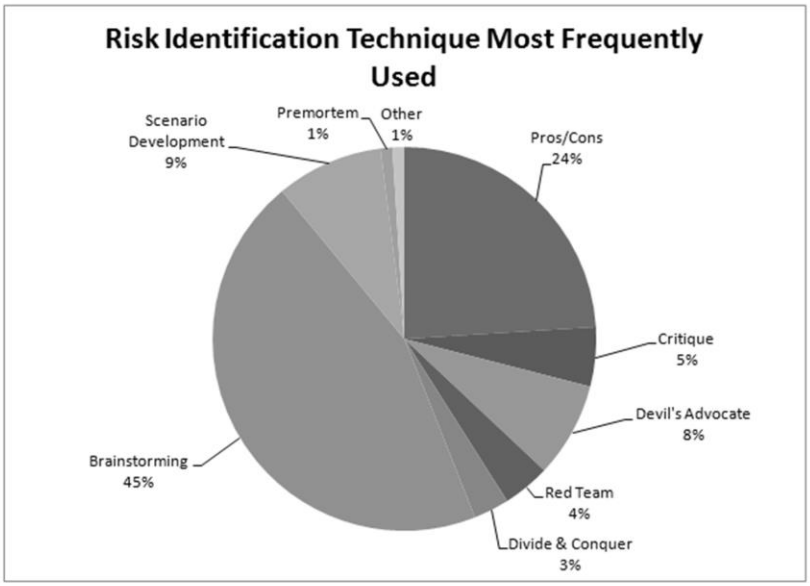


Figure 41 Subjects Most Used Risk ID Technique

The demographic data was used to determine if there was a statistical difference between those that performed the brainstorming treatment and those that performed the premortem treatment along the two demographic dimensions of risk identification experience and current duties reported in the subject survey. In other words, did one

treatment receive a statistically significant advantage by receiving more (or less) experienced subjects or more (or less) subjects that describe their job as technical? A Mann-Whitney hypothesis test for risk identification experience and current duties found no statistical difference between the two treatments for those demographic measures (Table 4).

Experience		Duties	
Mann-Whitney Test (General Version)		*Mann-Whitney Test (General Version)*	
Hypotheses		*Hypotheses*	
Null Hypothesis	Neither Dist. Smaller	Null Hypothesis	Neither Dist. Smaller
Alternative Hypothesis	Either Dist. Smaller	Alternative Hypothesis	Either Dist. Smaller
Ranking Information		*Ranking Information*	
Number of Values in Ranking	100	Number of Values in Ranking	100
Number of Tied Values	100	Number of Tied Values	100
Rank Sum for Variable 1 (Test Statistic)	2351	Rank Sum for Variable 1 (Test Statistic)	2500
Rank Sum for Variable 2	2699	Rank Sum for Variable 2	2550
p-Value Computation		*p-Value Computation*	
Normal Approximation (NA) Used	Yes	Normal Approximation (NA) Used	Yes
Ties Present, but Not Corrected For	No	Ties Present, but Not Corrected For	No
Mean for NA	2525	Mean for NA	2525
Std. Dev. for NA with Tie Correction	136.5192294	Std. Dev. for NA with Tie Correction	84.49941725
z-Statistic for NA with Tie Correction	-1.2709	z-Statistic for NA with Tie Correction	-0.2899
p-Value	0.2038	p-Value	0.7719
Significance Levels		*Significance Levels*	
Null Hypoth. at 10% Significance	Don't Reject	Null Hypoth. at 10% Significance	Don't Reject
Null Hypoth. at 5% Significance	Don't Reject	Null Hypoth. at 5% Significance	Don't Reject
Null Hypoth. at 1% Significance	Don't Reject	Null Hypoth. at 1% Significance	Don't Reject

Table 4 Hypothesis Test Showing No Difference in Treatment Demographics

There was no statistical difference between treatments regarding the team sizes. The mean size of the brainstorming teams was 5.0, and the premortem team was 5.1. The standard deviations were 0.9428 and 0.9444, respectively. I used the general version of the Mann-Whitney to test the difference between the two treatments regarding team size. The general version of the Mann-Whitney denies that either treatment tends to yield smaller values than the other. As opposed to the median version of the Mann-Whitney, the general version does not assume the two treatment distributions have the same shape.

The Mann-Whitney null hypothesis that the treatment team sizes were equal could not be rejected at 5 percent significance with a p-value of 0.8534. We can conclude that the two treatments' team sizes were not statistically different.

With 45 of the 100 respondents reporting that their most frequently used risk identification technique is the brainstorming technique; the brainstorming treatment may have had an inherent advantage over the premortem treatment. Nineteen of the 45 respondents that reported brainstorming to be their most frequently used technique did use brainstorming in the experiment. The one respondent that reported using the premortem technique most frequently actually performed brainstorming in the experiment. The premortem treatment did not have the advantage of subjects having frequent prior experience using the premortem in the experiments.

For the modified Delphi expert panel, I selectively recruited 30 DSMC faculty members based on experience and background. The biographies for those expert panel members who participated and gave me permission to add their biographies to this document can be found in Appendix A. All of the expert panel members were professors from DSMC. Their areas of subject matter expertise include the following:

- Program Management: space systems, foreign military sales, acquisition strategy development, Department of Defense acquisition policy, business case analysis, performance based logistics, risk and opportunity management
- Technical Management: systems engineering, design for exportability, human systems integration, open systems architecture, program protection planning, test and evaluation, cybersecurity, manufacturing readiness, requirements management, and quality management

- Business Management: contract management, Department of Defense budget process, and cost estimating
- Leadership: organizational dynamics, organizational design, stakeholder management, emotional intelligence, and strategic planning

Of the thirty expert panel members recruited, twenty-eight returned their score sheets (a 93 percent response rate). Two of the twenty-eight score sheets were excluded because the expert panel members failed to follow the established protocol. They failed to provide numerical responses where required or failed to grade all of the deliverable sets. Therefore, I had 26 of 30 completed and valid expert panel member score sheets for analysis (an 87 percent valid response rate).

5.2 Null Hypothesis #1

In this section, I will report on the results from the twenty experiments in terms of rejecting null hypothesis #1 and its subordinate null hypotheses.

- Null Hypothesis #1: Team collaboration using the premortem technique is equal to team collaboration using brainstorming. This null hypothesis can be decomposed into four subordinate null hypotheses.
    - Null Hypothesis #11: The number of individual differences in perspective discovered using the premortem technique is equal to the number of differences in perspective discovered using brainstorming.
    - Null Hypothesis #12: The number of ideas enriched using the premortem technique is equal to the number of ideas enriched using brainstorming.

- Null Hypothesis #13: The number of critiques applied using the premortem technique is equal to the number of critiques applied using brainstorming.
- Null Hypothesis #14: The number of original risk ideas offered using the premortem technique is equal to the number of original ideas offered using brainstorming.

The focus of the null hypotheses is to compare team collaboration when one set of ten teams used brainstorming for risk identification and another set of ten teams used the premortem technique for risk identification. We had a total of 101 subjects, fifty of whom used brainstorming and fifty-one of whom used the premortem technique. For hypothesis testing, I separated the 101 into their respective treatments and used each subject's total collaborative exchanges as a sample value. The collaborative behaviors were measured along five dimensions:

- Offered a unique risk idea offered
- Enriched a risk idea of another teammate
- Provided a different perspective on a risk idea offered by another teammate
- Critiqued a risk idea of another teammate
- The sum of all collaborative behaviors

I first tested the distribution of the data from the two treatments and the five collaborative behavior dimensions (ten total data sets) for normality using the Lilliefors test. If the Lilliefors test statistic is greater than the Lilliefors critical value then we reject the normality assumption. I could not assume normality for nine of the ten data sets. The

remaining data set marginally passes the normality test. A parametric test with non-normal data may lead to a flawed inference. Therefore, I performed nonparametric tests of the null hypotheses:

"the number of [collaborative behavior of interest] using the premortem technique is equal to the number of [the collaborative behavior of interest] using brainstorming."

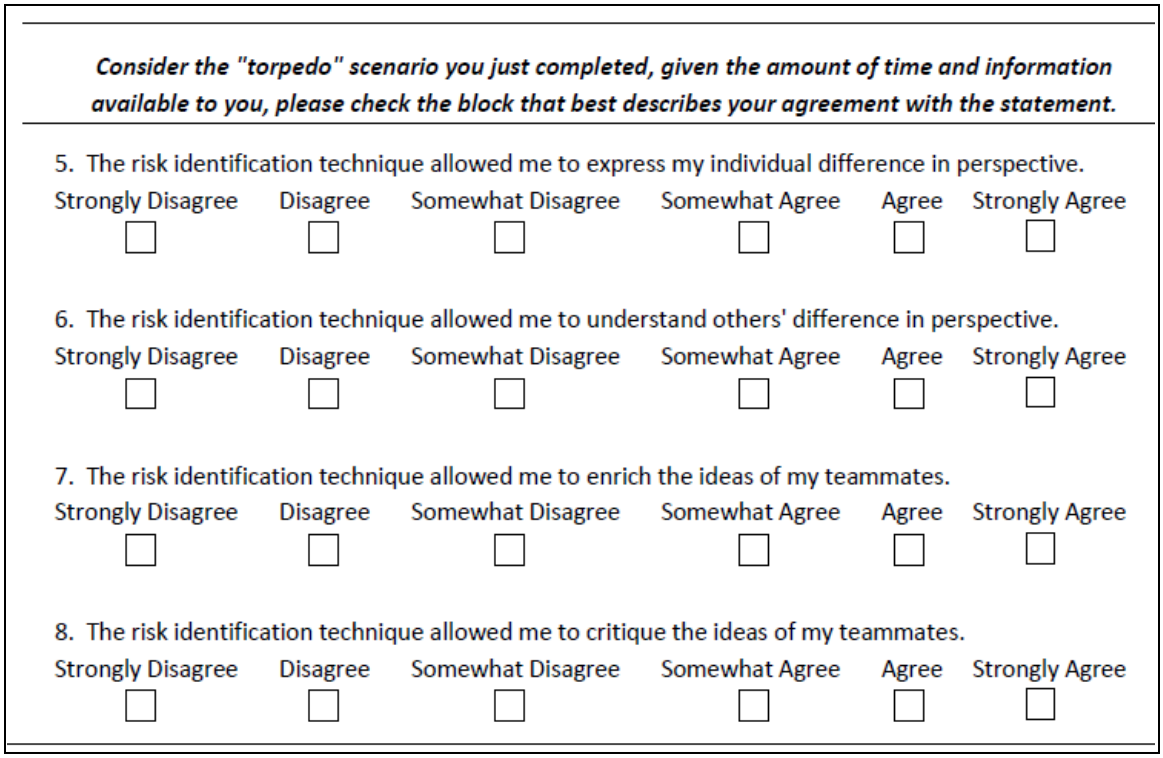
The Mann-Whitney test compares two populations to determine if one population tends to yield smaller values than the other. The general version of the Mann-Whitney test does not assume that the distributions have the same shape. The analysis indicates that the premortem generates more of the collaborative behaviors for four of the five dimensions at the 5 percent significance level. Null hypothesis #14 regarding the number of original risk ideas offered cannot be rejected according to the Mann-Whitney tests.

The nonparametric tests of the data from our experiments thus far would indicate the premortem technique has advantages in collaborative risk identification. However, parametric tests of the null hypotheses are preferred. I applied the Box-Cox transformation to address the normality assumption the t-test requires. Box-Cox is a logarithmic transformation. I added "1" to the sample values to eliminate "0" in the data prior to transformation since the log of 0 does not exist. The parametric tests on the transformed data confirm the nonparametric tests for three of the five dimensions. I was unable to reject the null hypothesis for the collaboration dimension, which is the sum of the other four dimensions, and the original risk ideas offered dimension. I conclude there is strong evidence that teams conducting the premortem technique build on one another's ideas, offer differing perspectives and critique each other more than teams conducting brainstorming. Table 5 is a summary of the test results.

Dimension	Collaborations		Risk Ideas Offered		Ideas Enriched		Differences in Perspective		Critiques	
Treatment	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*
Sample Size	50	51	50	51	50	51	50	51	50	51
Mean	7.080	11.176	5.140	6.098	1.060	2.922	0.4800	1.294	0.4000	0.863
Standard Deviation	2.791	6.241	2.373	4.095	1.252	2.189	0.8862	1.591	0.6389	1.000
Median	7	11	5	5	1	2	0	1	0	1
Mode	10	7	4	3	0	2	0	0	0	0
Minimum	2	2	1	1	0	0	0	0	0	0
Maximum	13	31	11	19	5	10	4	7	3	4
Lilliefors Test Statistic	0.1251	0.1191	0.1645	0.1468	0.2414	0.1729	0.4060	0.2400	0.3944	0.2567
Lilliefors Critical Value (5% significance)	**0.1246**	0.1234	**0.1246**	**0.1234**	**0.1246**	**0.1234**	**0.1246**	**0.1234**	**0.1246**	**0.1234**
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Reject		Don't Reject		Reject		Reject		Reject	
P-value	0.0002		0.2749		< 0.0001		0.0006		0.0060	
Box-Cox Mean	8.9780	2.8800	2.2994	2.1371	0.6406	1.6589	0.2460	0.6424	0.2515	0.5711
Box-Cox Standard Deviation	3.8730	0.7320	0.6591	0.8077	0.6606	0.9629	0.4011	0.6113	0.3684	0.5940
Box-Cox Median	8.7570	3.0113	2.3496	2.0496	0.7515	1.3796	0.0000	0.6956	0.0000	0.7650
Box-Cox Minimum	2.2210	1.1943	0.7677	0.7356	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Box-Cox Maximum	17.5480	4.5430	3.6404	3.9604	2.2174	4.0237	1.3657	2.1012	1.3302	2.0333
t-Test of the Null Hypothesis using Box-Cox Transformations (5% significance)	Don't Reject		Don't Reject		Reject		Reject		Reject	
P-value	1.0000		1.0000		0.0001		< 0.0001		0.0002	

Table 5 Testing Null Hypothesis #1 Based on Structured Observation

Subject responses to the survey questions #5 through #8 (Figure 42) can be analyzed to confirm or contradict the collaboration data from structured observation. Analyzing the 100 surveys (50 from brainstorming subjects and 50 from premortem subjects), I found the premortem subjects reporting a higher degree of perspective sharing, idea enrichment, and critiquing than their brainstorming counterparts.



Consider the "torpedo" scenario you just completed, given the amount of time and information available to you, please check the block that best describes your agreement with the statement.

5. The risk identification technique allowed me to express my individual difference in perspective.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

6. The risk identification technique allowed me to understand others' difference in perspective.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

7. The risk identification technique allowed me to enrich the ideas of my teammates.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

8. The risk identification technique allowed me to critique the ideas of my teammates.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

Figure 42 Subject Survey Questions #5 through #8 (Likert Scale "Strongly Disagree"=1 "Strongly Agree"=6)

The general version of the Mann-Whitney hypothesis test formed the null hypothesis that neither brainstorming nor premortem will yield lower Likert scores than the other. The alternative hypothesis using a one-tail test is: brainstorming yields lower Likert scores than the premortem. The null hypothesis was rejected in all four cases, indicating that the brainstorming subjects reported lower levels of perspective sharing, perspective understanding, idea enrichment, and critiquing than their premortem counterparts. This data in Table 6 confirms the structured observation findings that teams performing the premortem offered more perspectives, enrichments, and critiques.

Question	Express My Perspective		Understand Others' Perspective		Enrich Others' Ideas		Critique Others' Ideas	
Treatment	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*
Sample Size	50	50	50	50	50	50	50	50
Mean	4.840	5.320	4.780	5.300	4.4600	5.200	3.6800	4.660
Standard Deviation	0.955	0.5127	1.0753	0.5051	0.9730	0.6701	1.4350	0.9392
Mode	5	5	5	5	4	5	5	5
Minimum	1	4	1	4	2	4	1	2
Maximum	6	6	6	6	6	6	6	6
Lilliefors Test Statistic	0.3465	0.3937	0.2811	0.4037	0.2018	0.2773	0.2012	0.2613
Lilliefors Critical Value (5% significance)	**0.1246**	**0.1246**	**0.1246**	**0.1246**	**0.1246**	**0.1246**	**0.1246**	**0.1246**
Mann-Whitney Test of the Null Hypothesis (5% significance)	Reject		Reject		Reject		Reject	
P-value	0.0019		0.0046		<0.0001		0.0003	

Table 6 Survey Questions #5 through #8 Confirm Structured Observation Findings

From the data analysis, we can conclude that the alternative hypothesis for Null Hypothesis #1 is partially supported. The premortem technique enables a higher degree of team collaboration along the dimensions of perspective sharing, idea enrichment, and critiques. We cannot conclude from the data that the premortem generates more original risk ideas than brainstorming.

5.3 Null Hypothesis #2

At the conclusion of the experiments, the 26-member expert panel reviewed all twenty sets of team deliverables. There were ten sets of deliverables from teams that performed risk identification using brainstorming and ten sets of deliverables from teams that used the premortem technique. The expert panel members were blind to which deliverables were developed by which technique. The expert panel members reviewed the deliverables independently from each other. Their scores were the basis for the testing null hypotheses #2, #3, and #4. In this section, I will present the results testing *Null Hypothesis #2: The number of quality risks identified using the premortem technique*

is equal to the number of quality risks identified using brainstorming. I instructed the expert panel members to use a score of 70 as their quality cut-off. A score of 70 or higher for <u>both</u> dimensions (creativity and plausibility) of the risk represented a quality score. The expert panel members were unaware of the two-part requirement for a quality score.

Reviewing the data from the nonparametric Mann-Whitney test and the parametric test of the data transformed to normality using the Box-Cox method, I rejected the null hypothesis in both cases (Table-7). We conclude that the 26 expert panel members found the premortem technique identified more quality risks than the brainstorming technique.

	Total Quality Risks	
	Brainstorming	Premortem
Expert-1	3	4
Expert-2	14	19
Expert-3	0	0
Expert-4	22	29
Expert-5	0	7
Expert-6	5	9
Expert-7	2	3
Expert-8	0	0
Expert-9	1	3
Expert-10	7	12
Expert-11	9	12
Expert-12	7	10
Expert-13	19	21
Expert-14	13	24
Expert-15	17	25
Expert-16	0	4
Expert-17	31	33
Expert-18	9	18
Expert-19	9	34
Expert-20	19	21
Expert-21	14	11
Expert-22	6	10
Expert-23	7	13
Expert-24	0	4
Expert-25	8	19
Expert-26	6	3

	Quality Risks from Expert Panel Members	
Treatment	*Brainstorming*	*Premortem*
Sample Size	26	26
Mean	8.769	13.380
Standard Deviation	0.955	0.5127
Mode	0	3
Minimum	0	0
Maximum	31	34
Lilliefors Test Statistic	0.1807	0.1318
Lilliefors Critical Value (5% significance)	**0.1703**	0.1703
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Reject	
P-value	0.0432	
Box-Cox Mean	3.7500	5.3330
Box-Cox Standard Deviation	2.6900	3.0430
Box-Cox Median	3.7020	5.1550
Box-Cox Minimum	0.0000	0.0000
Box-Cox Maximum	9.5220	10.5320
Parametric Test of the Null Hypothesis using Box-Cox Transformations (5% significance)	Reject	
P-value	0.0030	

Table 7 Test Results for Null Hypothesis #2 Quality Risks

5.4 Null Hypothesis #3

Null hypothesis #3 addressed the quality of the changes to the plan. Like the quality of risks in null hypothesis #2, the expert panel scored the four changes to the plan for each team based on two dimensions. In order for a change to the plan to be considered quality, it had to receive a score of 70 or greater in both the creativity and usefulness dimensions. Null Hypothesis #3 stated: *The number of quality proposed changes to the plan using the premortem technique is equal to the number of quality proposed changes to the plan using brainstorming.* Both the Mann-Whitney nonparametric test and the parametric test indicate that we can reject the null hypothesis and conclude that the expert panel members found that teams using the premortem technique proposed more quality changes to the plan than the teams that used brainstorming (Table 8).

	Quality Changes to Plan	
	Brainstorming	Premortem
Expert-1	6	8
Expert-2	12	25
Expert-3	0	0
Expert-4	11	21
Expert-5	8	9
Expert-6	7	15
Expert-7	2	2
Expert-8	0	0
Expert-9	3	2
Expert-10	11	18
Expert-11	13	15
Expert-12	14	17
Expert-13	11	19
Expert-14	10	14
Expert-15	14	21
Expert-16	7	5
Expert-17	11	12
Expert-18	7	14
Expert-19	8	19
Expert-20	24	30
Expert-21	20	22
Expert-22	6	14
Expert-23	7	9
Expert-24	5	10
Expert-25	3	6
Expert-26	4	9

	Quality Changes to the Plan from Expert Panel Members	
Treatment	*Brainstorming*	*Premortem*
Sample Size	26	26
Mean	8.615	12.930
Standard Deviation	5.636	7.8530
Mode	7	9
Minimum	0	0
Maximum	24	30
Lilliefors Test Statistic	0.1204	0.0930
Lilliefors Critical Value (5% significance)	0.1703	0.1703
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Reject	
P-value	0.0153	
Parametric Test of the Null Hypothesis (5% significance)	Reject	
P-value	0.0003	

Table 8 Test Results for Null Hypothesis # 3 Quality Changes to the Plan

5.5 Null Hypothesis #4

Null hypothesis #4 addressed whether the four risks identified by the team matched one of the four black swans the team in the Mark-14 torpedo case missed. If one of the team's risks matched the one of the four black swans, the risk was classified as a "black swan hit." Null hypothesis #4 stated: *The number of historical risks identified (black swan hits) using the premortem technique is equal to the number of historical risks identified (black swan hits) using brainstorming.* Both the Mann-Whitney nonparametric and parametric tests indicate that we can reject the null hypothesis. Therefore, we conclude that teams using the premortem technique had more "black swan hits" than teams using the brainstorming technique (Table 9).

	Black Swan Hits	
	Brainstorming	Premortem
Expert-1	9	15
Expert-2	12	12
Expert-3	17	19
Expert-4	11	14
Expert-5	3	6
Expert-6	9	13
Expert-7	21	31
Expert-8	12	13
Expert-9	2	7
Expert-10	11	11
Expert-11	14	20
Expert-12	13	30
Expert-13	9	6
Expert-14	14	20
Expert-15	9	8
Expert-16	2	2
Expert-17	12	13
Expert-18	9	13
Expert-19	5	19
Expert-20	20	22
Expert-21	5	6
Expert-22	11	15
Expert-23	14	17
Expert-24	27	24
Expert-25	7	9
Expert-26	6	8

	Black Swan Hits from Expert Panel Members	
Treatment	*Brainstorming*	*Premortem*
Sample Size	26	26
Mean	10.923	14.346
Standard Deviation	5.885	7.3150
Mode	9	13
Minimum	2	2
Maximum	27	31
Lilliefors Test Statistic	0.1467	0.1182
Lilliefors Critical Value (5% significance)	0.1703	0.1703
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Reject	
P-value	0.0407	
Parametric Test of the Null Hypothesis (5% significance)	Reject	
P-value	0.0033	

Table 9 Test Results for Null Hypothesis #4 Black Swan Hits

5.6 Expert Panel Top-3 Votes on the Deliverable Sets

As the last step in the expert panel scoring process, each member was directed to review all twenty sets of deliverables and identify the three sets of deliverables they would most like to receive if they were the senior decision maker. The intent was to allow the expert panel members to apply their own criteria of quality. Would the expert panel members select more deliverable sets in their top three that were generated using premortem or brainstorming? Written as a null hypothesis:

The number deliverable sets receiving expert panel member votes as being in the Top-3 votes using the premortem technique is equal to the number deliverable sets receiving expert panel member votes as being in the top three using brainstorming. The alternative hypothesis is… the deliverable sets developed using the premortem technique receive more Top-3 votes that those developed using brainstorming.

The nonparametric Mann-Whitney test results indicate that we should reject the null hypothesis and conclude that the deliverables generated using the premortem technique receive more top-three votes from the expert panel members. After the Box-Cox transformation, the parametric test confirmed the Mann-Whitney test results (Table-10). These tests indicate that our expert panel members, in the role of senior decision makers, using their own measure of quality, preferred the risk lists and changes to the plan that teams using the premortem technique developed.

	Top-3 Expert Panel Member Votes	
	Brainstorming	Premortem
Expert-1	0	3
Expert-2	1	2
Expert-3	1	2
Expert-4	0	3
Expert-5	0	3
Expert-6	0	3
Expert-7	1	2
Expert-8	1	2
Expert-9	2	1
Expert-10	1	2
Expert-11	2	1
Expert-12	0	3
Expert-13	1	2
Expert-14	0	3
Expert-15	1	2
Expert-16	0	3
Expert-17	1	2
Expert-18	0	3
Expert-19	0	3
Expert-20	2	1
Expert-21	1	2
Expert-22	2	1
Expert-23	0	3
Expert-24	2	1
Expert-25	1	2
Expert-26	3	0

	Top-3 Deliverables as Scored By the Expert Panel	
Treatment	*Brainstorming*	*Premortem*
Sample Size	26	26
Mean	0.8846	2.1154
Standard Deviation	0.8638	0.8638
Mode	0	2
Minimum	0	0
Maximum	3	3
Lilliefors Test Statistic	0.2317	0.2317
Lilliefors Critical Value (5% significance)	**0.1703**	**0.1703**
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Reject	
P-value	0.0407	
Box-Cox Mean	0.7279	2.8910
Box-Cox Standard Deviation	0.6788	1.3200
Box-Cox Median	0.8724	2.6290
Box-Cox Minimum	0.0000	0.0000
Box-Cox Maximum	2.2319	4.2990
Parametric Test of the Null Hypothesis using Box-Cox Transformations (5% significance)	Reject	
P-value	<0.0001	

Table 10 Test Results for Expert Panel Member Top-3 Voting

5.7 Subject Survey Results for Items 9 through 12

Subject survey items 9 through 12 asked the subjects to compare the risk identification technique they used in the experiment (brainstorming or premortem) to any other risk identification technique they had previously used (Figure 43). The survey items asked the subject to compare the techniques along the following attributes:

- More risks
- Better risks
- More creative changes to the plan
- Ability to avoid of organizational blind spots

Items 9 thru 12 ask you to compare the risk identification technique you used in the "torpedo" scenario (Brainstorming or PreMortem) to other risk identification techniques you have used. Consider a risk identification activity in which you previously participated. Consider what went well and not so well. Please check the block that best describes your agreement with the statement.

9. If we had used the risk identification technique used in the "torpedo" scenario, we would have discovered more risks in the same amount of time.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

10. If we had used the risk identification technique used in the "torpedo" scenario, we would have discovered better risks in the same amount of time.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

11. If we had used the risk identification technique used in the "torpedo" scenario, we would have been able to propose more creative improvements to the plan in the same amount of time.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

12. If we had used the risk identification technique used in the "torpedo" scenario, we would have been less vulnerable to organizational blind-spots.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

Figure 43 Subject Survey Items 9 thru 12

Analyzing the responses by treatment will confirm or contradict the expert panel results. As noted earlier, the expert panel results indicated teams that used the premortem technique presented more quality risks and quality changes to the plan than teams that used brainstorming. The subject's responses were not constrained by *creativity-plausibility* or *creativity-usefulness* dimensions as was the expert panel. The subjects were free to use whatever other risk identification technique they have used in the past as a comparative benchmark. Although the subjects did not report their benchmark technique, based on the demographic data, it is highly unlikely that subjects that used the brainstorming treatment compared that technique to the premortem technique as their

comparative benchmark. In hindsight, this is a remote possibility and a limitation in the wording of this survey item.

Using the Likert scale scores of 1 = *Strongly Disagree* to 6= *Strongly Agree*, survey item #9 regarding the quantity of risks indicates the mean score for those using the brainstorming technique was 3.92; close to *somewhat agree* that brainstorming identified more risks than the subject's benchmark technique. The mean score for those subjects using the premortem technique was 4.56, that is between *somewhat agree* and *agree*. Performing the nonparametric hypothesis test and a parametric test on the transformed Likert scores both indicated that the subjects that used the premortem technique believed the premortem identified more risks when compared to their personal benchmark (Table 11).

	Subject Survey Results Question-9 More Risks	
Treatment	*Brainstorming*	*Premortem*
Sample Size	50	50
Mean	3.9200	4.5600
Standard Deviation	1.0850	0.8609
Mode	4	5
Minimum	2	3
Maximum	6	6
Lilliefors Test Statistic	0.1694	0.2553
Lilliefors Critical Value (5% significance)	**0.1246**	**0.1246**
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Reject	
P-value	0.0014	
Box-Cox Mean	3.0530	5.0520
Box-Cox Standard Deviation	1.1600	1.4510
Box-Cox Median	3.1310	5.7650
Box-Cox Minimum	0.2100	2.5230
Box-Cox Maximum	5.2970	7.5800
Parametric Test of the Null Hypothesis using Box-Cox Transformations (5% significance)	Reject	
P-value	<0.0001	

Table 11 Test Results Subject Survey Item #9 "More Risks"

Survey item #10 asked the subject to compare the risk identification technique they used in the experiment with their personal benchmark technique in terms of quality of the risks identified. The subjects using the brainstorming technique had a mean Likert score of 3.9. The subjects using the premortem technique had a mean Likert score of 4.56. The nonparametric hypothesis test and the parametric hypothesis test of the transformed scores indicated the subjects that used the premortem technique reported that the premortem technique discovered better risks that the subject's personal benchmark. This data confirms the expert panel findings that the premortem technique generated more quality risks than the brainstorming technique (Table 12).

	Subject Survey Results Question-10 Better Risks	
Treatment	*Brainstorming*	*Premortem*
Sample Size	50	50
Mean	3.9000	4.5600
Standard Deviation	1.0150	0.8843
Mode	4	5
Minimum	2	1
Maximum	6	6
Lilliefors Test Statistic	0.2192	0.2906
Lilliefors Critical Value (5% significance)	**0.1246**	**0.1246**
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Reject	
P-value	0.0002	
Box-Cox Mean	3.5980	10.2800
Box-Cox Standard Deviation	1.3940	3.5940
Box-Cox Median	3.6930	12.0000
Box-Cox Minimum	1.0980	0.0000
Box-Cox Maximum	6.8320	17.5000
Parametric Test of the Null Hypothesis using Box-Cox Transformations (5% significance)	Reject	
P-value	<0.0001	

Table 12 Test Results Subject Survey Item #10 "Better Risks"

Survey item #11 asked the subjects to compare the risk identification technique they used in the experiment against their personal benchmark technique with regard to the technique's ability to generate creative improvements to the plan. Those subjects that used the brainstorming technique reported a mean Likert score of 3.92, while those that performed the premortem technique reported a mean Likert score of 4.78. The nonparametric test and the parametric tests on the transformed scores both indicate a statistical difference in the subject response values. Those subjects that used the premortem technique reported that it generated more creative improvements to the plan than their personal benchmark. This data confirms the expert panel findings that the premortem technique identifies more quality changes to the plan than the brainstorming technique (Table 13).

	Subject Survey Results Question-11 More Creative Improvements to Plan	
Treatment	*Brainstorming*	*Premortem*
Sample Size	50	50
Mean	3.9200	4.7800
Standard Deviation	0.9223	0.9957
Mode	4	5
Minimum	2	1
Maximum	6	6
Lilliefors Test Statistic	0.2346	0.3074
Lilliefors Critical Value (5% significance)	**0.1246**	**0.1246**
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Reject	
P-value	<0.0001	
Box-Cox Mean	2.9456	11.4100
Box-Cox Standard Deviation	0.9347	4.2070
Box-Cox Median	3.0256	12.0000
Box-Cox Minimum	1.0039	0.0000
Box-Cox Maximum	5.0579	17.5000
Parametric Test of the Null Hypothesis using Box-Cox Transformations (5% significance)	Reject	
P-value	<0.0001	

Table 13 Test Results Subject Survey Item #11 "Creative Improvements"

Subject survey item #12 attempted to address the subject's perspective on the treatment's ability to "catch the black swan." It was not practical to ask the subjects if the technique they used identified more black swans since the subjects were unaware of the black swans in the Mark-14 case. My approach was to ask them to assess the technique for its ability to *avoid organizational blind spots*. We know that organizational blind spots and deliberate ignorance are sources of black swan risks in general and in the Mark-14 case specifically.

	Subject Survey Results Question-12 Less Vulnerable to Org Blindspots	
Treatment	*Brainstorming*	*Premortem*
Sample Size	50	50
Mean	3.7200	4.8800
Standard Deviation	1.1440	1.0810
Mode	4	5
Minimum	2	2
Maximum	6	6
Lilliefors Test Statistic	0.1755	0.2442
Lilliefors Critical Value (5% significance)	**0.1246**	**0.1246**
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Reject	
P-value	<0.0001	
Box-Cox Mean	2.3955	11.9800
Box-Cox Standard Deviation	0.9417	4.7410
Box-Cox Median	2.6459	12.0000
Box-Cox Minimum	0.9441	1.5000
Box-Cox Maximum	4.2188	17.5000
Parametric Test of the Null Hypothesis using Box-Cox Transformations (5% significance)	Reject	
P-value	<0.0001	

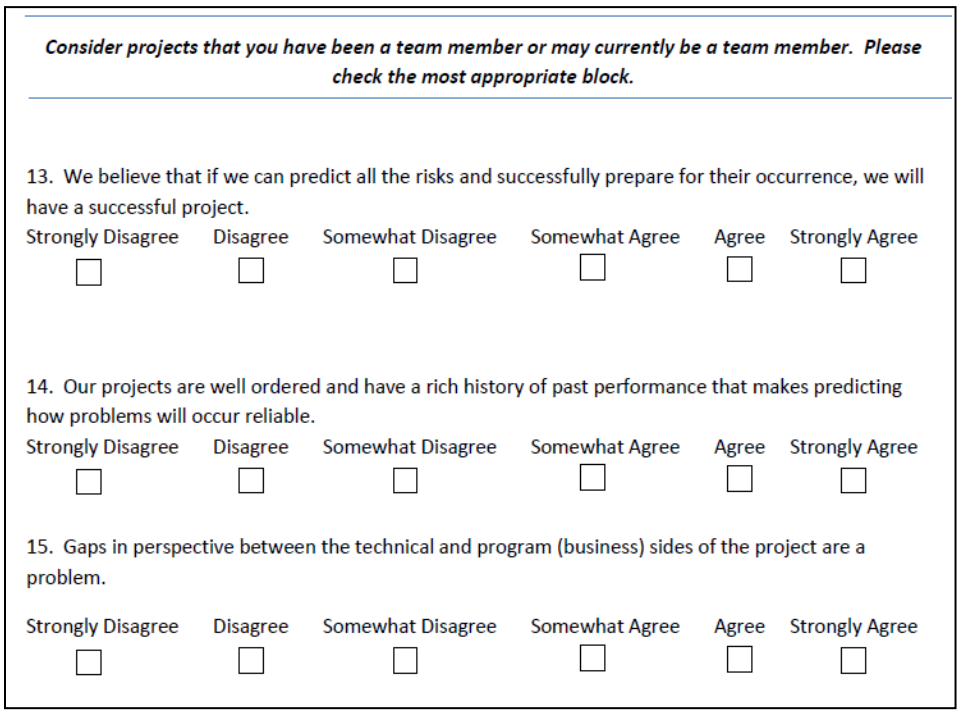
Table 14 Test Results Subject Survey Item #12 "Avoid Organizational Blind Spots"

The mean Likert response for the subjects using the brainstorming technique was 3.72 while the mean Likert score for subjects using the premortem technique was 4.88. This represents the largest difference between the two treatment means for the four subject survey items considered in this section. Both the nonparametric test and the parametric test of the transformed data indicate a statistical difference in the subject responses based on the technique they used in the experiment. The results indicated that the premortem technique is better than brainstorming at avoiding organizational blind spots. This is based on the subjects' comparison of the risk identification technique they used in the experiment with another risk identification technique they used in the past. The subject survey results for item #12 (Table-14) indirectly confirmed the expert panel's findings that the premortem technique identifies more black swans than the brainstorming technique.

5.8 Subject Survey Items 13 through 15

The last three subject survey items attempted to assess the subject's belief regarding risk attitudes in their organization. No comparison is made between the treatments in this survey section (Figure 44). Survey item #13 addressed the notion that if organizations are able to predict risks and prepare for their occurrence that they will have a successful project. This is the philosophy of project risk management (PMI 2009). That philosophy does not help with the detection of black swans that come from outside a team's experience base. Responses to this survey item may indicate our black swan vulnerability. The mean response to survey item #13 was 4.65 on our Likert scale with a

standard deviation of 0.9987. This would indicate the subjects tended towards agreement with the statement. It is interesting to note that none of the subjects were in *Strong Disagreement* with the statement. The modal response was to *Agree* with the statement.



Consider projects that you have been a team member or may currently be a team member. Please check the most appropriate block.

13. We believe that if we can predict all the risks and successfully prepare for their occurrence, we will have a successful project.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

14. Our projects are well ordered and have a rich history of past performance that makes predicting how problems will occur reliable.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

15. Gaps in perspective between the technical and program (business) sides of the project are a problem.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

Figure 44 Subject Survey Items #13 thru #15

Survey item #14 addressed the appropriateness of quantitative risk identification techniques for the projects supported by our subjects. We know that quantitative risk identification tools require a history of past performance and a degree of stability to make predicting the future reliable. The responses to survey item #14 should provide an indication of how appropriate quantitative tools are for the subject's project. A *Strongly Agree* response would indicate that quantitative tools would be appropriate for the perceived project environment (well-ordered and predictable). The mean response to survey item #13 was 3.31 with a standard deviation of 1.237 indicating the subjects

tended to *Somewhat Disagree* but the responses were more varied than survey item #13. The modal response was to *Somewhat Agree*. Results are in Table 15.

One Variable Summary	Predict is Success Organizational Questions 13, 14, and 15	Well Ordered Organizational Questions 13, 14, and 15	Gaps in Perspective Organizational Questions 13, 14, and 15
Mean	4.6500	3.310	4.480
Std. Dev.	0.9987	1.237	1.059
Mode	5.0000	4.000	5.000
Minimum	2.0000	1.000	2.000
Maximum	6.0000	6.000	6.000
Count	100	100	100

Table 15 Summary Statistics for Subject Survey Items #13, #14, and #15.

The relationship between responses to survey item #13 and survey item #14 deserves some analysis. Survey item #13 concerned the strategy employed for project success—namely, if we predict risks and plan for their occurrence, we will have successful projects. Survey item #14 concerned the perceived risk environment. If the responses to survey items #13 are *Strongly Agree* that the predict and plan strategy leads to project success but *Strongly Disagree* that our projects are well ordered and predictable, we would have a high degree of dissonance between the strategy for project success and our ability for that strategy to be successful. A Lilliefors test of the distributions of both sets of responses found both to fail the normality test at the 5 percent significance level. The Box-Cox transformation was applied to the data to conform to normality in order to perform a one-way ANOVA. The one-way ANOVA indicated a reliable main effect of the survey item ("predict and plan is success" and "well ordered and predictable projects") on the respondents' levels of agreement $F(1, 100) = 71.066$, $p < 0.0001$. This indicates a statistically significant dissonance between the strategy the

respondents use to achieve project success through "plan and predict" and the nature of their projects as being not well ordered or predictable. Results are in Table 16.

One-Way ANOVA for Selected Unstacked Variables

ANOVA Summary

Total Sample Size	200
Grand Mean	3.9800
Pooled Std Dev	1.1240
Pooled Variance	1.2633
Number of Samples	2
Confidence Level	95.00%

ANOVA Sample Stats	Predict is Success Organizational Questions 13, 14, and 15	Well Ordered Organizational Questions 13, 14, and 15
Sample Size	100	100
Sample Mean	4.6500	3.310
Sample Std Dev	0.9987	1.237
Sample Variance	0.9975	1.529
Pooling Weight	0.5000	0.5000

OneWay ANOVA Table	Sum of Squares	Degrees of Freedom	Mean Squares	F-Ratio	p-Value
Between Variation	89.7800	1	89.7800	71.0660	< 0.0001
Within Variation	250.1400	198	1.2633		
Total Variation	339.9200	199			

Confidence Interval Tests	Difference of Means	No Correction Lower	No Correction Upper	Bonferroni Lower	Bonferroni Upper	Tukey Lower	Tukey Upper	Sch Lower
Predict is Success-Well Ordered	1.3400	1.026538119	1.653461881	1.026538119	1.653461881	1.028448427	1.651551573	1.026538119

Table 16 ANOVA for Subject Survey Items #13 and #14

Survey item #15 addressed the perceived gap between the technical and programmatic sides of projects. This is one of the problem areas INCOSE identified in *SE Vision 2025*. However, INCOSE's statement is based on anecdotal reports (INCOSE 2015a). The subject survey results reported here provide some empirical measures to the assertion that the gaps between the technical and programmatic sides of a project are a problem. The responses to subject survey item #15 also provide an indication of our organizations' vulnerability to the black swan that results from the deliberate ignorance elements of untopicality, undecidability, or taboo topics. The mean response to survey item #15 was 4.48 with a standard deviation of 1.059, indicating the subjects were tending to *Agree* that the gaps between technical and programmatic sides of a project are

a problem. The modal response was to *Agree* with the statement. This provides empirical evidence of INCOSE's assertion that a gap exists and is a problem.

5.9 Group Performance (Student, Government, and Industry)

The twenty experiments have been analyzed based on the two treatments of brainstorming and the premortem technique. The experiments can be further analyzed to determine if the three groups of subject teams performed differently in the experiment. The three groups based on organization type are: the *Student* teams (DSMC and the Eisenhower School), the *Government* teams (Future Frigate and US Coast Guard) and the *Industry* teams from Textron (Figure 45). There were a total of seven *Student* teams, seven *Government* teams, and six *Industry* teams. Four of the *Student* teams performed the experiment using the brainstorming technique and three of the *Student* teams performed the experiment using the premortem technique. *Student* teams consisted of experienced professionals receiving executive level training in an academic environment. The *Government* teams were experienced professionals from a program management office or headquarters. The *Industry* teams were experienced professionals from a variety of programs, positions, and geographic locations within the Textron Corporation.

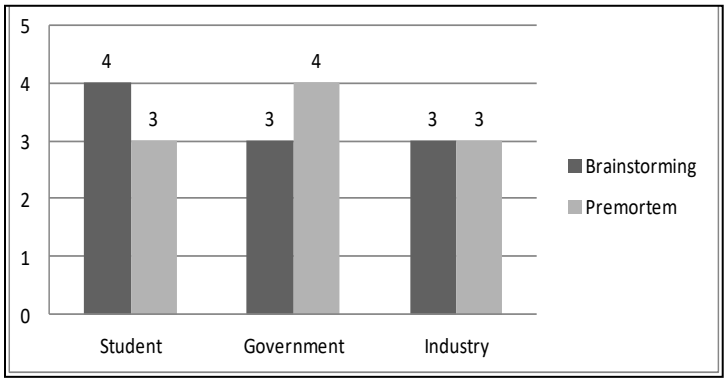


Figure 45 Number of Experiments by Group and Treatment

We can begin the comparison of performance between the three groups and two `treatments by examining the expert panel votes for the top-three deliverables. Each expert panel member used their own criteria to select the three best team deliverable sets. Earlier in this chapter we concluded that the premortem showed a statistical advantage over the brainstorming technique in producing deliverables in the top three. Examining the total top-three votes by group (*Student*, *Government*, or *Industry*) in Figure 46 and further separating the results into the two treatments (brainstorming or premortem), we see that the combination of *Government* teams using the premortem technique received the most top-three votes.

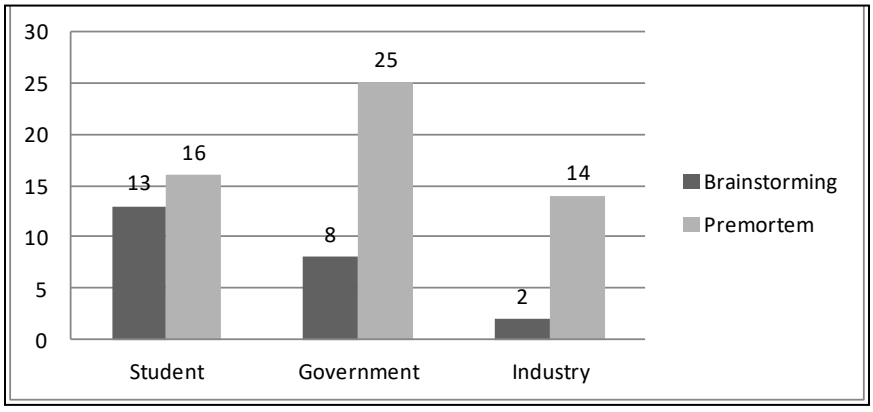


Figure 46 Total Top-3 Votes from the 26 Expert Panel Members by Group and Treatment

The group and treatment combinations in the rank order from highest to lowest are:

- Government teams using Premortem

- Student teams using Premortem
- Industry teams using Premortem
- Student teams using Brainstorming
- Government teams using Brainstorming
- Industry teams using Brainstorming

Since the *Student* and *Government* groups were not equal in the number of experiments per treatment, the top-three results were then indexed to a "per experiment" (Figure 47) basis to remove this discrepancy.

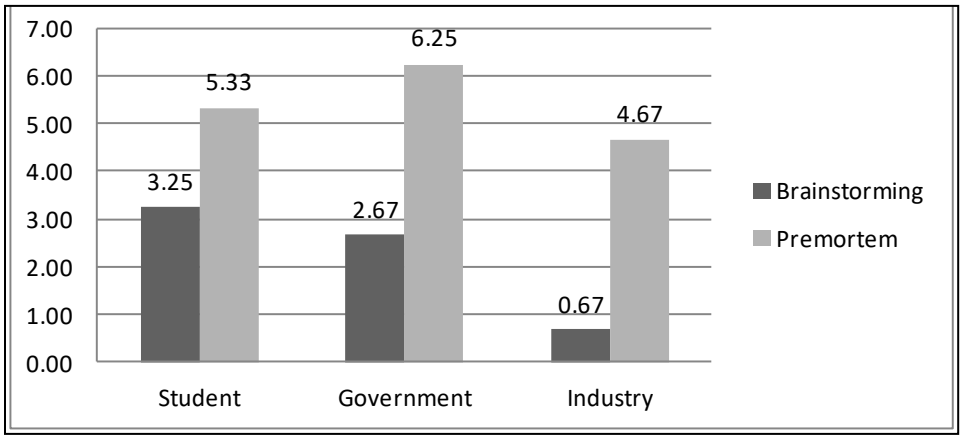


Figure 47 Average Expert Panel Top-3 Votes by Group and Treatment per Experiment

The *Government* group had a similar advantage with the additional premortem experiment. After indexing the total top-three votes, we see that the relative rank order of the groups and treatment combinations remained the same. All groups that used the premortem technique outperformed the counterparts based on the expert panel Top-3 votes. We can conclude that the premortem technique provides an advantage in delivering quality sets of risks and changes to the plan regardless of group based on the

26 expert panel member top-three votes. When considering the group results for brainstorming we see that the *Student* and *Government* teams performed significantly better than the *Industry* teams. However, a one-way ANOVA (Table 17) of the average expert panel top-three votes by organization and treatment confirms that the main effect is between the treatments and not between the organization types.

One-Way ANOVA for Selected Unstacked Variables

ANOVA Summary	
Total Sample Size	6
Grand Mean	3.8056
Pooled Std Dev	1.1108
Pooled Variance	1.2338
Number of Samples	2
Confidence Level	95.00%

	Average Top-3 Votes	
	Brainstorming	Premortem
Student	3.25	5.33
Government	2.67	6.25
Industry	0.67	4.67

ANOVA Sample Stats	Brainstorming Avg Top-3 by Org and Treatment	Premortem Avg Top-3 by Org and Treatment
Sample Size	3	3
Sample Mean	2.194	5.4167
Sample Std Dev	1.355	0.7949
Sample Variance	1.836	0.6319
Pooling Weight	0.5000	0.5000

OneWay ANOVA Table	Sum of Squares	Degrees of Freedom	Mean Squares	F-Ratio	p-Value
Between Variation	15.5741	1	15.5741	12.6229	0.0237
Within Variation	4.9352	4	1.2338		
Total Variation	20.5093	5			

		No Correction		Bonferroni		Tukey		Scheffe	
Confidence Interval Tests	Difference of Means	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Brainstorming-Premortem	-3.2222	-5.740277112	-0.704167333	-5.740277112	-0.704167333	-5.743596471	-0.700847974	-5.740277112	-0.704167333

Table 17 One-Way ANOVA Results for Average Top-3 Votes

The differences between the groups can be further analyzed using the expert panel scores for Quality Risks, Quality Changes to the Plan, and Black Swan Hits. Does one group benefit more from the premortem technique along the three scoring dimensions by the expert panel members? Recall that a Quality Risk is a risk that received a score of 70 or greater for both creativity and plausibility. A Quality Change to the Plan must have received of 70 or greater for both creativity and usefulness. A Black Swan Hit is a risk

identified by the team that, according to the expert panel member matched one of the four black swans from the Mark-14 case. Mann-Whitney nonparametric tests of the three groups (*Student, Government, and Industry*) for each of the three scoring dimensions (Quality Risks, Quality Changes to the Plan, and Black Swan Hits) yields some interesting results. In the nonparametric test, the alternative hypothesis is "brainstorming results are less than the premortem results." In the case of the *Student* group (Table 18), the alternative hypothesis cannot be statistically supported for all three of the scoring dimensions. It does not appear that the *Student* group gained any statistically significant advantage from using the premortem technique.

Group	Student		Student		Student	
Total Based on Expert Panel Scoring	Quality Risks		Quality Changes to the Plan		Black Swan Hits	
Treatment	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*
Sample Size	26	26	26	26	26	26
Mean	3.692	4.115	3.385	3.462	4.577	4.115
Standard Deviation	3.308	3.115	2.968	2.370	2.517	1.925
Median	3.0	3.5	3.5	4.0	4.0	4.0
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Don't Reject		Don't Reject		Don't Reject	
P-value	0.2927		0.3836		0.7817	

Table 18 Student Group Scores by Treatment with the Alternative Hypothesis that Brainstorming Scores are Less than Premortem

The same nonparametric test was applied to the results from the *Government* group (Table 19). The alternative hypothesis was not statistically supported for both the Quality Risk and Quality Changes to the Plan scoring dimensions. The alternative hypothesis was supported for the Government group for Black Swan Hits. It appears that

in the case of the *Government* group, teams that used the premortem technique had a statistical advantage in catching the black swan.

	Group	Government		Government		Government	
Total Based on Expert Panel Scoring		Quality Risks		Quality Changes to the Plan		Black Swan Hits	
	Treatment	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*
	Sample Size	26	26	26	26	26	26
	Mean	2.692	2.577	3.308	4.038	2.6540	6.577
Standard Deviation		2.604	2.831	2.589	2.615	2.1710	4.002
	Median	2.5	2.0	3.0	4.0	2.5	5.0
	Mann-WhitneyTest of the Null Hypothesis (5% significance)	Don't Reject		Don't Reject		Reject	
	P-value	0.6762		0.1107		<0.0001	

Table 19 Government Group Scores by Treatment with the Alternative Hypothesis that Brainstorming Scores are Less than Premortem

The *Industry* group (Table 20) appeared to benefit the most from using the premortem technique. The alternative hypothesis, using the nonparametric tests, was supported across all three of the scoring dimensions. We can conclude that of the three groups, the *Industry* teams saw a statistically significant advantage from using the premortem technique over the brainstorming technique across the three scoring dimensions. Is the *Industry* group's apparent benefit due to its increase in performance from using the premortem technique or underperformance when using the brainstorming technique?

Group	Industry		Industry		Industry	
Total Based on Expert Panel Scoring	Quality Risks		Quality Changes to the Plan		Black Swan Hits	
Treatment	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*
Sample Size	26	26	26	26	26	26
Mean	2.385	4.885	2.000	4.154	3.692	8.308
Standard Deviation	2.729	3.819	2.059	3.552	2.379	4.946
Median	1.5	4.5	2.0	4.5	4.0	8.5
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Reject		Reject		Reject	
P-value	0.0081		0.1290		0.0003	

Table 20 Industry Group Scores by Treatment with the Alternative Hypothesis that Brainstorming Scores are Less than Premortem

This question is answered by comparing the differences in the mean values for all three scoring dimensions, the three groups and the two treatments. Subtracting the Industry group mean from the *Student* and *Government* group means shows (Table-21) that when using the brainstorming technique, the Industry group performed within 1.385 points of the other two groups. The Industry group performed better using brainstorming (on average) than the *Government* group in the Quality Risks and Black Swan Hits scoring dimensions. A similar analysis between the three groups using the premortem technique shows that the Industry group performed better (on average) than both the *Student* and *Government* groups across all three scoring dimensions. This would indicate that the *Industry* group's performance came more from the benefit of using the premortem technique rather than a handicap in using the brainstorming technique.

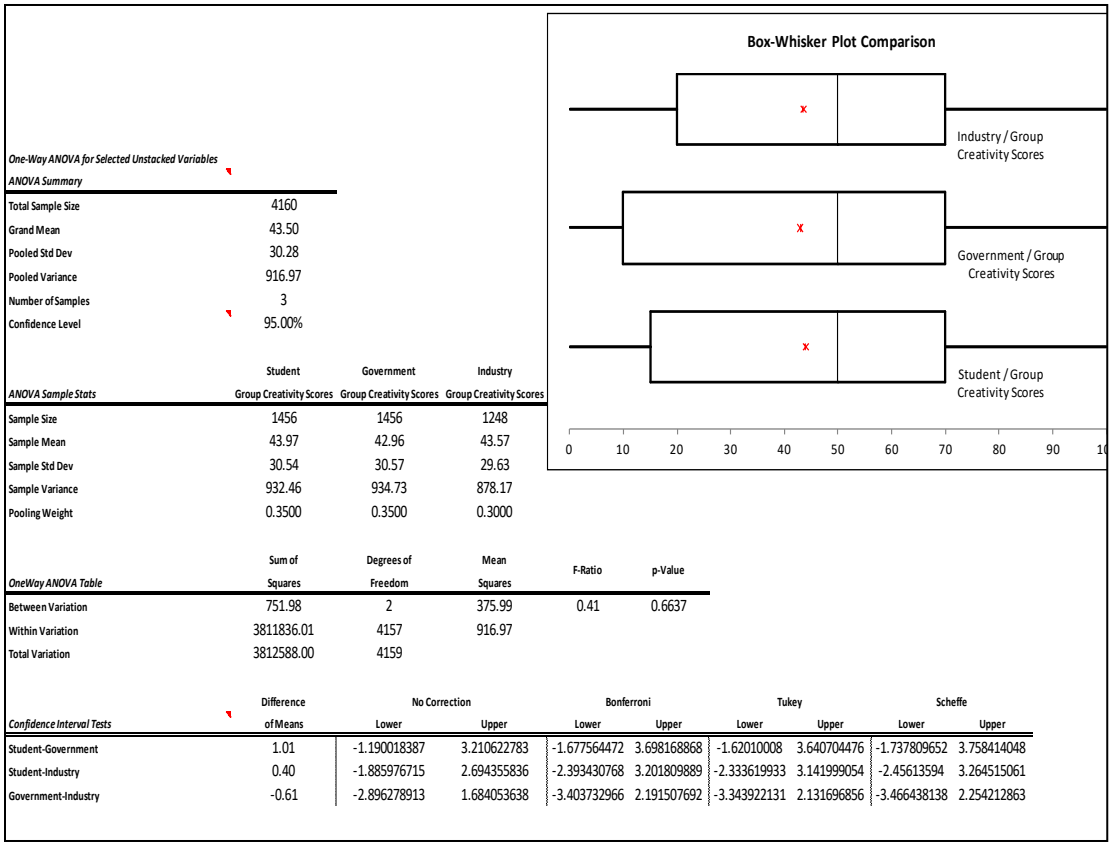
Industry Group Comparison of Means using Brainstorming			
	Quality Risks	Quality Changes to the Plan	Black Swan Hits
Student	0.000	-1.385	-0.885
Government	1.000	-1.308	1.038
Industry Brainstorming μ - Other Brainstorming μ			

Industry Group Comparison of Means using Premortem			
	Quality Risks	Quality Changes to the Plan	Black Swan Hits
Student	0.770	0.692	4.193
Government	2.308	0.116	1.731
Industry Premortem μ - Other Premortem μ			

Table 21 Comparing the Student and the Government Group Means to the Industry Group Means

Central to this research is the concept that the more creative we are in identifying risks and then changing the plan, the better we will be at reducing our vulnerability to black swan risks. Each experiment generated four risks and four changes to the plan. Every risk and every change to the plan received a creativity score from 0 to 100 from the 26 expert panel members. Therefore, every experiment generated 208 *creativity* data points. Analyzing those data points may reveal differences in creativity by group and treatment. Was one group more creative than the others? Did the premortem technique enable a greater degree of creativity than brainstorming?

A one-way ANOVA of all of the creativity scores (both treatments) by group showed no statistical difference in creativity between the groups (Table 22). The box-whisker plot within Table 22 provides a good visual representation of the data. With the exception that the *Government* group's interquartile range is slightly larger than the other two groups, the plots are nearly identical. The data indicated that no group was more creative than another.



One-Way ANOVA for Selected Unstacked Variables

ANOVA Summary

Total Sample Size	4160
Grand Mean	43.50
Pooled Std Dev	30.28
Pooled Variance	916.97
Number of Samples	3
Confidence Level	95.00%

ANOVA Sample Stats	Student Group Creativity Scores	Government Group Creativity Scores	Industry Group Creativity Scores
Sample Size	1456	1456	1248
Sample Mean	43.97	42.96	43.57
Sample Std Dev	30.54	30.57	29.63
Sample Variance	932.46	934.73	878.17
Pooling Weight	0.3500	0.3500	0.3000

OneWay ANOVA Table	Sum of Squares	Degrees of Freedom	Mean Squares	F-Ratio	p-Value
Between Variation	751.98	2	375.99	0.41	0.6637
Within Variation	3811836.01	4157	916.97		
Total Variation	3812588.00	4159			

Confidence Interval Tests	Difference of Means	No Correction Lower	No Correction Upper	Bonferroni Lower	Bonferroni Upper	Tukey Lower	Tukey Upper	Scheffe Lower	Scheffe Upper
Student-Government	1.01	-1.190018387	3.210622783	-1.677564472	3.698168868	-1.62010008	3.640704476	-1.737809652	3.758414048
Student-Industry	0.40	-1.885976715	2.694355836	-2.393430768	3.201809889	-2.333619933	3.141999054	-2.45613594	3.264515061
Government-Industry	-0.61	-2.896278913	1.684053638	-3.403732966	2.191507692	-3.343922131	2.131696856	-3.466438138	2.254212863

Table 22 One-Way ANOVA and Box Plot Comparing Group Creativity

Using the nonparametric Mann-Whitney hypothesis test in which the alternative hypothesis is that creativity using brainstorming is less than creativity using premortem, we find some very conclusive results (Table 23). For every group, the premortem technique enabled a statistically significant higher level of creativity within the teams. If creativity is our best weapon against the black swan, the data would suggest the premortem technique provides an advantage over brainstorming. Comparing treatment medians in Figure 48, the data indicates that the *Government* group experienced the greatest difference in creativity between teams that used brainstorming and teams that used the premortem. The *Student* group experienced the least difference in creativity between teams that used brainstorming and teams that used the premortem.

Creativity Scores By Group & Treatment						
Group	**Student**		**Government**		**Industry**	
Treatment	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*	*Brainstorming*	*Premortem*
Sample Size	832	624	624	832	624	624
Mean	38.61	51.13	35.58	48.50	36.97	50.17
Standard Deviation	30.24	29.47	30.19	29.69	28.76	29.04
Median	40	55	25	50	30	50
Mann-WhitneyTest of the Null Hypothesis (5% significance)	Reject		Reject		Reject	
P-value	<0.0001		<0.0001		<0.0001	

Table 23 Test Results Comparing Creativity by Group and Treatment (Alternative Hypothesis is Brainstorming Creativity Scores are Lower than Premortem Creativity Scores)

5.10 Goodman and Kruskal's Gamma, Strength of Association

The expert panel members selected their top three deliverable sets based on their own measure of quality in the role of the senior decision maker. Goodman and Kruskal's Gamma provides insight into the strength of association between the expert panel's level of deliverable quality (high, medium, or low) and other measures such as team size (large, medium, small or in our case six, five, four). A Gamma of "1" indicates a perfectly, positive association between the top-three ranking and the other associated measure. A Gamma of "0" would indicate no association between the two. A Gamma of "-1" would indicate a perfectly, inverse relationship between the top-three ranking and the other associated measure. The results are summarized in Table 24.

Gamma scores were recorded for the association between the team's top-three ranking and the quality measures for the deliverables (quality risks, quality changes to the plan, and black swan hits). Gamma scores were recorded for the top-three association with the team demographics of collective experience and team size. Gamma scores were

also recorded for the association between the team's top-three ranking and the team's collaboration and the four dimensions of collaboration recorded during structured observation.

Gamma Strength of Association with Top-3 Expert Panel Votes											
Associated Measure	Quality Risks	Quality Changes to the Plan	Black Swan Hits	Team Size	Team Experience	Team Total Collaborations	Collaborations per Person	Risk Ideas Offered	Enrichments Made	Critiques Made	Differences in Perspective
Sample Size	20	20	20	20	20	20	20	20	20	20	20
Gamma Score	0.7778	0.8851	0.6338	0.1471	0.2368	0.4063	0.3962	0.1692	0.2973	0.2667	0.7273
t-score	2.3479	3.9656	1.5439	0.2741	0.4752	0.7953	0.7025	0.3095	0.5989	0.4792	1.9249
Critical Value of t at 0.05	1.729	1.729	1.729	1.729	1.729	1.729	1.729	1.729	1.729	1.729	1.729
Statistically Significant	Yes	Yes	No	No	No	No	No	No	No	No	Yes

Table 24 Gamma Scores for Associations with Expert Panel Rankings Based on Top-3 Votes

With regard to the association between the expert panel top-three ranking and the quality scoring criteria, we find two strong and positive associations. For a risk or change to plan to receive a "quality" score, the expert panel member must have graded the two components for quality a 70 or higher. The components of quality for a risk were *creative and plausible*. The components for a quality change to the plan were *creative and useful*. We see a very strong and statistically significant association with the number of quality changes to the plan the team generated and the expert panel's Top-3 ranking for the team with a Gamma of 0.8851. This would indicate that the expert panel members (as the senior decision makers) found the quality of the change to the plan a significant factor in their top-three deliverables. In other words, if the team delivered highly creative and useful changes to the plan, it was likely to receive a top-three vote from an expert panel member. The inverse is also true. The expert panel top-three ranking and the number of quality risks also showed a positive and statistically significant association.

The Gamma of 0.7778 suggests that the expert panel looked favorably upon those risks that were both creative and plausible when considering their top-three deliverable sets. The number of black swan hits has a positive association with the expert panel ranking but weaker with a Gamma score of 0.6338 that is not statistically significant.

The association between the expert panel top-three rankings and the team demographics of team size and collective experience shows a positive and weak association with a Gamma of 0.1471 and 0.6338, respectively. Both scores are not statistically significant but imply that the differences in team size had little impact on creating a deliverable set that would receive an expert panel Top-3 score. A team's collective experience level has a higher association with the expert panel Top-3 ranking but the results are not statistically significant.

Of all of the dimensions of collaboration measured during structured observation, the total number of differences of perspective offered among the team members during the experiment had the strongest association with the expert panel member top-three rankings. The statistically significant Gamma of 0.7273 implies that those teams that had a high-level of sharing differences in perspective generated deliverable sets that the expert panel ranked among the highest quality. Though the score is not statistically significant, the very low Gamma of 0.1692 for the association between the top-three ranking and the total number of risk ideas the team generated during the experiment appears to run counter to the notion that generating a large volume of risks will lead to better risks.

5.11 Data Validity

Effort was taken to ensure the validity of the data collected. That effort began with multiple rounds of peer reviews and pilot testing to ensure consistency in the workshop and experiment conduct and the data collection techniques. Data collection instruments followed best practices in the design and were tested and revised before their use by the experiment subjects and the expert panel members. Subjects were recruited from a select target audience of experienced professionals. The expert panel members were selectively recruited based on their experience and subject area expertise. The expert panel members were anonymous to each other but known to the researcher. Expert panel score sheets were exchanged directly between the expert panel member and the researcher.

Teams were randomly assigned to their treatment. The teams were retrospectively analyzed for differences in subject experience, subject background, and the number of persons in the team. There was no statistical difference in those measures between teams performing the brainstorming treatment and the premortem treatment. Clear protocols for the experiments and the expert panel scoring were established and enforced. Three experiments were canceled because they failed to meet the minimum team size protocol. One set of experiment data has been excluded because the subject team broke protocol during the experiment. Two expert panel score sheets were excluded because the expert panel members failed to comply with the score sheet instructions.

Structured observation was the most difficult of the data collection methods used for this study. It is also the most subjective. The peer review feedback helped mitigate observer biases. The subject surveys and expert panel scores were used to either confirm or contradict the structured observation results.

Results related to hypothesis testing were obtained through both nonparametric and parametric means. Where the nonparametric and parametric test results differed, the results were reported to be inconclusive. Goodman and Kruskal's Gamma was used to study the strength of associations between results of different data collection methods.

5.12 Anecdotal Subject Reports and Researcher Observations

Over the course of the experiments, the subjects offered unsolicited comments regarding the risk identification techniques. Please recall that all the subjects experienced the premortem technique. The subjects that performed brainstorming in the experiment also practiced the premortem. The brainstorming subjects were in the position to compare and contrast both techniques. I paraphrase the subject reports below:

- While using the premortem technique, without a good facilitator, the team can easily revert to using brainstorming.
- Premortem might be best used when the team is under a time pressure. It is when the team is under time pressure that brainstorming is most likely to lead to group-think.
- Brainstorming requires an upfront understanding of the project context before the team can conduct risk identification. Using the premortem technique, the project context is less important.
- Brainstorming is likely to remain the method we use simply because everyone is familiar with the process.
- The premortem technique makes it nearly impossible to explain away problems.

- After using the premortem to consider changes to the plan, we never said, "there is nothing we can do about this risk." We found a way to change the plan to deal with even the most improbable risks.
- Changes to the plan to address the most creative yet plausible risks often addressed multiple risks that were more typical as well.

Below are my observations from the experiments using brainstorming for risk identification:

- Although brainstorming is popular and its application varies widely, the primary rule of recording ideas first without criticism or discussion before collaboration does not appear to be universally practiced. Several subjects wanted to debate the merits of another subject's identified risk as it was stated. On those occasions, the team leader (acting as the facilitator) had to enforce the rule.
- Brainstorming quickly generated risk ideas but the team often struggled with what to do with those ideas. Some groups spent a good portion of the 45-minute experiment discussing the similarities between the risks identified and how to properly bin those risks. Those bins would frequently evolve into the common cost, schedule, and performance categories.
- Subjects would often offer risk ideas that appeared to be from their personal favorite or familiar topics.

- None of the brainstorming teams compared the risks the team identified with the set of risks that the notional project sponsor in the scenario had declared to be unacceptable.
- If a team member did not have a risk idea to offer during brainstorming and "passed," the remainder of the team members "passed" within the next round, ending the brainstorming portion of the experiment and beginning the team collaboration phase. The quick transition from the first "pass" to ending the brainstorming session may be a lesser known pitfall of brainstorming.
- When the team completed the risks using brainstorming, new risk ideas would be offered and discussed but rarely were those new risks added to the risk list.

Below are my observations from the experiments using the premortem technique for risk identification:

- The team leader's lack of familiarity with the premortem technique caused the team leader to refer to the process worksheet frequently. The methodical approach allowed most premortem teams to spend the majority of the 45-minute experiment in useful discussion rather than comparing similarities between the risks.
- Most teams using the premortem technique referred back to the risks in the scenario that the notional project sponsor had declared to be unacceptable. It appears the teams used those risks as benchmarks for risks that were not creative.

- Premortem teams added to their risk lists during collaboration. Even with adding new risks late in the experiment, premortem teams had no problem narrowing down the risks to their top four.
- Premortem teams rarely discussed binning risk ideas or changes to the plan in accordance with the common cost, schedule, and performance areas.
- Some subjects appeared to be uncomfortable with the premortem technique. Those subjects failed to provide significant contribution throughout the entire session or contributed after the other team members had identified the risks. In other words, some subjects opted out of the creativity portion of the method (which is nearly impossible to do in brainstorming).

I performed a brief follow-up with select expert panel members and recorded the following observations:

- When asked why an expert panel member scored the risks consistently low, that expert panel member said that the risks were not in "if, then" statement form. This is the doctrinal method of describing risks in defense acquisition programs.
- Expert panel members reported that the 36-inch-by-24-inch paper scoring sheet made it easier to remain consistent in scoring the twenty deliverable sets. This would have been more difficult to do in electronic form.

5.13 Summary of Results

The results of this study indicate that the premortem technique is better than the brainstorming technique at identifying risk when used in a systems engineering context.

- **Alternative hypothesis # 1 is partially supported.** Teams conducting the premortem technique build on one another's ideas, offer differing perspectives, and critique each other more than teams conducting brainstorming. We cannot conclude that teams using the premortem technique generate more risk ideas than teams using the brainstorming technique. Subject survey responses confirm these results.
- **Alternative hypothesis #2 is supported.** The twenty-six expert panel members found the premortem technique identified more quality risks than the brainstorming technique.
- **Alternative hypothesis #3 is supported.** The twenty-six expert panel members found that teams using the premortem technique proposed more quality changes to the plan than teams using the brainstorming technique.
- **Alternative hypothesis #4 is supported**. The twenty-six expert panel members found that teams using the premortem technique had more black swan hits than teams using the brainstorming technique.
- The deliverables generated using the premortem technique received more top-three votes from the twenty-six expert panel members than the deliverables generated by teams using the brainstorming technique. This indicates a higher degree of senior leader acceptability when the team generates risks and changes to the plan using the premortem technique.

- Subjects using the premortem technique reported that technique generated more risks when compared to their personal benchmark than those subjects using the brainstorming technique.
- Subjects using the premortem technique reported that technique generated better risks when compared to their personal benchmark than those subjects using the brainstorming technique.
- Subjects using the premortem technique reported that technique generated more creative improvements to the plan when compared to their personal benchmark than those subjects using the brainstorming technique.
- Subjects using the premortem technique indicated that the premortem would help avoid organizational blind spots more than those subjects using the brainstorming technique.
- Based on survey responses, there is a statistically significant dissonance between the commonly perceived strategy to achieve program success through predicting risks and planning for their occurrence and the predictability and orderliness of our projects that would allow that strategy to be successful.
- Based on the survey responses, the subjects believed that gaps between the programmatic and technical sides of a project are a problem.
- The premortem technique provided an advantage in delivering quality risks and changes to the plan for all three groups (Student, Government, and Industry).
- Government teams had a statistical advantage in black swan hits (caught more black swans) when they used the premortem technique.

- Industry teams that used the premortem technique received the largest statistical advantage over their brainstorming counterparts.
- Based on Goodman and Kruskal's Gamma, there was no statistical difference between groups regarding their creativity scores. However, every group was more creative when they used the premortem technique.
- Based on Goodman and Kruskal's Gamma, there is a strong, positive association between what the expert panel members (surrogates for senior decision makers) regard as a top-deliverable set and the quality of the changes to the plan in the deliverable.
- Based on Goodman and Kruskal's Gamma, there is a strong, positive association between what the expert panel members (surrogates for senior decision makers) regard as a top-deliverable set and the creativity and plausibility of the risks identified the deliverable.
- Based on Goodman and Kruskal's Gamma, there is a strong, positive association between what the expert panel members (surrogates for senior decision makers) regard as a top-deliverable set and those teams that had a high-level of sharing differences in perspective.

Chapter 6 - Conclusions and Recommendations

6.1 Conclusions

The results reported in Chapter 5 can be examined in terms of potential opportunities for career education providers, professional organizations, senior decision makers, project managers, engineering managers and systems engineers. The target audience for the recommendations for education providers would be advanced degree colleges and universities that include risk management in their curriculum, Department of Defense senior military education providers such as the service war colleges and federal government institutions such as the Federal Acquisition Institute, the National Defense University, and the Defense Acquisition University. The target audience for the recommendations for professional organizations is the professional certification providers for systems engineers (INCOSE) and project managers (PMI). The target audience for recommendations for senior decision makers, project managers, engineering managers, and systems engineers are those individuals working in the defense and aerospace fields.

Synthesizing the analytical and anecdotal results from the brainstorming versus premortem experiments and the literature review, I offer the following conclusions:

- Knowing the *Taxonomy of Ignorance* and the symptoms of deliberate ignorance can lead to more productive collaborations.
- Brainstorming continues to be the "go to" tool for creativity, but it is misapplied.

- Senior decision makers expect teams to bring them creative solutions, but there is very little support to teams to help them be more creative.
- There are very few empirical studies on risk identification techniques requiring risk management practitioners to resort to trial and error.
- Quantity of risks does not guarantee quality but may result in a false sense of security.
- The premortem technique makes teams think and act differently.
- Creativity-based risk identification may be our best defense against black swan risks.

6.2 Recommendations

6.2.1 Taxonomy of Ignorance Awareness

Realized risks due to distorted or incomplete information are unavoidable. However, deliberate ignorance is a human behavior that we can study and manage. Risks due to the components of deliberate ignorance (untopicality, undecidability, taboo topics, and suspension of belief) are avoidable. Career education institutions should develop curriculum that allows students to practice identifying the symptoms of deliberate ignorance. This topic could be covered in a series of case studies and practical exercises. Project managers are on the front lines in the fight against deliberate ignorance and perhaps have the most to lose when risks emerge as black swans from deliberate ignorance. PMI should consider presenting the *Taxonomy of Ignorance* in the Project Management Body of Knowledge and disseminate best practices in dealing with deliberate ignorance. Senior decision makers normally make decisions from a list of recommendations that came from a team. The senior decision maker should understand

the team's processes used to generate those recommendations. If the processes did not actively counter deliberate ignorance, there may be a black swan hidden in the recommendation. Project managers, engineering managers, and systems engineers should establish team rules and roles to maintain a team's self awareness regarding black swans caused by deliberate ignorance.

6.2.2 Brainstorming Training and Discipline

Brainstorming is likely to remain the "go to" tool for team creativity. Among the participants in this study it was the most common technique for risk identification. Its ubiquity and simplicity may be its greatest weakness as risk identification tool. It appeared many of the brainstorming participants had performed the technique numerous times. The subjects using brainstorming rarely referenced the process worksheet provided to them. This is a sign of their comfort and perhaps overconfidence in the technique. However, some teams failed to follow the primary rule of no criticism or discussion until all the ideas are exhausted. If teams continue to prefer to use brainstorming, career education providers should make a point to train the technique properly and reinforce adherence to the primary rule. Both INCOSE and PMI should address the capabilities and limitations of brainstorming in their publications and training. Senior decision makers should know that risks generated through brainstorming are likely to be typical risks. This will require senior decision makers to challenge their teams to expand their field of view to consider the atypical risks. Project managers, engineering managers, and systems engineers should avoid ad hoc brainstorming. Brainstorming requires training, practice, and disciplined execution to achieve results.

6.2.3 Senior Decision Makers Expect Creative Solutions

Assuming the expert panel members in this study were adequate surrogates, senior decision makers expect their teams to deliver creative changes to the plan as part of risk identification. Too often, when our teams are faced with risks they spend their time and energies trying to fortify the plan. Teams will often accept a risk, saying there is nothing to be done. If this risk happens the project is dead. Clearly our surrogate senior decision makers preferred creative and useful changes to the plan. Career education providers need to train professionals to consider changing the plan before they consider fortifying the plan. Curriculum that develops creativity skills will help professionals deliver those creative changes senior decision makers expect. PMI and INCOSE should consider adding "change the plan" as a risk handling step along side mitigate, transfer, and control and accept. Project managers, engineering managers, and systems engineers should become more comfortable with creativity tools. This can only come with training and practice.

6.2.4 More Study Needed on Risk Identification Techniques

We saw in Chapter 2 that there are few empirical studies on risk identification techniques. Risk management practitioners are left to trial and error in selecting a risk identification technique for a project. More studies are needed to develop risk identification processes that are result-based and appropriate to the situation. Career education providers should encourage students to perform empirical studies on risk identification techniques. Professional organizations can compile best practices and track

the results of specific risk identification techniques as they are practiced within their fields.

6.2.5 More Risks Does Not Guarantee Better Risks

The premise of brainstorming is that by generating a large quantity of ideas followed by group collaboration, a few high quality ideas will emerge. This study indicates that the premise did not hold true for our risk identification scenario. For career education institutions, they should offer alternative approaches to brainstorming risks such as the premortem. INCOSE and PMI should provide caveats in their respective handbooks identifying the limitations of brainstorming as risk identification tool. Senior decision makers should question the quality of risks presented to them if brainstorming was the primary means of generating the risks. A large quantity of risk items may lead the senior decision maker to conclude that the team has covered all the possible risks. This can lead to overconfidence in the plan and black swan blindness. Project managers, engineering managers, and systems engineers should calculate the costs and benefits of identifying risks via brainstorming. Brainstorming is likely to generate a large quantity of risks, each with its own costs for management. If those risks are typical and predictable, they may not be worth the cost of managing. Developing changes to the plan to address the risks that are the most creative yet plausible may be more cost effective. When a team changes the plan to account for atypical risks, they are likely to be simultaneously addressing several typical risks.

6.2.6 Teams Think and Act Differently Using Premortem

Teams using the premortem appeared to think and act differently using the premortem technique than their brainstorming counterparts. They seemed to generate risk ideas more efficiently by delivering more quality risks in the same amount of time. Perhaps the act of having to focus on the steps in a novel process contributed to the efficiency. For career education institutions, it may be very beneficial to provide students with process worksheets as quick reference job aids to make it easier for students to apply new tools within their organization. Professional institutions such as INCOSE and PMI should offer continuing education on novel processes such as the premortem. Senior decision makers should encourage their managers to adopt new processes to stimulate creative thinking in teams. Project managers, engineering managers, and systems engineers should experiment with new tools and techniques by taking a "split squad" approach. For example, have half of the team review a plan for risks using brainstorming and half do the same task using premortem. Then, have each squad present their findings to the other squad and discuss the differences. Although it is possible to have the whole team use one technique followed by the whole team using an alternate technique, the team may be less receptive to repeating the process on a second technique. Even if the second technique is novel, one could expect the team to feel as though they are plowing the same field again.

6.2.7 Creativity-Based, Our Best Defense Against the Black Swan

Quantitative methods or risk identification based on historical information, updated with project status reports should continue to be important elements of the project manager, engineering manager, and systems engineer tool kits. As we have seen, these methods lack the ability to detect unprecedented risks or those risks out of the

team's field of view. Teams appear to be largely receptive to creativity-based approaches such as brainstorming and premortem but are uncomfortable using these techniques. If creativity-based approaches are our best defense against black swan risks, teams will need to get more comfortable employing individual creativity. For career education institutions, there is a need for a cadre of trained facilitators to help teams use creativity-based approaches. Curriculum should balance creative and quantitative techniques. INCOSE and PMI should add creativity-based risk identification as a process in their respective handbooks. The two professional organizations can also train and certify facilitators in creativity-based risk identification techniques. Senior decision makers should insist team use both quantitative and creativity- based risk identification techniques as a broader approach to risk identification. Project managers, engineering managers, and systems engineers should evaluate their plans using creativity-based approaches. The findings in this study show the premortem is a good technique that overcomes the problems of deliberate ignorance. A team that collaborates more and better by sharing differences in perspective is less likely to be blindsided by black swans.

6.3 Recommendations for Future Research

This study was limited in time and scope. There are many areas for future research regarding creativity-based risk identification, brainstorming, and the premortem technique.

6.3.1 Before and After Study

This study used a classic Treatment-A and Treatment-B experimental design. A limitation of this approach is that it lacks a baseline to measure a change in a team's

performance after the treatment is introduced. A before and after study that measures improved risk identification performance or a change in attitudes toward collaboration on the same team after premortem is introduced would be a valuable complement to the findings in this study.

6.3.2 Change the Scenario

I used a brownfield, technology insertion scenario for the experiments. Would the results of the experiments be similar if the team used a greenfield, new product development scenario? The scenario in this study's experiments was based on the Mark-14 torpedo scenario and the black swans that blindsided the team in the 1940s. Would the results change if the scenario was more contemporary? Would the results change if the team used their assigned treatment on a project in which they had a stake in the outcome?

6.3.3 Integrate the Team

The experiments in this study used experienced professionals, but they were not a mix of government and industry personnel working on their own project. Mixing of government and industry is typical in defense and aerospace programs. Would the results differ in an integrated product team environment?

6.3.4 Longitudinal Study

We do not know if there are additional benefits or perhaps costs associated with continued, repeated use of the premortem technique. One of the limitations of the premortem technique is risk aversion. Does repeated and continued use of the premortem

make naturally risk adverse organizations, staunch risk avoiders? This behavior could paralyze an organization. A longitudinal study of the impacts to an organization of continued, repeated use of the premortem could answer that question.

6.3.5 A Different Treatment-A

In this study, the premortem technique was compared to brainstorming. I mentioned the reasons for selecting brainstorming as a comparison to the premortem in Chapter 2. Among those reasons were the similarities between the two techniques. Would the results in this study hold true if a radically different technique were used for comparison with the premortem? Staying within the creativity-based approaches, the devil's advocate would be an interesting comparison. It would also be interesting to test different instantiations of brainstorming such as asynchronous or electronically facilitated brainstorming.

6.3.6 Further Analysis of the Structured Observation Data

The data from structured observation could be further analyzed to compare the premortem technique against the brainstorming pitfalls. Does the data indicate the premortem technique is less susceptible to the common brainstorming pitfalls? The collaboration data could also be used to develop utility functions for both techniques. With the utility function, one could determine the optimal team size and duration of a premortem session.

References

Abkowitz, Mark. 2008. *Operational Risk Management: A Case Study Approach to Effective Planning and Response*. Hoboken, Nj: John Wiley & Sons.

ACP. 2014. "A Primer on Before/After Studies: Evaluating a Report of a 'Successful Intervention.'" *Effective Clinical Practice*. http://ecp.acponline.org/primers/sepoct99.htm.

Altabbakh, Hanan, Susan Murray, Katie Grantham, and Siddharth Damle. 2013. "Variations in Risk Management Models : A Comparitive Study of the Space Shuttle Challenger Disaster." *Engineering Management Journal* 25 (2): 13–24.

Amer, Muhammad, Tugrul U. Daim, and Antonie Jetter. 2013. "A Review of Scenario Planning." *Futures* 46: 23–40. doi:10.1016/j.futures.2012.10.003.

Antonio, Brian. 2015. "PEO LCS Program Update." NAVSEA Small Business Industry Day. Washington, DC.

Ashton, Kevin. 2015. *How to Fly a Horse: The Secret History of Creation, Invention and Discovery*. New York: Knopf Doubleday Publishing Group.

Aven, Terje. 2015. "Implications of Black Swans to the Foundations and Practice of Risk Assessment and Management." *Reliability Engineering and System Safety* 134. Elsevier: 83–91. doi:10.1016/j.ress.2014.10.004.

Aven, Terje, and Bodil S Krohn. 2013. "A New Perspective on How to Understand , Assess and Manage Risk and the Unforeseen." *Reliability Engineering and System Safety* 121. Elsevier: 1–10. doi:10.1016/j.ress.2013.07.005.

Ayton, Peter, and Alastair G. R. McClelland. 1997. “How Real Is Overconfidence?” *Journal of Behavioral Decision Making* 10 (3): 279–85. doi:10.1002/(SICI)1099-0771(199709)10:3<279::AID-BDM280>3.3.CO;2-E.

Barker, Joel. 2000. “Act like an Adult : The Truth behind Unintended Consequences.” *The Journal for Quality and Participation* 23 (1): 46–47.

Bartram, Peter. 2015. “Seven Quick Tips for Preparing for the Worst.” *CGMA Journal*, 6–8.

Beasley, Mark, Bruce Branson, and Bonnie Hancock. 2015. “2015 Report on the Current State of Enterprise Risk Oversight: Update and Trends.” www.erm.ncsu.edu.

Bernstein, Peter. 1998. *Against the Gods: The Remarkable Story of Risk*. New York: John Wiley & Sons.

Brassard, Michael, and Diane Ritter. 1994. *Memory Jogger II: A Pocket Guide of Toolsfor Continuous Improvement& Effective Planning*. Edited by Francine Oddo. First. Salem, NH: Goal/QPC. http://www.goalqpc.com.

Briggs, Robert O., and Bruce a. Reinig. 2010. “Bounded Ideation Theory.” *Journal of Management Information Systems* 27 (1): 123–44. doi:10.2753/MIS0742-1222270106.

Brown, Vincent, and Paul Paulus. 2015. “A Simple Dynamic Model of Social Factors in Group Brainstorming.” *Small Group Research* 27 (1): 91–114. doi:10.1177/1056492611432802.

Brown, Vincent, Michael Tumeo, Timothy Larey, and Paul Paulus. 1998. “Modeling

Cognitive Interactions During Group Brainstorming." *Small Group Research* 29 (4): 495–526. doi:10.1177/1056492611432802.

Buchanan, Mark. 2000. *Ubiquity*. London, UK: Widenfeld & Nicolson.

Buhl, Hans (University of Augsburg). 2012. "The Contribution of Business and Information Systems Engineering to the Early Recognition and Avoidance of ' Black Swans ." *Business & Information Systems Engineering* 2: 55–59. doi:10.1007/s12599-012-0206-8.

Carifio, James, and Rocco J. Perla. 2007. "Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and Their Antidotes." *Journal of Social Sciences* 3 (3): 106–16. doi:10.3844/jssp.2007.106.116.

Carpenter, Tina D. 2014. "Audit Team Brainstorming , Fraud Risk and Fraud Risk Identification , Assessment : Of SAS No . 99 Implications." *The Accounting Review* 82 (5): 1119–40.

Chen, Fang, Jay F. Nunamaker, Robert O. Briggs, Gail Corbitt, James Sager, and Stanley C. Gardiner. 2013. "An Application of Focus Theory to Project Management Processes." *Group Decision and Negotiation* 23 (5): 961–78. doi:10.1007/s10726-012-9323-5.

Chen, Tiffani R., Daniel B. Shore, Stephen J. Zaccaro, Reeshad S. Dalal, Lois E. Tetrick, and Aiva K. Gorab. 2014. "An Organizational Psychology Perspective to Examining Computer Security Incident Response Teams." *IEEE Security & Privacy* 12 (5): 61–67. doi:10.1109/MSP.2014.85.

Chiles, James. 2002. *Inviting Disaster: Lessons From the Edge of Technology*. Paperback. New York: Harper Collins Publishers.

Chorafas, Dimitris. 1994. *Chaos Theory in Financial Markets*. Chicago: Probus.

Conway, Patrick. 2012. "Red Team: How the Neoconservatives Helped Cause the Iraq Intelligence Failure." *Intelligence and National Security* 27 (4): 488–512. doi:10.1080/02684527.2012.688304.

Copi, Irving, and Carl Cohen. 2009. *Introduction to Logic*. Strayer Un. New York: Pearson Custom Publishing.

Craig, Susan. 2007. "Reflections from a Red Team Leader." *Military Review*, no. March-April: 57–60.

DAU. 2014. "Courseware Review, Changes, and Publication Process." DAU Directive 702. Fort Belvoir, VA.

DeBakker, Karel, Albert Boonstra, and Hans Wortmann. 2011. "Risk Management Affecting IS/IT Project Success Through Communicative Action." *Project Management Journal* 42 (3): 75–90. doi:10.1002/pmj.20242.

DeBono, Edward. 1999. *Six Thinking Hats*. Second. New York: Back Bay Books.

DeCarlo, Doug. 2004. *Extreme Project Management: Using Leadership, Principles, and Tools to Deliver Value in the Face of Volatility*. First. SanFrancisco, CA: Jossey-Bass.

"Defense Systems Management College." 2016. Accessed April 29. http://www.dau.mil.

Diehl, M, and W Stroebe. 1987. "Productivity Loss In Brainstorming Groups: Toward the

Solution of a Riddle." *Journal of Personality and Social Psychology* 53 (3): 497–509.

Donnelly, Richard, Jason Clement, Richard Le Heron, and John St George. 2012. "Redesigning Risk Frameworks and Registers to Support the Assessment and Communication of Risk in the Corporate Context: Lessons from a Corporate Risk Manager in Action." *Risk Management* 14 (3). Nature Publishing Group: 222–47. doi:10.1057/rm.2012.3.

DSMC. 2012. "Instructor Peer Review Process." Fort Belvoir, VA.

El Badawy, Tarek A., Rania M. Marwan, and Mariam M. Magdy. 2015. "The Impact of Emerging Technologies on Knowledge Management in Organizations." *International Business Research* 8 (5): 111–20. doi:10.5539/ibr.v8n5p111.

Elm, Joseph P, and Dennis R Goldenson. 2012. "The Business Case for Systems Engineering Study: Results of the Systems Engineering Effectiveness Survey: Special Report." Hanscom AFB, MA. http://www.sei.cmu.edu.

Eriksen, Christoffer. 2015. "ON CREATIVITY: A CASE STUDY OF MILITARY INNOVATION." NAVAL POSTGRADUATE SCHOOL.

Everett, Stuart, and Tony Mulholland. 2006. "Challenger: The Untold Story." USA: National Geographic Television.

Faily, Shamal, Simon Parkin, and John Lyle. 2014. "Evaluating the Implications of Attack and Security Patterns with Premortems." In *Cyberpatterns:Unifying Design Patterns with Security and Attack Patterns*, edited by Clive Blackwell and Hong

Zhu. Springer.

Farrington, Ted, Christian Crews, and Jennifer Blenkle. 2013. “Implications Wheels: Structured Brainstorming about the Future.” *Research Technology Management* 56 (3): 56–58. doi:10.5437/08956308X5603008.

Farrington, Ted, Keith Henson, and Christian Crews. 2012. “Research Foresights: The Use of Strategic Foresight Methods for Ideation and Portfolio Management.” *Research Technology Management* 55 (2): 26–33. doi:10.5437/08956308X5502023.

Forest, Ed. 2016. “The ADDIE Model: Instructional Design.” *Educational Technology*. Accessed May 3. http://educationaltechnology.net/the-addie-model-instructional-design/.

Foster, H H. 1978. “The Devil’s Advocate.” *Bull Am Acad Psychiatry Law* 6 (2): 256–58.

Fowler, Floyd. 1995. *Improving Survey Questions: Design and Evaluation*. Applied So. New York: Sage.

Franklin, Benjamin. 1772. “Benjamin Franklin’s 1772 Letter to Joseph Priestley.” *Mr. Franklin: A Selection from His Personal Letters.* http://www.procon.org/view.background-resource.php?resourceID=1474.

French, Simon. 2015. “Cynefin: Uncertainty, Small Worlds and Scenarios.” *Journal of the Operational Research Society* 66 (10): 1635–45. doi:10.1057/jors.2015.21.

Gallupe, R. B., a. R. Dennis, W. H. Cooper, J. S. Valacich, L. M. Bastianutti, and J. F. Nunamaker. 1992. “Electronic Brainstorming and Group Size.” *Academy of Management Journal* 35 (2): 350–69. doi:10.2307/256377.

GAO. 2015. “Assessments of Selected Weapon Programs.” *Report to Congressional Committees* GAO-15-342 (March): 1–194. doi:GAO-06-391.

Gobble, MaryAnne M. 2014a. “Beyond Brainstorming.” *Research Technology Management* 57 (2): 60–62. doi:10.5437/08956308X5702005.

———. 2014b. “The Persistence of Brainstorming.” *Research Technology Management* 57 (1): 64–66. doi:10.5437/08956308X5701005.

Gold, Ted, and Bob Hermann. 2003. “The Role and Status of DoD Red Teaming Activities.” *Defense Science Board* 298 (September).

Goodman, Leo, and William Kruskal. 1963. “Measures of Association for Cross Classifications III: Approximate Sampling Theory.” *Journal of American Statis* 58 (302): 310–64.

Green, Leslie. 2002. *What’s Left of Critique?* Edited by Wendy Brown and Janet Halley. Durham, NC: Duke University Press.

Greenberg, Leslie S., and Michael C. Webster. 1982. “Resolving Decisional Conflict by Gestalt Two-Chair Dialogue: Relating Process to Outcome.” *Journal of Counseling Psychology* 29 (5): 468–77. doi:10.1037/0022-0167.29.5.468.

Haddou, Hassan Ait, Guy Camilleri, and Pascale Zaraté. 2014. “Prediction of Ideas Number During a Brainstorming Session.” *Group Decision and Negotiation* 23 (2): 271–98. doi:10.1007/s10726-012-9312-8.

Hagel, Jack. 2012. “The Merits of Thinking the Unthinkable.” *Journal of Accountancy* 214 (4): 18.

Heard, Wayne. 2016. “Brainstorm: Soliciting Variety of Ideas Yields Better Results.” *ARMY Magazine,* April.

Herazo, Benjamin, Gonzalo Lizarralde, and Raymond Paquin. 2012. “Sustainable Development in the Building Sector; A Canadian Case Study on the Alignment of Strategic and Tactical Management.” *Project Management Journal* 43 (2): 84–100. doi:10.1002/pmj.

Hobbs, D. 1994. “Adapting Strategic Aircraft Assets to a Changing World--Technology Insertion to Provide Flexibility.” Maxwell Airforce Base, AL.

Hossain, Liaquat, and Shahadat Uddin. 2012. “Design Patterns: Coordination in Complex and Dynamic Environments.” *Disaster Prevention and Management* 21 (3): 336–50. doi:10.1108/09653561211234516.

Howat, Colin. 2008. “Fault Tree Analysis.” *Introduction to Fault Tree Analysis*. Lawrence, Kansas.

Hubrecht, Robert, and James Kirkwood. 2010. *The UFAW Handbook on the Care and Management of Laboratory and Other Research Animals*. Edited by R Hubrecht and J Kirkwood. 8th ed. Ames, Iowa: Wiley-Blackwell.

Humphrey, Albert S. 2005. “SWOT Analysis for Management Consulting.” *SRI Alumni Association Newsletter,* no. December: 7,8. http://www.sri.com/sites/default/files/brochures/dec-05.pdf.

IBM. 2006. *SWOT Analysis*. *IBM Corporation SWOT Analysis*. doi:10.1109/ICEEE.2010.5660252.

INCOSE. 2014. “A World in Motion: Systems Engineering Vision 2025.” www.incose.org.

———. 2015a. “INCOSE AT 25.” San Diegp, CA.

———. 2015b. *Systems Engineering Handbook*.

J Brooks, David. 2011. “Security Risk Management: A Psychometric Map of Expert Knowledge Structure.” *Risk Management* 13 (1-2). Nature Publishing Group: 17–41. doi:10.1057/rm.2010.7.

James, K., and D. Drown. 2008. “Whether ‘malevolent’ or ‘negative,’ Creativity Is Relevant to Terrorism Prevention: Lessons from 9/11 and Hazardous Material Trucking.” *Creativity Research Journal* 20 (2): 120–27. doi:10.1080/10400410802059648.

Kahneman, Daniel. 2003. “Maps of Bounded Rationality: Maps of Bounded Rationality: Psychology for Behavioral Economicst.” *The American Economic Review* 93 (5): 1449–75. http://www.jstor.org/stable/3132137.

Kaufmann, Lutz. 2011. “Findings Sweet and Sour.” *JSCM Journal of Supply Chain Management* 47 (4): 17–18. doi:10.1111/j.1745-493X.2011.03247.x.

Keeney, Ralph. 2013. “Foundations for Group Decision Analysis.” *Decision Analysis* 10 (2): 103–20. http://dx.doi.org/10.1287/deca.2013.0265.

Kerr, C, R Phaal, and D Probert. 2008. “Technology Insertion in the Defence Industry : A Primer” 222 (i): 1009–24. doi:10.1243/09544054JEM1080.

Khakzad, Nima, Faisal Khan, and Paul Amyotte. 2011. “Safety Analysis in Process

Facilities: Comparison of Fault Tree and Bayesian Network Approaches." *Reliability Engineering and System Safety* 96 (8). Elsevier: 925–32. doi:10.1016/j.ress.2011.03.012.

Kirschner, Paul, John Sweller, and Richard Clark. 2006. "Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching." *Educational Psychologist* 41 (2): 75–86. doi:10.1207/s15326985ep4102_1.

Klein, Gary. 2004. *The Power of Intuition: How to Use Your Gut Feelings to Make Better Decisions at Work*. New York: Currency Books.

———. 2007a. "Performing a Project PreMortem." *Harvard Business Review*. www.hbr.org.

———. 2007b. "Performing a Project Premortem." *Harvard Business Review* September (Reprint F0709A): 18–19.

———. 2008. "Naturalistic Decision Making." *Human Factors* 50 (3): 456–60. doi:10.1518/001872008X288385.

———. 2009. *Streetlights and Shadow: Searching for the Keys to Adaptive Decision Making*. Cambridge, MA: MIT Press.

———. 2015. "Premortem versus Fault Tree Analysis."

Kohavi, Ron, and Roger Longbotham. 2015. "Online Controlled Experiments and A/B Tests." *Encyclopedia of Machine Learning and Data Mining*. http://bit.ly/onlineControlledExperiments.

Kohavi, Ron, Roger Longbotham, Dan Sommerfield, and Randal M. Henne. 2009. "Controlled Experiments on the Web: Survey and Practical Guide." *Data Mining and Knowledge Discovery* 18 (1): 140–81. doi:10.1007/s10618-008-0114-1.

Kohn, Nicholas W., and Steven M. Smith. 2011. "Collaborative Fixation: Effects of Others' Ideas on Brainstorming." *Applied Cognitive Psychology* 25 (3): 359–71. doi:10.1002/acp.1699.

Konate, J, A Sahraoui, and Gwendolyn L Kolfschoten. 2013. "Collaborative Requirements Elicitation: A Process-Centred Approach." *Group Decision and Negotiation* online fir: 847–77. doi:10.1007/s10726-013-9350-x.

Konno, Noboru, Ikujiro Nonaka, and Jay Ogilvy. 2014. "Scenario Planning: The Basics." *World Futures* 70 (1): 37–41. doi:10.1080/02604027.2014.875720.

Kutsch, Elmar, David Denyer, Mark Hall, and Elizabeth (Liz) Lee-Kelley. 2013. "Does Risk Matter? Disengagement from Risk Management Practices in Information Systems Projects." *European Journal of Information Systems* 22 (6). Nature Publishing Group: 637–49. doi:10.1057/ejis.2012.6.

Kutsch, Elmar, and Mark Hall. 2010. "Deliberate Ignorance in Project Risk Management." *International Journal of Project Management* 28 (3). Elsevier Ltd and IPMA: 245–55. doi:10.1016/j.ijproman.2009.05.003.

Leckie, Gloria, Karen Pettigrew, and Christian Sylvian. 1996. "MODELING THE INFORMATION SEEKING OF PROFESSIONALS: A GENERAL MODEL DERIVED FROM RESEARCH ON ENGINEERS, HEALTH CARE PROFESSIONALS, AND LAWYERS." *The Library Quarterly: Information,*

Community, Policy 66 (2): 161–93. http://www.jstor.org/stable/4309109.

LeMerle, Matthew. 2011. “How to Prepare for a Black Swan: Disrupter Analysis.” *Strategy+Business*.

Lilliefors, Hubert. 1967. “ON THE KOLMOGOROV-SMIRNOV TEST FOR NORMALITY WITH MEAN AND VARIANCE UNKNOWN.” *Journal of American Statistical Association* 62 (318): 399–402.

Lin, Yuri, Jean-Baptiste Michel, Erez Lieberman Aiden, Jon Orwant, Will Brockman, and Slav Petrov. 2012. “Syntactic Annotations for the Google Books Ngram Corpus.” In *Proceedings of the ACL 2012 System Demonstrations*, 169–74. http://www.aclweb.org/anthology/P12-3029.

Lloyd, Henry Martin. 2012. “Power , Resistance , and the Foucauldian Technologies.” *Philosophy Today* 56 (1): 26–38. doi:10.5840/philtoday201256122.

Maier, Norman R. F. 1972. “Effects of Training on Decision-Making.” *Psychological Reports* 30 (1): 159–64. doi:10.2466/pr0.1972.30.1.159.

Malone, Timothy G., and Reagan E. Schaupp. 2002. “The ‘Red Team’ Forging a Well-Conceived Contingency Plan.” *Aerospace Power Journal* 16 (2).

Mandelbrot, Benoit B, and Nassim Taleb. 2010. “Mild Vs Wild Randomness Focusing on Those Risks That Matter.” In *The Known, the Unknown, and the Unknowable in Financial Risk Management : Measurement and Theory Advancing Practice*, edited by Richard Herring, Neil Doherty, and Francis Diebold, 47–56. Princenton, NJ. doi:10.1017/CBO9781107415324.004.

Margolis, Howard. 2008. “A Note on Neglect Defaulting.” *Judgement and Decision Making* 3 (4): 355–63. doi:10.1016/j.nutres.2015.04.014.

Martínez León, Hilda C., Jennifer A. Farris, Geert Letens, and Alberto Hernandez. 2013. “An Analytical Management Framework for New Product Development Processes Featuring Uncertain Iterations.” *Journal of Engineering and Technology Management* 30 (1): 45–71. doi:10.1016/j.jengtecman.2012.11.004.

Massie, MJ. 2011. “Risk Acceptance Personality Paradigm: How We View What We Don’t Know We Don't Know.” *American Institute of Aeronautics and Astronautics*, 1–18. http://ntrs.nasa.gov/search.jsp?R=20110010000.

Masys, A. 2012. “Black Swans to Grey Swans: Revealing the Uncertainty.” *Disaster Prevention and Management* 21 (3): 320–35. doi:10.1108/09653561211234507.

Mateski, Mark. 2009. “Red Teaming : A Short Introduction,” no. June: 0–48. http://redteamjournal.com/resources/.

Matta, N, and R Ashkenas. 2003. “Why Good Projects Fail Anyway.” *Harvard Business Review* 81 (9): 109–16.

Mayseless, Ofra, and Arie W Kruglanski. 1987. “What Makes You so Sure? Effects of Epistemic Motivations on Judgmental Confidence.” *Organizational Behavior and Human Decision Processes* 39 (2): 162–83. doi:10.1016/0749-5978(87)90036-7.

McGlynn, Richard P., Dennis McGurk, Vicki Sprague Effland, Nancy L. Johll, and Deborah J. Harding. 2004. “Brainstorming and Task Performance in Groups Constrained by Evidence.” *Organizational Behavior and Human Decision*

Processes 93 (1): 75–87. doi:10.1016/j.obhdp.2003.09.003.

Menkhoff, Lukas, Maik Schmeling, and Ulrich Schmidt. 2013. "Overconfidence, Experience, and Professionalism: An Experimental Study." *Journal of Economic Behavior and Organization* 86. doi:10.1016/j.jebo.2012.12.022.

Merrick, Jason R W. 2008. "Getting the Right Mix of Experts." *Decision Analysis* 5 (1): 43–52. doi:10.1287/deca.1080.0108.

Merrill, M, L Drake, M Lacy, and J Pratt. 1996. "Reclaiming Instructional Design." *Educational Technology* 36 (5): 5–7.

Meyer, Mary, and Jane Booker. 1991. *Eliciting and Analyzing Expert Judgement: A Practical Guide*. San Diego, CA: Academic Press Limited.

Mitchell, Deborah J, J Edward Russo, and Nancy Pennington. 1989. "Back to the Future: Temporal Perspective in the Explanation of Events." *Journal of Behavioral Decision Making*. doi:10.1002/bdm.3960020103.

Mlodinow, Leonard. 2008. *The Drunkard's Walk: How Randomness Rules Our Lives*. New York: Pantheon Books.

Mullen, Brian, Craig Johnson, and Eduardo Salas. 1991. "Productivity Loss in Brainstorming Groups: A Meta-Analytic Integration." *Basic and Applied Social Psychology* 12 (1): 3–23. doi:10.1207/s15324834basp1201_1.

NASA. 2003. "Columbia Accident Investigation Board Report Volume I." Washington, DC.

NDU. 2016. "Eisenhower School."

Nemeth, Charlan, Keith Brown, and John Rogers. 2001. “Devil’s Advocate versus Authentic Dissent: Stimulating Quantity and Quality.” *European Journal of Social Psychology* 31 (6): 707–20. doi:10.1002/ejsp.58.

Nemeth, Charlan, and J Wachtler. 1983. “Creative Problem Solving as a Result of Majority vs. Minority Influence.” *European Journal of Social Psychology* 13 (1): 45–55.

Noble, David, and Michael Letsky. 2001. “Cognitive-Based Metrics to Evaluate Collaboration Effectiveness.” *Interactions*.

Olson, Bill. 2012. “Problem Management : A System Engineering Management Framework.” The George Washington University.

Osita, Ifediora, Idoko Onyebuchi, and Justina Nzekwe. 2014. “Organization’s Stability and Productivity: The Role of SWOT Analysis an Acronym for Strength, Weakness, Opportunities and Threat.” *International Journal of Innovative and Applied Research* 2 ((9): 23- 32): 1–12.

Ott, R, and M Longnecker. 2010. *An Introduction to Statistical Mehtods and Data Analysis*. Sixth. Belmont, CA: Brooks/Cole Cengage learning.

OUSD(AT&L). 2006. “Risk Management Guide for DoD Acquisition.” Washington, DC.

Palisade. 2015. *User’s Guide StatTools Version 7*. Vol. 14850. Ithaca, New York: Palisade Corporation. http://www.palisade.com.

Patil, Rahul, Katie Grantham, and David Steele. 2012. “Business Risk in Early Design: A Business Risk Assessment Approach.” *Engineering Managment Journal* 24 (1): 35–

46.

Paulus, Paul B., Runa M. Korde, Jubilee J. Dickson, Abraham Carmeli, and Ravit Cohen-Meitar. 2015. “Asynchronous Brainstorming in an Industrial Setting: Exploratory Studies.” *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57 (6): 1076–94. doi:10.1177/0018720815570374.

Perminova, Olga, Magnus Gustafsson, and Kim Wikström. 2008. “Defining Uncertainty in Projects – a New Perspective.” *International Journal of Project Management* 26 (1): 73–79. doi:10.1016/j.ijproman.2007.08.005.

PMI. 2009. *Practice Standard for Project Risk Management*. Newtown Square, PA: Project Management Institute. www.pmi.org.

Porath, Christine L., and Amir Erez. 2009. “Overlooked but Not Untouched: How Rudeness Reduces Onlookers’ Performance on Routine and Creative Tasks.” *Organizational Behavior and Human Decision Processes* 109 (1). Elsevier Inc.: 29–44. doi:10.1016/j.obhdp.2009.01.003.

Posner, Kenneth. 2010. “Chapter 6 Mitigating Information Assymetry.” In *Stalking the Black Swan*, 137–54. New York: Columbia University Press. URL: http://www.jstor.org/stable/10.7312/posn15048.10.

Prestwich, Andrew, Rebecca Lawton, and Mark Conner. 2003. “The Use of Implementation Intentions and the Decision Balance Sheet in Promoting Exercise and Behaviour.” *Psychology & Health* 18 (6): 707–21.

Prochaska, James O, Wayne F Velicer, Joseph S Rossi, Michael G Goldstein, Bess M

Marcus, William Rakowski, Christine Fiore, et al. 1994. “Stages of Change and Decisional Balance for 12 Problem Behaviors.” *Health Psychology* 13 (April 2016): 39–46. doi:10.1037/0278-6133.13.1.39.

Puccio, Gerard J., John F. Cabra, J. Michael Fox, and Helene Cahen. 2010. “Creativity on Demand: Historical Approaches and Future Trends.” *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 24 (02): 153. doi:10.1017/S0890060410000028.

Ralston, Bill, and Ian Wilson. 2006. *The Scenario Planning Handbook: Developing Strategies in Uncertain Times*. First. Crawfordsville, Indiana: Thomson.

Reason, James. 1997. *Managing the Risks of Organizational Accidents*. New York: Asgate.

Reeves, John D. 2012. “The Impact of Early Design Phase Risk Identification Biases on Space System Project Performance.” *Systems Conference (SysCon), 2012 IEEE International*, 1–8. doi:10.1109/SysCon.2012.6189498.

Reeves, John D., Tim Eveleigh, Thomas H. Holzer, and Shahryar Sarkani. 2013. “Risk Identification Biases and Their Impact to Space System Development Project Performance.” *Engineering Management Journal* 25 (2): 3–12.

Reinig, Bruce a., and Robert O. Briggs. 2013. “Putting Quality First in Ideation Research.” *Group Decision and Negotiation* 22 (5): 943–73. doi:10.1007/s10726-012-9338-y.

Reinig, Bruce, Robert Briggs, and Jay Nunamaker. 2007. “On the Measurement of

Ideation Quality." *Journal of Management Information Systems* 23 (4): 143–61. doi:10.2753/MIS0742-1222230407.

Robson, Colin. 2011. *Real World Research: A Resource for Users of Social Research Methods in Applied Settings*. Third. Cornwall, Uk: Wiley.

Ross, John. 1999. *The Polar Bear Strategy: Reflections on Risk in Modern Life*. 1st ed. Reading, Massachusetts: Perseus Books.

Rouwette, Etienne, Ingrid Bastings, and Hans Blokker. 2011. "A Comparison of Group Model Building and Strategic Options Development and Analysis." *Group Decision and Negotiation* 20 (6): 781–803. doi:10.1007/s10726-010-9207-5.

Ruxton, Graeme D., M Neuhauser, and Markus Neuhäuser. 2010. "When Should We Use One-Tailed Hypothesis Testing?" *Methods in Ecology and Evolution* 1 (2): 114–17. doi:10.1111/j.2041-210X.2010.00014.x.

Sakia, R M. 1992. "The Box-Cox Transformation Technique : A Review." *The Statistician* 41: 169–78.

Sarkar, Debasis, and Goutam Dutta. 2012. "A Framework for Project Risk Management for the Underground Corridor Construction of Metro Rail." *International Journal of Construction Project Management* 4 (1): 21–38.

Schwenk, Charles R. 1988. "Effects of Devil's Advocacy on Escalating Commitment." *Human Relations* 41 (10): 769–83. doi:10.1177/001872678804101005.

Scott, David. 2016. "Box-Cox Transformations." Accessed May 6. http://onlinestatbook.com/2/transformations/box-cox.html.

Shenkir, Wiliam G., and Paul L. Walker. 2007. “Enterprise Risk Management: Tools and Techniques for Effective Implementation.” *Institute of Management Accountants*, 1–31.

Sheth, Mitez. 2012. “THE AMAZING CONCEPT OF SIX THINKING HATS.” *INTERNATIONAL JOURNAL OF MANAGEMENT RESEARCH AND REVIEW* 2 (3): 316–34. http://www.ijmrr.com/admin/upload_data/journal_NaserMohammadoghli___8sep13mrr.pdf.

Shirouyehzad, Hadi, Reza Dabestani, and Mostafa Badakhshian. 2011. “The FMEA Approach to Identification of Critical Failure Factors in ERP Implementation.” *International Business Research* 4 (3): 254–63. doi:10.5539/ibr.v4n3p254.

Sinfield, Joseph V., Tim Gustafson, and Brian Hindo. 2013. “The Discipline of Creativity.” *MIT Sloan Management Review*, no. 55215. http://sloanreview.mit.edu/article/the-discipline-of-creativity/.

Slovic, Paul, Melissa L Finucane, Ellen Peters, and Donald G MacGregor. 2004. “Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality.” *Risk Analysis : An Official Publication of the Society for Risk Analysis* 24 (2): 311–22. doi:10.1111/j.0272-4332.2004.00433.x.

Smith, H.W. 1991. *Strategies of Social Research*. 3rd ed. New York: Holt, Reinhart and Winston.

Smithson, Michael. 1989. *Ignorance and Uncertainty: Emerging Paradigms (Cognitive Science)*. 1st ed. Springer - Verlag.

Stake, Robert. 2010. *Qualitative Research: Studying How Things Work*. New York: Guilford Press.

Stamatelatos, Michael, and Jose Caraballo. 2002. “Fault Tree Handbook with Aerospace Applications.” Washington, DC.

Starr, Chauncey, and Chris Whipple. 1980. “Risks of Risk Decisions.” *Science* 208 (4448): 1114–19.

Strausz, Erzsebet. 2011. “Foucault’s Critique: A Topology of Thought.” *Law and Critique* 22 (2): 119–33. doi:10.1007/s10978-011-9082-5.

Susskind, Lawrence. 2012. “Responding to the Risks Posed by Climate Change.” *Town Planning Review* 81 (3): 217–35. doi:10.3828/tpr.2010.5.

Taleb, Nassim. 2003. “The Black Swan: Why Don’t We Learn That We Don't Learn?” In *Highland Forum 23*, 1–44. Las Vegas, NV. www.thehighlandsforum.org.

———. 2009. “Ten Principles for a Black Swan-Proof World.” *Financial Times*, 1–2.

———. 2010. *The Black Swan: The Impact of the Highly Improbable*. Second. New York: Random House. www.atrandom.com.

Taleb, Nassim, Daniel Goldstein, and Mark Spitznagel. 2009. “The Six Mistakes Executives Make in Risk Management.” *Harvard Business Review*, no. October: 78–81. doi:Article.

Tarrant, M. 2010. “The Organisation: Risk, Resilience and Governance.” *The Australian Journal of Emergency Management* 25 (2): 13–17.

Textron. 2012. “Textron Systems Overview.” http://www.textron.com.

Tjosvold, D, and D Deemer. 1980. “Effects of Controversy Within a Cooperative or Competitive Context on Organizational Decision-Making.” *Journal of Applied Psychology* 65: 590–95.

Tseng, Yen Hua-Hui. 2008. “Motivation of Participation in Inservice Training Based on Problem Solving: A Modified Delphi Study.” University of Phoenix.

Tumer, Irem, Robert Stone, and Rory Roberts. 2003. “Decomposition-Base Failure Mode Identification Method.”

Turner, M, and A Pratkanis. 1997. “Mitigating Groupthink by Stimulating Constructive Conflict.” In *Using Conflict in Organizations*, edited by C DeDreu and E Van de Vliert, 53–71. London, UK: Sage.

USCG. 1998. “Effective Mission Analysis.” In *Team Coordination Training Student Guide*, 1–34. Washington, DC: USCG. http://www.uscg.mil/auxiliary/training/tct/chap3.pdf.

———. 2016. “United States Coast Guard Acquisition Programs.” http://www.uscg.mil/acquisition/programs/acquisitionprograms.asp.

Valukas, Anton. 2014. “Report to Board of Directors of General Motors Company Regarding Ignition Switch Recalls (Redacted).” http://www.projectauditors.com/Papers/GM_investigation_report_Valukas_ignition_switch.pdf.

Varvoglis, M P. 2003. “Scientists, Shamans, and Sages: Gazing through Six Hats.” *Journal of Parapsychology* 67 (1): 3–16.

Vaughan, Diane. 1996. *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. 1st ed. Chicago: University of Chicago Press.

Veinott, Beth, Gary A. Klein, and Sterling Wiggins. 2010. "Evaluating the Effectiveness of the PreMortem Technique on Plan Confidence." *Proceedings of the 7th International Conference on Information Systems for Crisis Response and Management*, no. May: 1–9.

Velicer, W F, C C DiClemente, J O Prochaska, and N Brandenburg. 1985. "Decisional Balance Measure for Assessing and Predicting Smoking Status." *Journal of Personality and Social Psychology* 48 (5): 1279–89. doi:10.1037/0022-3514.48.5.1279.

Vesely, Bill. 2002. "Fault Tree Analysis (FTA): Concepts and Applications." Washington, DC: NASA.

Wagner, Cynthia G. 2005. "Futurists Gather to Develop Foresight, Innovation, and Strategy." *Futurist* 39 (6).

Walker, Paul, William Shenkir, and Thomas Barton. 2014. "A Risk Challenge Culture." London. http://www.ima.org.

Walters, Helen. 2008. "Apple's Design Process." *Bloomberg Business: The Tech Beat*. http://www.businessweek.com/the_thread/techbeat/archives/2008/03/apples_design_p.html.

Wessa, P. 2015. "Box-Cox Normality Plot." Free Statistic (Office for Research Development and Education). http://www.wessa.net/rwasp_boxcoxnorm.wasp/.

Wilson, Chauncey. 2013. “Chapter 1 - Brainstorming.” In *Brainstorming and Beyond*, 1–41. doi:http://dx.doi.org/10.1016/B978-0-12-407157-5.00001-4.

Wright, G, and P Goodwin. 2009. “Decision Making and Planning under Low Levels of Predictability: Enhancing the Scenario Method.” *International Journal of Forecasting* 25 (4): 813–25. doi:10.1063/1.2756072.

Yule, Peter, and Derek Woolner. 2008. *The Collins Class Submarine Story: Steel, Spies and Spin*. Cambridge: Cambridge University Press.

Zhang, Honglianh (Peking University). 2011. “Two Schools of Risk Analysis : A Review of Past Research on Project Risk.” *Project Management Journal*, no. July: 5–18. doi:10.1002/pmj.

Zhao, Zhenguo, and Jialin Hou. 2010. “The Study on Influencing Factors of Team Brainstorming Effectiveness.” *International Journal of Business and Management* 5 (1): 181–84. http://media.proquest.com.ezaccess.libraries.psu.edu/media/pq/classic/doc/2223171301/fmt/pi/rep/NONE?hl=brainstorming&cit:auth=Zhao,+Zhenguo;Hou,+Jialin&cit:title=The+Study+on+Influencing+Factors+of+Team+Brainstorming+Effectiveness&cit:pub=International+J.

Appendix A - Expert Panel Member Biographies

(Excludes five panel members that did not consent to include their biographies)

Patrick K. Barker
Professor
Defense Systems Management College
Defense Acquisition University (DAU)

Patrick Barker joined the Defense Systems Management College in February 2012.

He is a certified and experienced DAU Executive Coach, course manager for ACQ-405 (Executive Refresher Course), Learning Asset Manager for all DSMC instructional material for Earned Value Management (EVM), Scheduling, Planning and Executive Decision-making classes. Mr. Barker is also a key asset in DAU Mission Assist (consulting) engagements with other agencies and program management offices.

Previously he worked over 14 years as a contractor providing support and consulting to Defense, non-Defense Federal and commercial organizations in areas such as EVM, integrated baseline reviews (IBR), scheduling, risk management, test & evaluation and systems engineering. Clients included the US Navy Cobra Judy Replacement ACAT 1D program, USD (AT&L)/ARA, the Federal Aviation Administration's Surveillance and Broadcast System Program, programs in Department of Homeland Security, Department of Justice, Small Business Administration and multiple small and mid-sized commercial businesses. Mr. Barker has served as program control/integration lead on a major program and, as a consultant, often led small teams that prepared clients for IBRs and/or set up EVM system implementations "from scratch," often in the middle of program execution. He also led EVM assessments of numerous programs to include ACAT 1's.

Prior to becoming a contractor, he spent a 20-year career in the USAF and USAF Reserves. Assignments included serving as instructor electronic warfare officer in B-52s, assistant professor of history and airmanship instructor at the USAF Academy, and functional manager in technology transfer for the Defense Intelligence Agency.

A native of Pennsylvania, Mr. Barker earned a Bachelor of Science in Mechanical Engineering (1986) and a Masters of Arts in History (1994), both from Lehigh University. He continues to serve as adjunct professor at American University where he established a senior program manager certification program to meet Federal Acquisition Institute requirements. He is certified DAWIA Level III in Program Management.

Mr. Michael S. Bohn
Professor of Acquisition Management
Defense Systems Management College
School of Program Managers
Defense Acquisition University (DAU)

Mike Bohn joined DAU in November 2009 as a Professor of Acquisition Management.

Prior to joining DAU, Mr. Bohn was the Chief of Staff and Director of the Technology Division at the Marine Corps Warfighting Lab. In this capacity, he managed the day to day activities of the Warfighting Lab and was responsible for discovering and managing emerging technologies to support future Warfighting experiments and current Warfighter needs.

From June 2004 to June 2008, Mr. Bohn served as the Director, Marine Corps Operational Test and Evaluation Activity, Quantico, VA. In this capacity he served as the Marine Corps Operational Test Agency Commander, responsible for all Marine Corps operational test activities and policies.

Prior to attending Top Level School with the State Department, Mr. Bohn was the Program Manager, Assault Amphibian Vehicles, Marine Corps Systems Command from February 2001 until August 2003.

He was the Operational Test Project Officer for the Advanced Amphibious Assault Vehicle, Marine Corps Operational Test and Evaluation Activity from May 2000 to February2001, in preparation for the program's Milestone B decision.

From June 1998 to May 2000, Mr. Bohn was the Director, Amphibious Vehicle Test Branch, Camp Pendleton, CA. Prior to June of 1998, Mr. Bohn served in a variety of Operating Forces and Supporting Establishment billets as a Marine Corps Assault Amphibian Officer.

Mr. Bohn is a 1980 graduate of the University of Colorado, with a BA degree in History and 1992 graduate of the University of Nebraska, with a MA in History. He is a 2001 graduate of DSMC's Advanced Program Management Course and is DAWIA Level III certified in Program Management.

Mr. Brian P. Brodfuehrer
Professor of Program Management
Defense Systems Management College (DSMC)
Defense Acquisition University (DAU)

Mr. Brodfuehrer joined DSMC faculty in July 2007 after spending five years working for the private sector.

While in the private sector, he provided satellite and information technology related systems engineering and program management support to the National Geospatial Information Agency and to the National Reconnaissance Office. Prior to working for industry, Mr. Brodfuehrer spent twenty two years in the Air Force acquisition community where he had assignments at Ft. Meade, MD developing microelectronics for NSA, at Los Angeles Air Force Station, CA in the MILSTAR satellite Joint Program Office, at Wright Patterson AFB, OH in the B-2 Program Office, and at the HQ Defense Contract Management Agency, Ft. Belvoir, VA. He spent his last assignment in the Air Force at DAU where he taught Systems Engineering and was part of the team that started the PMT-401 course.

Mr. Brodfuehrer is a PMC graduate (88-2) and is Level III certified in Program Management and Systems Engineering. He holds a BA in International Relations and a BS in Electrical Engineering from the University of Notre Dame, an MS in Electrical Engineering from the University of Maryland, and an MS in National Security Resource Strategy from the National Defense University, Industrial College of the Armed forces, Class of 1995. In August 2012 he was awarded the Master of Arts in Theology from the Notre Dame Graduate School, Alexandria, VA. Additionally, he is a certified DAU Executive Coach; he is qualified to conduct and provide feedback on the Myers Briggs Type Indicator (MBTI) assessment and is a Certified Polarity Transformation and Assessment Consultant.

Mr. James O. Carter
Professor of Acquisition Management
Defense Systems Management College
School of Program Managers
Defense Acquisition University (DAU)

Mr. Carter is a Professor of Systems Acquisition Management.

For twenty years, Mr. Carter served in a variety of operational billets at sea and acquisition billets ashore. At the Naval Sea Systems Command, he was the ASW system engineer for the AEGIS Cruiser and Destroyer Shipbuilding Program. At the Naval Air Systems Command, he was the combat system engineer for Lamps Mk III integration in the USS Ticonderoga. At Commander Operational Test & Evaluation Force, he was the operational test director for the Oliver Hazard Perry class of guided missile frigates. At the Naval Sea Systems Command, he was the combat system engineer for the AEGIS Cruiser Shipbuilding Program. Upon retirement, Mr. Carter founded C.E.I., a system engineering and program management consulting firm. He directed the business activities of C.E.I. for 17 years.

He is a graduate of the U.S. Naval Academy, the DAU Program Management Course, and the DAU Program Manager's Course. He holds two Master's Degrees from the Naval Postgraduate School: one in Operations Research/Systems Analysis and one in Weapons Systems Acquisition Management. He is Level III certified in Program Management and Test & Evaluation Engineering.

Christopher J. D'Ascenzo D.BA

Intermittent Professor of Program Management
Defense Systems Management College
Defense Acquisition University (DAU)

Dr D'Ascenzo joined the Defense Systems Management College faculty in 2014 as a Professor of Program Management teaching, developing curriculum, performing mission assistance and research supporting program managers, program executive officers or their deputies in the DOD.
Dr D'Ascenzo is an industry operations executive with a full career serving the defense and national security community. He has led sensitive communications and automated information and data systems development and lifecycle services programs including highly successful business expansions. From 2007 to 2013, Chris was Vice President and General Manager of the $300M/year Defense Intelligence Solutions business at BAE Systems Intelligence and Security providing enterprise IT systems and services to DIA, NSA, DLA, DTS, DISA, Air Force ISR, Army G2/G6, Army CECOM and the Combatant Commands. He expanded the cross domain cyber product development business and led the Army C4ISR full lifecycle software engineering sustainment work at Army CECOM.

Before joining BAE, Dr D'Ascenzo was employed by Lockheed Martin for 27 years in the large information and intelligence systems integration business. From 2001 to 2007 as a senior director he led successful corporate strategic C4I growth campaigns in Homeland Security, Law Enforcement Communications and the Navy's Next Generation Enterprise Network. Prior to that Mr D'Ascenzo was senior program director at Lockheed Martin's Information Systems and Global Solutions Sector where he managed several P&L portfolios including NGA/Army/Air Force image dissemination programs, IMINT and SIGINT collection and processing systems, NSA LMD/KP cryptographic software development, classified ground stations, Justice and Homeland Security communications systems. His early career contributions included electronic design engineering and project leadership roles on classified space NRO programs, Air Force DSCS satellite ground terminals and the MK21A missile radar.

Chris holds BS and MS degrees in Electrical Engineering from Drexel University, his MBA at Villanova and holds a Doctorate in Business Administration with his dissertation in early stage venture risk management. He completed BAE Systems' Gateway Global Leadership executive program at Wharton and Oxford and has a patent in 911 emergency IT. He holds PMI's PgMP and PfMP professional Program Management certifications, NCMA's CPCM Contract Management designation and ITIL and COBIT5 Information Technology credentials and is DAWIA certified in program management and IT.

Dr. Bobbie G. DeLeon
Professor of Acquisition Management
Defense Systems Management College
Defense Acquisition University (DAU)

Dr. DeLeon joined the DAU in December 2006 as a Professor of Acquisition Management. She has over 15 years of program management experience including logistics, acquisition and IT programs. Her primary focus is as an instructor and faculty advisor in the PMT-401 Program Manager's Course (PMC). Dr. DeLeon has extensive experience in program and project management, team building and executive development.

Before joining the DSMC/DAU faculty Dr. DeLeon served as Program Manager for the Defense Research and Engineering Program Analysis and Information Technology at the Defense Technical Information Center (DTIC). Prior to that she held several assignments at Naval Sea Systems Command (NAVSEA) including Acquisition Program Manager for the Test, Measurement and Diagnostic Equipment (TMDE) Program, Logistics Manager for the General Purpose Electronic Test Equipment (GPETE) and Calibration Standards (CAL STDS) Programs and Business Financial Manager for the Other Procurement, Navy (OPN) Outfitting Spares Program.

Dr. DeLeon holds both a Bachelor of Science Degree in Economics and a Bachelor of Science Degree in Business Administration from Tuskegee University, a Master of Science Degree in Technology Management from George Mason University, Doctor of Education Degree from the George Washington University and holds DoD Level III certification in Program Management. She is also a graduate of the Department of Navy Acquisition Logistics Intern Program, the General Services Administration (GSA) Chief Information Officer's (CIO) University and the PMC.

Mrs. Joni Forman
Associate Academic Affairs Dean
Defense Systems Management College
Defense Acquisition University (DAU)

Mrs. Joni Forman joined DAU in October 2001 as a Professor of Acquisition Management and Deputy Director for Curricula Development Support Center (CDSC). She transitioned to the DSMC School of Program Managers in January 2002 as the Director, Executive Curriculum Development leading the development of three executive level courses for the AT&L post level three certified workforce. In January 2007, she was selected as the Associate Academic Affairs Dean for DSMC responsible for management and delivery of the DSMC executive level curriculum. In addition to the Associate Academic Dean (AAD) responsibilities, Mrs. Forman was the Course Manager for the Defense Acquisition Executive Overview Workshop (DAEOW) and Course Director for the Acquisition Leadership and Management Curriculum for 2 years. In February 2013, Mrs. Forman relinquished her duties as the AAD and was assigned as the Program Learning Director for the Leadership and Management Curriculum.

Prior to joining DAU, Mrs. Forman was a senior program analyst with the Office of the Under Secretary of Defense (Acquisition and Technology) Acquisition Resources and Analysis, Acquisition Management. She served on the OSD staff from February 1988 to October 2001. During her employment with OSD(AT&L), Mrs. Forman was selected for an AT&L career development assignment and served from October 1997 to January 1999, as the Deputy Project Manager for PM Night Vision/Reconnaissance Surveillance and Target Acquisition Program located at Ft. Belvoir, VA.

Other government assignments included working for the Assistant Secretary of the Army (Research, Development & Acquisition), Weapons and Tracked Directorate as a Department of Army Systems Coordinator (DASC) from April 1985 to February 1988; program analyst for the Army Material Command and the Army Adjutant General from 1980 to 1985; and administrative staff positions for the U.S. Information Agency from 1974 to 1980.

Advanced educational degrees include a Masters in Public Policy from the American University Key Executive Program, May 1996, and a Bachelor of Science from the University of Maryland University College in Business and Technology Management.

Mr. Richard D. Hansen Jr.
Executive Coach and Professor of Program Management
Defense Systems Management College
Defense Acquisition University

Richard Hansen joined DAU in September 2008 as a Professor of Program Management where he contributes to the executive programs for the Department of Defense senior acquisition workforce. Specifically, Dick helped launch an award winning initiative to create and sustain a coaching culture among key acquisition leaders to improve their programmatic performance and organizational and leadership development. Using a refined, proven coaching model/process, Dick continues to train and qualify a cadre of 49 experienced acquisition practitioners, who have coached over 225 key leader clients to design and implement an extraordinary future and stretch beyond the predictable status-quo.

Prior to joining DAU, Dick spent twenty six years in the US Army in both the warfighting and acquisition fields. He held various command and staff positions in self- propelled, towed, and airborne artillery units and fire support elements with III Corps Artillery, the 2nd Infantry Division, and the 82d Airborne Division.

His research, development, and acquisition experience includes various assignments as a product, project, or program manager in electric guns, armored vehicles, precision guided munitions, and critical weapons, clothing, and equipment for aviation crewmen, dismounted Soldiers, and vehicle crewmen. Dick served in the requirements and resourcing fields as a member of the Army's Objective Force Task Force, the Army's action agency to coordinate, integrate, and synchronizing the Army's transformation efforts to the future force. He also focused on joint integrated air and missile defense architectures and interoperability challenges while assigned to the Joint Single Integrated Air Picture (SIAP) System Engineering Organization.

He has a Bachelor of Science degree in Business Administration from Trinity University in San Antonio, Texas, a Masters of Arts in Economics from the University of Oklahoma, a Masters in Business Administration from the University of Texas at Austin, and a Masters of Strategic Studies from the US Army War College. He is Level III certified in Program Management, a Six Sigma Green Belt, a Speed of Trust and Strength Deployment Inventory facilitator, and holds the International Coaching Federation credential of Associate Certified Coach.

Mr. Steve Henry
Professor of Acquisition Management
Defense Systems Management College
Defense Acquisition University (DAU)

Stephen A Henry joined the faulty as a Professor of Acquisition Management in July 2015. He is supporting the Missile Defense Agency Key Leadership Development Program and mission assist activities.

Steve was previously the director of Engineering for the Northrop Grumman Information Systems sector's Command and Control Division. Mr. Henry has more than 43 years of experience in designing, implementing, testing and managing large programs. He served 28 n the Air Force retiring at the rank of Colonel in 2000. He is a graduate of the USAF Test Pilot School and flew the first flight of the B-1B in 1984. Key acquisition and engineering assignments include A-10 Program Manager, B-1 Program Element Monitor, Executive Officer SAF/AQQ, Deputy Commandant of USAF TPS, Joint Stars Director of Engineering, and System Program Director for Global Air Traffic Operations and Mobility Command and Control (GATO-MC2) program office. As Director of Engineering, Mr. Henry led a group of senior engineers who helped apply effective systems engineering to start-up programs right. He also served as the Northrop Grumman Information Systems, Systems Engineering Center of Excellence lead. He provided significant leadership and capture support to major programs such as E-10 Battle Management Command and Control, CANES, Tanker, JLTV, AOC, and Joint STARS RECAP. In these roles he significantly improved the quality of system engineering personnel development and engineering practices. He is the past Chairman of the NDIA Systems Engineering Division and has supported numerous NDIA working groups to support DoD. In 2015 he was named the winner of the National Defense Industrial Association's Lt Gen Thomas R. Ferguson, Jr. Systems Engineering Excellence Award.

Mr. Henry is the recipient of the Air Force Material Command Chief Engineer of the Year, Air Force Association, Executive Division, Program Manager of the Year, and Armed Force Communications and Electronics Association's International Award for Excellence in Information Technology awards. Under his leadership, the GATO-MC2 program office was named best major program office at Electronic Systems Center for three consecutive years. Mr. Henry earned his Bachelor of Science in Aerospace Engineering from Texas A&M University and Masters of Science in Systems Management for the Air Force Institute of Technology.

Mr. Sean Herr
Professor of Acquisition Management
Defense Systems Management College
School of Program Managers
Defense Acquisition University (DAU)

Mr. Sean Herr joined the faculty at the Defense Acquisition University in July 2011 as a Professor of Acquisition Management in the DSMC School of Program Managers.

Mr. Herr retired from active duty with more than 24 years of service in the Air Force. Prior to his arrival at DAU, Mr. Herr was the Associate Dean of Faculty and Academic Programs at the Industrial College of the Armed Forces and an Associate Professor of Acquisition. From June 2003 until June 2005, he served as the E-10A and Multi-Platform Radar Technology Insertion Program (MP-RTIP) Program Element Monitor for the Office of the Assistant Secretary of the Air Force for Acquisition.

From 2000 to 2003, he was as an Air Officer Commanding at the US Air Force Academy and then a Deputy Group Commander. From 1997 to 1999, he was the Executive Officer in the Air Force Program Executive Office for Fighter and Bomber Programs. From 1993-1996, he served in a number of positions in the Defense Dissemination Program Office at Los Angeles Air Force Base. When the National Imagery and Mapping Agency was established in 1996, he relocated to Washington DC and was selected as Program Manager, Defense Dissemination System IV. From 1987-1993, he worked on a variety of program management positions in the Air Force Munitions Systems Division to include Deputy Program Manager for the Operational Concept Demonstration, a precursor to the JDAM program.

Mr. Herr is a graduate of the Industrial College of the Armed Forces, June 2006, and Chapman University, December 1996. He has a Master of Science in National Resource Strategy and a Masters in Business Administration. His undergraduate degree is from the United States Air Force Academy in Management. Mr. Herr is a 1999 graduate of DSMC's Advanced Program Management Course and PMT 401 (2009-2). He is level III certified in Program Management.

Mr. Mike Holbert
Defense Systems Management College
School of Program Managers
Defense Acquisition University (DAU)

Mr. Mike Holbert joined DAU in August 2009 as a Professor of Program Management. He presently serves as Center Director, Executive Programs, faculty member and certified Executive Coach. Prior to joining DAU, Mr. Holbert was the Director of the 84th Combat Sustainment Wing, Ogden Air Logistics Center, Hill Air Force Base, Utah.

Mr. Holbert retired from the Air force after over 26 years of service. He has had the opportunity to lead and work in various positions in the Air Force, including duties as a program manager, research and development engineer, logistics manager and aircraft maintenance officer. He served as an engineer/test engineer or program manager with various weapon systems including medium caliber ammunition/gun systems, air-to-ground precision munitions, the B-2 bomber, and GPS ground and space segments. He has aircraft maintenance experience on B-52s and KC-135s. He provided systems engineering and developmental test oversight for multiple missile defense programs. His logistic management experience covers multiple munitions, space and C3I systems, and aircraft commodities. His tenure in the Air Force has lead to assignments in the Office of the Secretary of Defense (Acquisition, Technology, and Logistics), Ballistic Missile Defense Office, Office of the Secretary of the Air Force, Air Force Material Command, Air Force Systems Command, Air Force Space Command, Strategic Air Command, and Air Education and Training Command (AF Center for Systems Engineering).

Mr. Holbert is a graduate of the Navy War College, June 2001 and Embry-Riddle Aerospace University, Aug 1992. He has a Master of Arts in National and Strategic Studies and a Master of Science in Aerospace/Aviation Management. His undergraduate degree is from North Dakota State University in Mechanical Engineering. Mr. Holbert is a 1998 graduate of DSMC's Advanced Program Management Course and PMT 401 (2007-1). He is level III certified in Program Management, SPRDE, Life Cycle Logistics, and Test and Evaluation.

Dr. D. Mark Husband
Professor of Acquisition Management
Defense Systems Management College
Defense Acquisition University (DAU)

Dr. Mark Husband first joined the DAU faculty in 2010 after 26 years in the US Air Force. While on active duty, he worked for over fifteen years in a variety of acquisition disciplines including cost estimating; systems program research, development and engineering (SPRDE); program management; and scientific and technical management. From May, 2013 to Jan, 2016, Dr. Husband served as Senior Advisor for Root Cause Analyses in the office of Performance Assessments and Root Cause Analyses (PARCA), where he conducted root cause analyses of critical Nunn-McCurdy breach programs and led a Better Buying Power Initiative on Eliminating Requirements Imposed on Industry. He re-joined DAU in Jan, 2016. Dr. Husband is Level III certified in SPRDE, Business-Cost Estimating and Business-Financial Management, Level II certified in Program Management, and is a certified DAU Executive Coach.

Dr. Husband began his career as a research engineer, specializing in solid propellant development and processing at Edwards AFB, CA and Karlsruhe, Germany and as a polymer processing engineer at the Materials Laboratory, Wright-Patterson AFB, OH. He then served as a Joint Operations Analyst at the U.S. European Command, where he conducted combat and logistics modeling supporting war plans development. During Operation ALLIED FORCE, he was Chief of the Contingency Plans Branch of the Kosovo Plans Group in the Operations Division and later served as Chief of the Operations Plans Group that developed two execution level contingency plans to respond to potential Serbian aggression in the Balkans. He served as the Military Assistant to the Commander, Allied Air Forces Southern Europe in Naples, Italy, and as squadron commander of the 426th Air Base Squadron, Stavanger, Norway.

From 2004 to 2007, Dr. Husband was an Operations Research Systems Analyst in the Office of Program Analysis and Evaluation (PA&E), Office of the Secretary of Defense. As a member of the Cost Analysis Improvement Group (CAIG), he conducted Independent Cost Estimates on Major Defense Acquisition Programs, including the USAF Combat Search and Rescue Helicopter, the USMC Expeditionary Fighting Vehicle, the USA Armed Reconnaissance Helicopter, UH-60M Upgrade, Stryker, and Global Combat Support System (a Major Automated Information System). Dr. Husband then served as Dean of the Air Command and Staff College resident program at Maxwell AFB, Alabama, followed by a second Pentagon tour as Deputy Director of Resource Analyses in HQ USAF/A9 (Studies & Analyses, Assessments and Lessons Learned).

Dr. Husband has Bachelor and Masters of Science degrees from the University of South Carolina and a Doctor of Engineering from the University of Karlsruhe, all in chemical engineering.

Mr. John Krieger
Professor of Contract Management
Defense Systems Management College
School of Program Managers
Defense Acquisition University (DAU)

Mr. Krieger is a Professor of Contract Management in the Defense Systems Management College's School of Program Managers, consultant in the Executive Department for PMT 402 Executive Program Manager's Course, ACQ 403 Defense Acquisition Executive Overview Workshop, ACQ 404 Systems Acquisition Management Course for General/Flag Officers, and ACQ 405 Executive Refresher Course. Previously, as the Director of the Contracting Center, he managed development and maintenance of DAU Contracting Courses.

Immediately prior to coming to DAU, he was the Assistant Commander (Contracts) for Marine Corps Systems Command, where he oversaw all aspects of the Command's contracting program, responsible for all "Green" systems acquisition for the Marine Corps.

He began his acquisition career with Defense Contract Administration Services as a Cost and Price Analyst. He has been a Contract Administrator, Contract Specialist and Procurement Analyst. He has worked for in both the Department of Defense and civilian agencies. As a contracting officer, he awarded major systems contracts for the Titan IV Complementary Expendable Launch Vehicle [Acquisition Streamlining Program] and, as part of the Space Launch Recovery Program, the Delta II Medium Launch Vehicle [Defense Enterprise Program].

EDUCATION AND TRAINING

1972 Bachelor of Arts degree in Speech Communication, California State University, Long Beach.
1973 Master of Arts degree in Communication, California State University, Long Beach.
1992 Master of Science degree in Contract and Acquisition Management, FIT.
2001 Master of Science Degree in National Resource Strategy, Industrial College of the Armed Forces.

Completed over 6,300 hours in Department of Defense and other courses in acquisition, contracting, logistics, and leadership. Courses included the Defense Acquisition University Senior Acquisition Course at the Industrial College of the Armed Forces and the Defense Systems Management College Program Manager's Course.

Dr. Glenn F. Lamartin
Professor of Acquisition Management
Defense Systems Management College

Dr. Lamartin joined the faculty of DSMC in October 2014 as a professor of acquisition management, working in the classroom and on curricula development, and assisting practitioners in the acquisition community. He is also an adjunct professor at Georgetown University's McCourt School of Public Policy where he teaches courses in national security and policy making.

Dr. Lamartin served a full career as a DoD acquisition professional, with success at the project, systems command, agency, and OSD levels. He started as a flight test engineer for the Navy and retired as a member of the Senior Executive Service on the OSD staff. Over the course of his career, he engaged in all phases of the acquisition lifecycle. He was a contributor to major reviews (e.g., QDRs); executive secretary for Defense Science Board studies; and the DoD focal point for GAO audits. As one of OUSD(AT&L)'s Overarching Integrated Product Team leaders, he handled a portfolio of 71 Major Defense Acquisition Programs including both strategic and tactical air, land, and sea systems. In this role, he directed numerous assessments of acquisition programs, policies, and practices; and he captured lessons and shared best practices. Of note, Dr. Lamartin presented more than 60 programs for Defense Acquisition Board consideration. To revitalize the practice of systems engineering, he stood up a new organization and set policy and defined processes for implementation. He also represented the U.S. government with allies to advance technology and develop weapons systems and testified to Congress to explain controversial decisions.

After his government service, Dr. Lamartin held executive positions at two premier aerospace and defense companies, Northrop Grumman and BAE Systems. There, he aligned industry products and services to meet DoD needs. He prepared strategic plans; managed the closure of a major manufacturing site; coordinated acquisition of an international company; aligned cross-sector mission assurance activities; and enhanced practices for system safety, engineering, and risk management. As Program Management Authority for a company sector, he made sure programs carried out contract requirements on schedule and within budget. He also led diagnosis, triage, and rapid turnaround of any with troubles. In addition, Dr. Lamartin saw to the education, placement, and coaching of program managers.

Dr. Lamartin is level III certified in Program Management. He holds a doctorate in public administration and master's in systems management from the University of Southern California, as well as a bachelor's degree in aerospace engineering from the University of Maryland. He lives in Great Falls, Virginia, with his wife Karen.

Mr. John E. Meeuwissen
Professor of Acquisition Management (International Cooperation)
Defense Systems Management College
Defense Acquisition University (DAU)

Mr. John Meeuwissen joined DAU in June 2009 as a Professor of Acquisition Management (International Cooperation); and is the Course Manager for ACQ 130, Fundamentals of Technology Security/Transfer (on-line) and ACQ 350, Advanced Technology Security/Control Workshop. He also has extensive curriculum development expertise and is the recipient of the 2013 DAU Curriculum Development Award.

Prior to joining DAU, Mr. Meeuwissen was a senior international analyst with DreamHammer, Inc. (September 2005 – June 2009) and AT&T Government Solutions (May 2003 – September 2005); providing on-site contractor support to the Office of the Under Secretary of Defense (Acquisition, Technology and Logistics) directorate of International Cooperation. His primary areas of expertise included technology transfer, export controls, national disclosure policy, international program management, and NATO standardization.

From December 1993 to May 2003, Mr. Meeuwissen was a senior research engineer and associate lab director at Georgia Tech Research Institute in Arlington VA. Technical, management and business development duties focused on managing GTRI's support activity for DOT&E and HQ USAF T&E in the areas of strategic planning, concept development, and methodology derivation for the effective and efficient use of test resources and infrastructure. Mr. Meeuwissen served in the U.S. Air Force from December 1971 to January 1994 as an instructor pilot, academic instructor, aircraft commander, research pilot, chief of safety, international R&D staff officer, and international cooperative programs manager.

Mr. Meeuwissen is a graduate of both the Joint PME Phase II at the Armed Forces Staff College and the Air War College at Maxwell AFB AL. His Master's Degree is in Industrial Relations and his Bachelor of Science is in Aerospace Engineering; both from West Virginia University.

Mr. John Mueller
Professor of Program Management
Defense Systems Management College
Defense Acquisition University (DAU)

John Mueller joined DAU in July 2010 as a Professor of Program Management in the executive acquisition faculty. He currently is the Center Director for the Executive Courses over seeing all the executive Program Management courses. He is also the DSMC Coaching Champion and a certified Executive Coach.

Prior to joining DAU, John served a 26 year Air Force career in a variety of acquisition / program management roles. This included working in program offices, logistics centers, major commands, and on the Air Staff focusing on delivering products to the warfighter.

His assignments included aircraft engine development and testing (Wright-Patterson AFB), space launch (Vandenberg AFB and Los Angeles AFB), program support & logistics (Kelly AFB). John also served as an F-22 Program Element Monitor on the Air Staff. He was selected as Commander, DCMA Greenville monitoring contractor compliance and quality for aircraft refurbishments. During this tour, he also served as a deployed Commander, DCMA Southern Iraq with responsibility for LOGCAP task orders on security, housing, and transportation. Following ICAF, John completed a joint duty tour with the Defense Threat Reduction Agency developing systems to counter Weapons of Mass Destruction. His final assignment was the Vice Commander, Space Based Infrared System (SBIRS) Wing with responsibility for the development, deployment, and certification of the newest DoD IR satellite space and ground systems.

John holds a Master of Science degree in National Resource Strategy from the Industrial College of the Armed Forces, June 2005 and a Master of Business Administration from the University of Dayton, May 1987. His undergraduate degree is from the University of Notre Dame in Mechanical Engineering. He is Level III certified in program management, acquisition logistics and SPRDE. Mr. Mueller is a graduate of the Program Manager's Course (98-2) and PMT-402 (09-4).

Greg Prothero
Professor of Requirements Management Training
Defense Systems Management College
Defense Acquisition University (DAU)

Greg Prothero joined DAU as a Professor of Requirements Management Training in July 2009. His Air Force service includes operational flying, training and platform instruction, Congressional Liaison and AF Requirements Development.

Greg flew as an Evaluator Navigator in C-130Es and H3s at Little Rock Air Force Base, AR and Pope Air Force Base, NC. He flew operational missions in Southwest Asia, East Africa, The Balkans and Latin America. He also had a tour as an instructor in Joint Undergraduate Navigator Training at Randolph Air Force Base, TX.

As an Assistant Professor of Management at the United States Air Force Academy, he taught and course directed the basic Management course required for all cadets. While at United States Air Force Academy, he was an Advance Agent for Air Force One as an additional duty.

Greg transferred to the Pentagon in 2004 and planned all Air Force sponsored Congressional travel as part of Air Force Legislative Liaison.

He spent his last assignment at Headquarters United States Air Force in A5 in requirements development. He ran the regular and special session of the Air Force Requirements for Operational Capabilities Council. After attending the pilot Requirements 110 program at DAU, Greg was responsible for all USAF course completion and reporting to J-8 and OSD.

Greg has a BS in Public Administration from Indiana University and a MS in Operations Management from the University of Arkansas.

Mr. Jim Ryan
Professor of Acquisition Management
Defense Systems Management College
School of Program Managers
Defense Acquisition University (DAU)

Mr. Ryan joined the School of Program Managers faculty in November of 2010, following retirement from the U.S. Air Force. Mr. Ryan has over 19 years of program management, manufacturing, requirements, contract pricing and legislative liaison experience.

Beginning in 1986, he served as a Program Manager and Manufacturing Manager within the Propulsion Systems Program Office. He also taught contract pricing and cost estimating at the Air Force Institute of Technology. Follow-on tours included managing the requirements for their Aircrew Training Systems and later the CV-22 Osprey aircraft program at HQ Air Force Special Operations Command; C-130J Program Element Monitor at Headquarters USAF, and later as Legislative Liaison for Airlift, Tanker, Transport and Special Operations aircraft to the Appropriations Committees.

Mr. Ryan also held positions as Executive Officer to the Vice Commander, United States Air Forces in Europe and as the Commander, 422nd Air Base Squadron, RAF Croughton in the United Kingdom. He culminated his military career as the Production Program Manager for the F-35 Lightning II, Joint Strike Fighter Program.

Mr. Ryan is a graduate of both the Program Management Course and Program Manager's Course. He is Level III certified in Program Management. Mr. Ryan holds a Bachelor of Science (Engineering) from the United States Air Force Academy and a Master of Science in Industrial Engineering from Texas A&M University.

Mr. S.L. "Dusty" Schilling
Professor of Program Management
Defense Systems Management College
School of Program Managers
Defense Acquisition University (DAU)

Mr. S. L. "Dusty" Schilling joined DAU in October 2008 as a Professor of Program Management, having spent 30 years in and around the US Navy, as a Naval Officer, a government civil servant, and then as a management consultant.

That career in the Defense industry brought Dusty "back through the customer chain" from operator to system delivery, to customer support, to production management to research, development, test & evaluation, to participating in setting the strategy of the Navy to advising acquisition managers.

Prior to joining DAU Dusty founded Centaur FCG, Inc., a small business that provided consulting services in management, strategic planning, business operations, and technical analysis.

As a government civil servant Dusty served in a number of positions in US Navy Program Executive Offices (PEOs) including as Deputy Director for Programs for Navy Theater Air Defense, as Director for Integrated Program Management for Theater Surface Combatants, and finally as Director of the Navy Liaison Team at the Missile Defense Agency.

Before his PEO tours Dusty was the Deputy Program Manager for the US Navy Ballistic Missile Defense Acquisition Category ID Major Defense Acquisition Program.

Dusty first served for seven years as a US Navy Surface Warfare officer, then entered Civil Service as the Waterfront Engineering Team lead for Pacific Fleet Aegis ships. After four years he joined the Aegis Program Office in Washington initially as the Production Manager for the US Navy Aegis Weapons System, and then as the Acquisition Manager.

Dusty's education includes an MBA from National University, a BA in General Physics from UCLA, completion of the Program Manager's class at the Defense Systems Management College and DoD Level III DAWIA certification in Program Management.

Mr. Harry L. Snodgrass
Professor of Acquisition Management
Defense Systems Management College
School of Program Managers
Defense Acquisition University (DAU)

Mr. Snodgrass joined the School of Program Managers faculty in August of 2007 after concluding a 26-year career in the United States Air Force (USAF).

Mr. Snodgrass has over 19 years of program management, manufacturing, and engineering acquisition experience. After commissioning in 1981, he had operational tours flying the B-52H in the Strategic Air Command and the T-43 aircraft in the Air Training Command. He entered the acquisition career field through the Education with Industry (EWI) program. Following the EWI program at Lockheed Martin, he was assigned to the F-22 System Program Office, Wright Patterson AFB, OH. Follow on tours included working at Special Access Programs; teaching manufacturing and engineering at DAU; and serving at Headquarters USAF, Pentagon.

Mr. Snodgrass is a graduate of both the Program Management Course and Program Manager's Course. He is Level III certified in Program Management, Systems Engineering, and Manufacturing. Mr. Snodgrass holds a Bachelor of Arts from the University of Tennessee and a Master of Science in Administration from Central Michigan University.

Lieutenant General (Retired) N. Ross Thompson III

Ross Thompson recently served as the Executive Vice President/ Chief Operating Officer for ABM Facility Solutions Group (FSG), an integrated facilities management company with over $700 million in annual revenue and 7,100 employees across the U.S. and in 10 international locations. While with ABM, he was also serving concurrently as the President, ABM Government Services, one of the three operating companies within the Facility Solutions Group. He provided overall direction and guidance to enterprise-wide activities, and led efforts to accomplish established corporate growth goals, both domestically and internationally. Revenue and profit had double digit growth every year for the last five years across the FSG companies.

Mr. Thompson led the efforts to integrate the on-site facilities services of ABM Engineering and Linc Facility Services into one company resulting from ABM's acquisition of The Linc Group in 2010. Mr. Thompson also guided ABM's growing business interests in the other FSG operating companies and had a senior leadership role in the post-acquisition integration projects for ten companies acquired by ABM during his time as EVP/COO.

Prior to joining The Linc Group in April 2010, Mr. Thompson served for 36 years in the United States Army, retiring as a Lieutenant General. His last assignment was as the Principal Military Deputy, Assistant Secretary of the Army (Acquisition, Logistics and Technology), the Army's senior acquisition officer. In this position he directed the Army's acquisition workforce, the program managers for all of the Army's weapons systems, and provided policy and programmatic direction for the Army's research and development program.

Mr. Thompson held numerous multidiscipline command and staff positions throughout his Army career. As the Army's Director of Program Analysis and Evaluation, he managed the allocation of financial resources for the annual budget that exceeded $200 billion, as well as the operations research/ systems analysis workforce. He also led the Army-wide deployment and training of the Lean Six Sigma program, personally receiving a Lean Six Sigma black belt for his project that made continuous improvement part of the Army culture.

Mr. Thompson's last command position was the Commanding General, United States Army Tank-Automotive and Armaments Command in Warren, Michigan, responsible for world-wide logistics, technology and contracting support for Army ground systems. In this position he led a workforce of more than 15,000 personnel with an annual budget authority over $10 billion. The Army's five ground system maintenance, manufacturing, and supply depots quadrupled their workload over three years, while implementing Lean Six Sigma, in response to wartime demands after 9-11.

Mr. Thompson's distinguished military career included extensive command and leadership experience, with a proven track record of success across the disciplines of logistics, operations, technology, procurement, and financial management. He continues

to emphasize values, continuous improvement, personnel development, and customer focus as he advises organizations on strategy, business development, operations and financial planning.

Mr. Thompson is a distinguished military graduate of the College of William and Mary, earning a degree in biology. He also has an MBA in operations research/systems analysis (engineering) from George Washington University, as well as an MA in international relations from Salve Regina University. He is a graduate of the Naval College of Command and Staff and the Industrial College of the Armed Forces, receiving master's degrees in National Security Decision Making and Resource Management. His military decorations include multiple awards of the Army Distinguished Service Medal, Legion of Merit, and Meritorious Service Medal, as well as the Senior Parachutist Badge and Ranger Tab.

Mr. Thompson lives in Alexandria, Virginia, with his wife and three children.

Appendix B - Technology Insertion Scenario - Instructions to Subjects

Technology Insertion Scenario
Instructions to Subjects (Rev-D)

The Student Investigator will read the following instructions to the subjects:

The Marlin torpedo has been a very reliable weapon system for the submarine fleet. The torpedo uses a contact fuse to detonate a high explosive charge that breeches the enemy ship's hull just below the waterline. The torpedo has two vulnerabilities. Since the torpedo travels close to the surface, the enemy can detect the torpedo's wake. This allows the enemy ship to take evasive maneuvers. The torpedo wake also indicates from which direction the torpedo was fired; giving away the firing submarine's position. The fact that the torpedo strikes the enemy ship just below the waterline is its second vulnerability. The enemy has added armor plating at the waterline to defeat the torpedo's high explosive charge.

The Sea Urchin anti-ship mine is also very reliable. The Sea Urchin is managed by the Sea Mines Program Office. The mine operates from an anchored position. A cable maintains the Sea Urchin at a fixed depth, well below the surface. The Sea Urchin detects the magnetic field of a ship passing above and detonates. The pressure from the explosion attacks the enemy ship at its weakest point and cracks the hull at the keel like an egg. The Marlin Project Office plans to integrate the magnetic fuse from the Sea Urchin to the Marlin torpedo while maintaining the contact fuse capability as a secondary mode of operation. This will allow the Marlin to run deeper, with no detectable wake and avoid the additional armor the enemy is adding to its ships at the waterline.

Admiral Tyrosine is concerned that the Lieutenant Bedford and the Marlin Project Office's identified risks are predictable and not very creative (see her sticky note on the project charter). You are employees of **Expert Risk International**, a company that specializes in risk identification techniques. Admiral Tyrosine has hired you to take a fresh look Lieutenant Bedford's plan for the Marlin Torpedo Upgrade. She wants to avoid being blind-sided by the technical and operational aspects of the plan. You have two deliverables. The first deliverable is a list of risks in the Marlin Torpedo Upgrade plan. The second deliverable is a set of recommended changes to the plan. Admiral Tyrosine has also hired a company called **Risk Geniuses** to do the same task. She will award a follow-on contract to the team with risks that are creative and plausible and changes to the plan that are creative and useful. You must turn-in your deliverables sheet on time to be considered.

Use the (Brainstorming/PreMortem) technique for risk identification to develop your deliverables. Please appoint a team lead and follow the steps for your assigned technique for risk identification:

Brainstorming Technique (ref. Goal/QPC, 1994)	**PreMortem Technique (ref. Klein, 2007)**
1. Get familiar with the plan. (3-mins)	1. Get familiar with the plan. (3-mins)
2. Team Lead writes the central question on the butcher paper.	2. "Imagine the plan has failed, a complete fiasco. What could have caused this?"
You Have 45-mins for the Remaining Steps	**You Have 45-mins for the Remaining Steps**
3. Each team member gives an idea in turn. No idea is criticized. Team Lead records on the butcher paper. Do this three times or until all ideas are exhausted (whatever occurs first).	3. Individually, list the reasons for failure. Recommend you take 3-minutes for this task.
4. Review the list for clarity, discard any duplicates, and discuss. Prepare your risk list and changes to the plan.	4. Each team member provides one item form their list in turn (without discussion), recording on the butcher paper. Do this three times or until all ideas are exhausted.
5. Complete your deliverable sheet.	5. Discuss and collaborate. Prepare your risk list and changes to the plan.
	6. Complete your deliverable sheet.

The only information you have is the project charter. Any questions?

Appendix C - Marlin Torpedo Project Charter

Marlin Upgrade Project Charter (version 1.2)

Project Name: Marlin Torpedo Upgrade	
Prepared by: C. Bar (Project Administrator)	
Date: 31 October 2015	
Initiation:	*Includes the name of the project and the designated project manager* Marlin Torpedo Upgrade (magnetic-sensing, proximity fuse integration) Project Manager: Lieutenant A. Bedford (USNR)
Synopsis:	*Briefly summarize the salient aspects of the project by answering the questions "Why?" (purpose), "What?" (product description/ scope), "When?" (time), and "How much?" (resources)* This project will integrate the fielded M211 proximity fuse from the Sea Urchin anti-ship mine. This will be an upgrade to the Marlin torpedo which currently depends on a high-reliability contact fuse. There is an urgent need for the upgraded Marlin—initial operational capability is scheduled for May 2017. The budget for integration and test is $1.8M. The budget to field the initial set of 20 Upgraded Marlins is $1.1M. Since the Government has surplus Sea Urchin mines and Marlin torpedoes (and the technical data for both), this will be a Government-led effort.
Purpose/Need:	*Identify the customers who are to receive and benefit from the product developed by the project and the need the product is intended to meet (either as a problem to solve or as an opportunity to exploit)* The Marlin torpedo attacks enemy ships just below the waterline. A contact fuse initiates a high explosive charge that penetrates the ship's hull. This method of attack has proven to be highly reliable. However, the enemy can visually detect the incoming Marlin riding close to the surface and perform evasive maneuvers. We also have reports of the enemy reinforcing ship hulls below the waterline to increase survivability against the Marlin. The upgraded Marlin will retain the high reliability contact fuse while integrating the proven M211 proximity fuse from the Sea Urchin mine. The proximity fuse detects the magnetic field of an enemy ship and detonates the high explosive charge underneath the ship. The water pressure from the detonation cracks the enemy ship's hull. The benefits are: (1) Reduces Marlin detection by traveling deeper (2) Attacks the weakest part of the ship and avoids the armor the enemy is deploying near the waterline (3) Provides the operator with selectable attack modes; retaining the contact fuse capability as a secondary mode.

Product Description and Deliverables:	*Identify what product is to be delivered at the end of the project, and at any interim delivery points. Describe the product sufficiently to enable the project team to create it, and for agreement to be reached at product delivery time that the product has been correctly produced* Two developmental test articles by Feb 2016. Four operational test articles by June 2016. First deliveries to the fleet October 2016. Initial Operational Capability May 2017.
Risks	*Identify the top risks for the project.* (1) If the Marlin PMO does not receive $1.8M then the Marlin Upgrade integration and test will be delayed or curtailed. (2) If integration and test are curtailed then technical problems will not be discovered before operational use. (3) If we discover too many technical problems during integration and test then we will not make Initial Operational Capability in 2017. (4) If the schedule slips because of technical problems found in integration and test then we will exceed the $1.8M budget. (5) If the operational concept or threat changes, then technical changes will be necessary. (6) If technical changes are necessary then we will not meet schedule and budget.
Approval	Project Manager: A. Bedford, Lieutenant, USNR Sponsor: S. Tyrosine, RADM, USN

Appendix D - Subject Questionnaire

Subject Questionnaire (Rev-B) dated 10-1-15

1. The risk identification technique my team used in the "torpedo" scenario was:

☐ Brainstorming ☐ PreMortem Technique

2. I have been conducting risk identification in a program management or systems engineering setting for: ☐ less than 10 years ☐ between 10 and 15 years ☐ more than 15 years

3. My job within the organization would best be described as: ☐ Programmatic ☐ Technical

4. The technique for risk identification that I have used most frequently is:

☐ Pros and Cons ☐ Critique ☐ Devil's Advocate ☐ Red Team ☐ Divide and Conquer

☐ Brainstorming ☐ Scenario Development ☐ PreMortem ☐ Other _____

Consider the "torpedo" scenario you just completed, given the amount of time and information available to you, please check the block that best describes your agreement with the statement.

---

5. The risk identification technique allowed me to express my individual difference in perspective.

Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐

6. The risk identification technique allowed me to understand others' difference in perspective.

Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐

7. The risk identification technique allowed me to enrich the ideas of my teammates.

Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐

8. The risk identification technique allowed me to critique the ideas of my teammates.

Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐

Items 9 thru 12 ask you to compare the risk identification technique you used in the "torpedo" scenario (Brainstorming or PreMortem) to other risk identification techniques you have used. Consider a risk identification activity in which you previously participated. Consider what went well and not so well. Please check the block that best describes your agreement with the statement.

---

9. If we had used the risk identification technique used in the "torpedo" scenario, we would have discovered more risks in the same amount of time.

Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐

10. If we had used the risk identification technique used in the "torpedo" scenario, we would have discovered better risks in the same amount of time.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

11. If we had used the risk identification technique used in the "torpedo" scenario, we would have been able to propose more creative improvements to the plan in the same amount of time.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

12. If we had used the risk identification technique used in the "torpedo" scenario, we would have been less vulnerable to organizational blind-spots.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

Consider projects that you have been a team member or may currently be a team member. Please check the most appropriate block.

---

13. We believe that if we can predict all the risks and successfully prepare for their occurrence, we will have a successful project.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

14. Our projects are well ordered and have a rich history of past performance that makes predicting how problems will occur reliable.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

15. Gaps in perspective between the technical and program (business) sides of the project are a problem.

Strongly Disagree ☐ Disagree ☐ Somewhat Disagree ☐ Somewhat Agree ☐ Agree ☐ Strongly Agree ☐

Appendix E - Team Deliverable Templates

Risk List Template
Marlin Torpedo Upgrade Plan Risk Identification Deliverable

The team offers the following **risks** for the Admiral's consideration. They are not in priority order:

A	
B	
C	
D	

Changes to the Plan Template
Marlin Torpedo Upgrade

The team offers the following changes to the Marlin Torpedo Upgrade for the Admiral's consideration. They are not in priority order:

E	
F	
G	
H	

Appendix F - Expert Panel Instructions Sheet

Step-1 Read the Consent Form

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC
Consent Statement for Exempt Research (Expert Panel)

Title of Study: Benefits of the PreMortem Technique for Risk Identification in a Systems Engineering Context
IRB #: 1155 8
Principal Investigator Name: Dr. T. Mazzuchi
Version Date: 10-1-2015

You are invited to participate in a research study under the direction of Dr. T. Mazzuci of the Department of Engineering and Applied Sciences, George Washington University (GWU). Taking part in this research is entirely voluntary. Your academic standing or the status of your employment will not, in any way, be affected should you choose not to participate or if you decide to withdraw from the study at any time. Further information regarding this study may be obtained by contacting David L. Gallop, the student investigator, at telephone number 703-804-5943.

The purpose of this study is to find out if the PreMortem Technique works better than brainstorming for risk identification in a systems engineering context.

If you choose to take part in this study, you will: (1) Review and grade sets of products/deliverables developed by subjects in a risk identification workshop. (research activity) (2) Review the risk lists developed in the workshops in the against an expected risk list. (research activity) (3) Identify the top three sets of deliverables. (research activity)

Your participation and expert panel input (grading) will be known to the research team. You will not know the workshop participant identities nor will they know your identity. The total amount of time you will spend in connection with this study is *two-hours*. You may stop your participation in this study at any time.

Possible risks or discomforts you could experience during this study include:

Risk #1: Waste the panel member's time. The likelihood of this risk is minimal. This risk will be controlled by providing a description of the research in advance and various to opt out.

Risk #2: Loss of Confidentiality. The likelihood of this risk is minimal. This risk will be controlled by not asking for personally identifying information on the questionnaires; securing the questionnaires in a locked container; and destroying the questionnaires at the no later than May 31, 2017.

Risk #3: Anxiety in assigning grades. The likelihood of this risk is minimal. This risk will be controlled by preparing participants for the activity and providing the opportunity to opt-out.

You will not benefit directly from your participation in the study. The benefits to science and humankind that might result from this study are: *The application of a new technique for risk identification in the systems engineering field.*

Every effort will be made to keep your information confidential, however, this can not be guaranteed. *The score sheet will not ask you for personally identifying information. Completed score sheets will be destroyed at the conclusion of the research or by 31 May 2017 (whichever comes first).* If results of this research study are reported in journals or at scientific meetings, the people who participated in this study will not be named or identified.

The Office of Human Research of George Washington University, at telephone number (202) 994-2715, can provide further information about your rights as a research participant.

To ensure anonymity your signature is not required, unless you prefer to sign it. Your willingness to participate in this research study is implied if you proceed.
*Please keep a copy of this document in case you want to read it again.

Step-2 Read the Scenario

Technology Insertion Scenario
Instructions to Subjects (Rev-D) 10-1-15

e Student Investigator will read the following instructions to the subjects:

e Marlin torpedo has been a very reliable weapon system for the submarine fleet. The torpedo uses a contact fuse to ionate a high explosive charge that breeches the enemy ship's hull just below the waterline. The torpedo has two lnerabilities. Since the torpedo travels close to the surface, the enemy can detect the torpedo's wake. This allows the emy ship to take evasive maneuvers. The torpedo wake also indicates from which direction the torpedo was fired, ving enemy the firing submarine's position. The fact that the torpedo strikes the enemy ship just below the waterline is i second vulnerability. The enemy has added armor plating at the waterline to defeat the torpedo's high explosive arge.

e Sea Urchin anti-ship mine is also very reliable. The Sea Urchin is managed by the Sea Mines Program Office. The ine operates from an anchored position. A cable maintains the Sea Urchin at a fixed depth, well below the surface. e Sea Urchin detects the magnetic field of a ship passing above and detonates. The pressure from the explosion tacks the enemy ship at its weakest point and cracks the hull at the keel like an egg. The Marlin Project Office plans to tegrate the magnetic fuse from the Sea Urchin to the Marlin torpedo while maintaining the contact fuse capability as a econdary mode of operation. This will allow the Marlin to run deeper, with no detectable wake and avoid the dditional armor the enemy is adding to its ships at the waterline.

dmiral Tyrosine is concerned that the Lieutenant Bedford and the Marlin Project Office's identified risks are predictable nd not very creative (see her sticky note on the project charter). You are employees of **Expert Risk International**, a ompany that specializes in risk identification techniques. Admiral Tyrosine has hired you to take a fresh look Lieutenant Bedford's plan for the Marlin Torpedo Upgrade. She wants to avoid being blind-sided by the technical and perational aspects of the plan. You have two deliverables. The first deliverable is a list of risks in the Marlin Torpedo ograde plan. The second deliverable is a set of recommended changes to the plan. Admiral Tyrosine has also hired a company called **Risk Geniuses** to do the same task. She will award a follow-on contract to the team with risks that are reative and plausible and changes to the plan that are creative and useful. You must turn-in your deliverables sheet on ime to be considered. ***Use the (Brainstorming/PreMortem) technique for risk identification to develop your deliverables. Please appoint a team lead and follow the steps for your assigned technique for risk identification: *the echnique that does not apply will be blanked-out****

Brainstorming Technique (ref. Goal/GPC, 1994)	PreMortem Technique (ref. Klein, 2007)
1. Get familiar with the plan. (5-mins)	1. Get familiar with the plan. (5-mins)
2. Team Lead writes the central question on the butcher paper.	2. *"Imagine the plan has failed, a complete fiasco. What could have caused this?"*
You Have 45-mins for the Remaining Steps	**You Have 45-mins for the Remaining Steps**
3. Each team member gives an idea in turn. No idea is criticized. Team Lead records on the butcher paper. Do this until all ideas are exhausted.	3. Individually, list the reasons for failure. (Recommend you take 3-minutes for this task.)
4. Review the list for clarity, discard any duplicates, and discuss.	4. Each team member provides one item from their list in turn (without discussion), recording on the butcher paper. Do this three times or until all ideas are exhausted.
5. Prepare your risk list and changes to the plan.	5. Discuss and collaborate. Prepare your risk list and changes to the plan.
6. Complete your deliverable sheets.	6. Complete your deliverable sheets.

The only information you have is the project charter. Any questions?

Step-3 Read the Marlin Upgrade Project Charter

Marlin Upgrade Project Charter (version 1.2) **10-1-15**

Project Name: Marlin Torpedo Upgrade	
Prepared by: C. Bar (Project Administrator)	
Date: 31 October 2015	
Initiation:	*Includes the name of the project and the designated project manager* Marlin Torpedo Upgrade (magnetic-sensing, proximity fuse integration) Project Manager: Lieutenant A. Bedford (USNR)
Synopsis:	*Briefly summarize the salient aspects of the project by answering the questions "Why?" (purpose), "What?" (product description/ scope), "When?" (time), and "How much?" (resources)* This project will integrate the fielded M211 proximity fuse from the Sea Urchin anti-ship mine. This will be an upgrade to the Marlin torpedo which currently depends on a high-reliability contact fuse. There is an urgent need for the upgraded Marlin—initial operational capability is scheduled for May 2017. The budget for integration and test is $1.8M. The budget to field the initial set of 20 Upgraded Marlins is $1.1M. Since the Government has surplus Sea Urchin mines and Marlin torpedoes (and the technical data for both), this will be a Government-led effort.
Purpose/Need:	*Identify the customers who are to receive and benefit from the product developed by the project and the need the product is intended to meet (either as a problem to solve or as an opportunity to exploit)* The Marlin torpedo attacks enemy ships just below the waterline. A contact fuse initiates a high explosive charge that penetrates the ship's hull. This method of attack has proven to be highly reliable. However, the enemy can visually detect the incoming Marlin riding close to the surface and perform evasive maneuvers. We also have reports of the enemy reinforcing ship hulls below the waterline to increase survivability against the Marlin. The upgraded Marlin will retain the high reliability contact fuse while integrating the proven M211 proximity fuse from the Sea Urchin mine. The proximity fuse detects the magnetic field of an enemy ship and detonates the high explosive charge underneath the ship. The water pressure from the detonation cracks the enemy ship's hull. The benefits are: (1) Reduces Marlin detection by traveling deeper (2) Attacks the weakest part of the ship and avoids the armor the enemy is deploying near the waterline (3) Provides the operator with selectable attack modes; retaining the contact fuse capability as a secondary mode.
Product	*Identify what product is to be delivered at the end of the project, and at any*
Description and Deliverables:	*interim delivery points. Describe the product sufficiently to enable the project team to create it, and for agreement to be reached at product delivery time that the product has been correctly produced* Two developmental test articles by Feb 2016. Four operational test articles by June 2016. First deliveries to the fleet October 2016. Initial Operational Capability May 2017.
Risks	*Identify the top risks for the project.* (1) If the Marlin PMO does not receive $1.8M thes the Marlin Upgrade integration and test will be delayed or curtailed. (2) If integration and test are curtailed then technical problems will not be discovered before operational use. (3) If we discover too many technical problems during integration and test then we will not make Initial Operational Capability in 2017. (4) If the schedule slips because of technical problems found in integration and test then we will exceed the $1.8M budget. (5) If the operational concept or threat changes, then technical changes will be necessary. (6) If technical changes are necessary then we will not meet schedule and budget.
Approval	Project Manager: A. Bedford, Lieutenant, USNR Sponsor: S. Tyrosine, RADM, USN

Lieut. Bedford:
I am disappointed in the risks you have identified. They are predictable and not very creative. I don't want to be blind sided by technical or operational issues.
-Tyrosine

The Instructions Sheet

Step-4 Read the Expert Panel Scoring Instructions

Expert Panel Scoring Instructions

- Assume the role of Admiral Tyrosine.
- You will be scoring 20 sets of deliverables.
- Each deliverable has two parts– "Risk List" and "Changes to the Plan"
- The Risk List has four risks. You will be scoring each risk along two dimensions– Creative and Plausible. Use a 0 to 100 grading scale. You will also mark whether or not the each of the four risks match the four Black Swans.
- The Changes to the Plan has four changes. You will be scoring each change along two dimensions– Creative and Useful. Use a 0 to 100 grading scale.
- You do not have to justify your scores. A grade of 70 or greater should represent "quality" score.

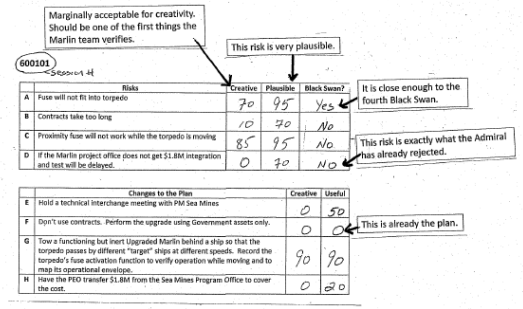
Definitions
Creative- having or showing imagination and artistic or intellectual inventiveness.
Plausible- seemingly true, acceptable.
Useful- having practical utility.

Step-5 Become Familiar with the Four "Black Swans"

"Black Swans": Risks that the real world/historical team missed, but in hindsight should have caught.

- Tested torpedoes without explosives in order to conserve limited test articles.
- Tested torpedoes without hitting objects to recover and reuse the torpedoes for more testing.
- The torpedoes ran too deep to sense the magnetic field of the enemy ship.
- Integration of the magnetic fuse interfered with the contact detonation mechanism causing torpedoes to bounce-off the ships without detonating when in the "contact fuse" mode.

Step-6 Review the Example Score Sheet with Comments



	Risks	Creative	Plausible	Black Swan?
A	Fuse will not fit into torpedo	70	95	Yes
B	Contracts take too long	10	70	No
C	Proximity fuse will not work while the torpedo is moving	85	95	No
D	If the Marlin project office does not get $1.8M integration and test will be delayed.	0	70	NO

	Changes to the Plan	Creative	Useful
E	Hold a technical interchange meeting with PM Sea Mines	0	50
F	Don't use contracts. Perform the upgrade using Government assets only.	0	0
G	Tow a functioning but inert Upgraded Marlin behind a ship so that the torpedo passes by different "target" ships at different speeds. Record the torpedo's fuse activation function to verify operation while moving and to map its operational envelope.	90	90
H	Have the PEO transfer $1.8M from the Sea Mines Program Office to cover the cost.	0	20

Appendix G - Expert Panel Score Sheet

Step-7 Score the Deliverables

532901

	Risk	Creative	Plausible	Black Swan?
A	Enemy would develop a hull design that makes standard/conventional assault missiles hull-on-floating flies over the length of the ship.			
B	Larger warhead required for stand-off attack—does it fit in the launch tubes? Is the new warhead feasible in a torpedo?			
C	These two systems were not designed to be placed in the same space. Therefore combining components from both may result in unacceptable performance degradation due to material incompatibility.			
D	Connecting the two guidance electronic packages may require more wires, power, cooling etc. than is available in the current torpedo design.			

	Changes to the Plan	Creative	Useful
1	Look at additional attack vectors—vertical or at extreme. Work with your community to see if there is a way to both design changes and decrease max change needed by most likely hulls the torpedo would be used against.		
2	Look into shaping charges to mitigate cons, weight, and power issues. Such changes could be change to the projectile or else new chemical armament to be used in conjunction with it.		
3	Work with Torpedo Material Science Lab(s) to review materials list. Try to find other materials on bulk value side—use tolerance of alternate materials.		
4	Perform an early study or frequency for information to balance to understand feasibility. Incentivize contractors to reduce size, power, cooling of electronics package. Consider vertical axis propeller/ duct retrofit.		

538801

	Risk	Creative	Plausible	Black Swan?
A	If no room for fuse, then a major redesign is required.			
B	If fuse wires are not available, then the program will be delayed.			
C	If the explosability of the fuse (EX) is too violent for mesh, then additional capability is required.			
D	If the payload is insufficient to accomplish the mission, then additional modifications will be needed.			

	Changes to the Plan	Creative	Useful
A	Will need to redesign the Mantis Torpedo. It will require an extension to the T&E schedule and additional $ for ~$100,000 +/- to design the missile.		
B	To ensure facilities are available, include a requirement for facility in the RFP and conduct site visits.		
C	Team will need to redefine, or re-design the weight of payload, the following balance the two items. – ships funding + additional time may be required. – look into/ware with Dive Commander to relax requirements or doesn't impact mantic or lower change. – Curtail the program.		
D	PM required to MALO to determine if the existing warhead payload needs additional requirements.		

538901

	Risk	Creative	Plausible	Black Swan?
A	If the upgrades to the Mantis decrease the probability of the kill then the upgrade will not be successful.			
B	If the integration of the new fuse reduces safety related detonations then we could inflict unintended consequences to our own ship and others in close proximity.			
C	If there is not a cyber-protection plan developed then the Mantis is susceptible to cyber-attack.			
D	If the sensors for countermeasure detection abilities then they could compromise the capabilities of our system.			

	Changes to the Plan	Creative	Useful
E	Ensure the test at sea to determine the warhead is effective with the possible fuse. (i.e. it kills the enemy part of the ship).		
F	Do cyber testing to determine whether the magnetic proximity fuse can differentiate friendly from enemy ships and is not a danger to the firing ship.		
G	Create an initiative to attack our system via cyber capabilities and implement changes as they are identified.		
H	Combine trial changes in Mantic makes decision.		

539001

	Risk	Creative	Plausible	Black Swan?
A	Sensor is not to life. Enemy can find large sea animals such as whales or dolphins will interrupt transmission and electronic signals.			
B	There is concerned that new design technology may contain rare components or material that is limited or hard to obtain. This may necessitate a more flexible design or an increase in domestic capability to supply such.			
C	Enemy has new technology (such as stealth) to defeat capacity rendering this device obsolete before IOC.			
D	Technology of the MK51 proximity fuse does not transfer from the static environment of a floating mine to the dynamic environment of a rapidly moving torpedo such that a failure of the torpedo is detected.			

	Changes to the Plan	Creative	Useful
E	Get additional funding for testing to ensure there is no damage to nearby life—sealife or otherwise—along all navigational tools identified and killing marine life.		
F	Recommend that market research is conducted for alternate materials that enable a more flexible design. Plan for technology insertion and pursuit for domestic availability of our raw earth materials.		
G	Create a red team to simulate enemy capabilities to ensure torpedo will be effective. Solicit support of intel agencies to help identify if new enemy detection technologies.		
H	Develop and complete test plan to evaluate the MK51 in a dynamic environment. This could be done by using a stationary platform placed in water flowing and the mine/ sensor being re-housed to using the torpedo body in open water.		

539401

	Risk	Creative	Plausible	Black Swan?
A	Technology, 1) reference efficacy of combination 2) impact on targets.			
B	Enemy technology defeats signals.			
C	Operational impacts 1) Cost, schedule of 2) Level			
D	Lack of internal stakeholder support.			

	Changes to the Plan	Creative	Useful
E	Configuration control of torpedo dimension. Request more money.		
F	Conduct total (EW) for latest on enemy capabilities.		
G	Ability to turn on magnetic/ being for remote control.		
H	Congressional/allied engagement. Industry support/ engagement. Pentagon/press briefing/Congress.		

539301

	Risk	Creative	Plausible	Black Swan?
A	If the design modification, including the revised (thinner) attack profile makes the torpedo vulnerable to countermeasures or interference from new ship, or would require a redesign of the warhead to be effective, then the weapon would not achieve its goal to meet operational need.			
B	There is a concern that sufficient time is available for the production/creation of the new torpedo. This includes the people available to do the work/experts, materials available (maybe focus on available/ready-for-issue), production capacity and timelines that impact technical approach. If these are not in place, the projected timeline would be impacted to meet fleet operation capability needs.			
C	Technical. If additional operator training and facilities are required, then funding will need to be authorized to support. Current budget does not have funding allocated for additional operator training or additional training facilities.			
D	Technical risk exists in the physical integration of the core fuse into a conflict torpedo. If the Mantic sensor support a second fuse (physical or RHS), does not provide an adequate warhead, or cannot be reliably controlled at depth, additional design and testing efforts will be needed, impacting schedule and cost.			

	Changes to the Plan	Creative	Useful
A	Study the following: warhead design effectiveness at depth (work with other torpedo designer for this profile), known magnetic countermeasures, tactics and weapons employ to mitigate new ship interference, incorporate corporate knowledge for water fuse design.		
B	Money identified to fund a study looking at counter force ways identifying any issues. Set experts on contract to help work on identifying issues. Layout production plan to see if schedule. Work with technical experts and engineers to see if there is requirements?		
C	Stop testing and begin to support fleetwood attack mode. May not be able to say it is "FOW" for validated, so will need to reallocate or extend out procurement of expended systems. Leverage testing that exists for values.		
D	Ensure testing of the thin fuse at depth is included in the Developmental Test (DT) plan 9 create additional DT unit/ or do flooded hard testing in a way that does not destroy the test articles until the torpedo body has been tested at depth. Model testing to look at credible approaches also.			

535101

	Risks	Creative	Plausible	Black Swan?
A	Lack of user ship shore interface			
B	Changes in training, logistics, storage			
C	There are insufficient assets to make the conversion from the mine fuse and not enough available torpedoes to make the changes. They got used for higher priority operational needs.			
D	Personnel trained and available/transitioned with magnetic sensor and fuse such as mining torpedo.			

	Changes to the Plan	Creative	Useful
E	Understand employment of weapon/other platforms with conv'd		
F	Planning drill set analysis and available. Understand changes in range—to influence logistics and storage. Update task documentation and employment documentation. Repair and maintenance manuals.		
G	Resource (per headquarters) augmentees to a dedicated number of Test and Conversion assets. Plan for portable modification kit that could be used to upgrade the torpedoes out to the field. Plan for building new parts for the modification kits (sensors, fuse, etc.)		
H	Evaluate how weapon will be used in scenario and threats. Evaluate technical/resource based on above bullet. Evaluate the current mines torpedo dimensions.			

535501

	Risks	Creative	Plausible	Black Swan?
A	May be structural or integration issues that would cause extended post-install efforts.			
B	Cost of performing modification outweighs perceived benefits (i.e. the cost of torpedoes isn't bad).			
C	Do not have enough trained engineering talent and resources to properly conduct integration.			
D	Changes in system performance in testing depth could have adverse effects on system performance or mine operations.			

	Changes to the Plan	Creative	Useful
E	Go to a single fuse (for magnetic fuse from this mine)		
F	Fly to an exercise. Go to a single fuse (the magnetic fuse from the mine). Perform a Business Case Analysis. Issue a Request for Information.		
G	Contract for engineering.		
H	Build new software. Ensure depth analysis incorporated in test plan by development and specialized testing.		

606901

	Risk	Creative	Plausible	Black Swan?
A	Risk of enemy countermeasures which would reverse the torpedo to impact a false target. Enemy could release small bulb that emit a magnetic signature that draws the torpedo to it instead of the intended target.			
B	Carrying the altered torpedoes on our vessels could change the magnetic signature of our vessels, potentially affecting sensitive ship's electronics.			
C	When pulling the magnetic fuse from the torpedo is a threat that the fuse cannot be reset.			
D	Magnetic fuse design guidance human rendering torpedoes useless.			

	Changes to the Plan	Creative	Useful
E	Prepared change in this risk would be to implement a two-stage detonation sequence. The first stage would be to notify the before arming the torpedo.		
F	Subjects testing of systems to identify possible conflicts (navigation, communication, other weapons systems). If conflicts exist, identify possible countermeasures/changes to mitigate impacts.		
G	Changes in the plan were discussed and the following courses of actions based per fuse market research for alternative sources for fuses; produce limited research to determine if other producers with similar characteristics with previous allocations to incorporate to the design of the fuse.		
H	Determine and validate requirements to ensure guidance system is unaffected. Explore options or methods or recording magnetic fuse to understand current compatibility and/or potential upgrades.			

606902

	Risks	Creative	Plausible	Black Swan?
A	Design may be ineffective.			
B	Inadequate follow-on/logistics support.			
C	Unsafe.			
D	Existing infrastructure inadequate.			

	Changes to the Plan	Creative	Useful
E	Consider range of design alternatives to ensure effectiveness.		
F	Conduct competitive assessment of existing inventory.		
G	Develop integrated logistics support plan to include training, maintenance and sustainment, resources, etc.		
H	Develop a comprehensive safety plan to test and field the selected system.		
I	Ensure alternative analysis and design considers impact to the launching platform and its systems and components.			

602001

	Risks	Creative	Plausible	Black Swan?
A	Fuse incompatibility.			
B	Unknown technologies.			
C	Magnetic fuse sensitivity			
D	Magnetic fuse won't want to be working torpedo.			

	Changes to the Plan	Creative	Useful
E	Fly test with technical experts and MOU to test fuse compatibility scenarios. Design a sensitivity select switch depending on operations. Perform extensive T&E. Design a selectable fuse capability with no competition between fuse (prior to launch). Perform integration testing of conduct fuse and MOU1 pass fuse with torpedo controls and use controls. Investigate fuse designs as a separate subsystem and subject to simulated environment and get sensitivity. Testing to determine fuse compatibility possible issues/ resolved.		
F	Testing to ensure no conflict procedures and capabilities. Fly. Perform complete evaluation of CONOPS with new modified torpedo and perform trade-offs with new fuse. Perform testing during tests for MK54 torpedoes to ensure they can operate near TSF input for use of mine torpedo. Perform operations change. Change tactics/procedures that are focused on fuse capabilities. Run this scenarios of CONOPS with operators to identify and define constraints for use of two different fuses (system operating procedures).		
G	Partner with technical experts to integrate to test fuse compatibility scenarios. Design a sensitivity select switch depending on operations. Fuse software (no problems) and sensitivity testing cycle. Perform real/ready data of MK51 fuse within ship system (specifically sub-carrying the liberty). Benchmark the magnetic field of the torpedo environment in testable shipping/handling/transportation/employment/operational environment (as intended) threat engagement. Testing to determine magnetic compatibility with the safety standard 9. Testing to determine magnetic compatibility with the fuse system and torpedo.		
H	Integrate fuse early testing of incoming MK51 detonator. Change the design to an impact type sensor. Partner with industry or independent test experts to conduct engineering characteristics of mag fuse to testing torpedo. Investigate multiple designs.			

606101

	Risks	Creative	Plausible	Black Swan?
A	Unable to function at high speed			
B	Faulty detection blows up wrong ship			
C	Crew rush built with unobtainable material			
D	Blows up before launch			

	Changes to the Plan	Creative	Useful
E	Retrofit sensor to function at high speeds		
F	Install operational controls to ensure no friendly vessels are in vicinity		
G	Modify technology to detect vessel vs. material		
H	Install multiple safety devices to prevent detonation on self – time delay.		

605202

	Risks	Creative	Plausible	Work Swan?
A	Adding the new fuse (weight, etc) could change the operational characteristics of the torpedo.			
B	Fail to either ship in the area			
C	Are the weapon systems using the same charge and battery systems			
D	Proximity to firing ship			

	Changes to the Plan	Creative	Useful
E	Fully fit the original torpedo spec without requirements and test the new torpedo against the original requirements		
F	Planning and coordination with other ships—Tactics, Techniques and Procedures.		
G	Design and testing built-weapon systems for compatibility and reliability. Expand coverage and evaluation.		
H	Tailor sensor to torpedo/ fuse and testing		

603301

	Risks	Creative	Plausible	Black Swan?
A	Going in a wrong line, like in the magnetic field of the target area will need to be analyzed.			
B	Existing weight will diminish the speed of the torpedo.			
C	Premature detonation of torpedo due to sensitivity found in the torpedo/fuse.			
D	If the contact fuse causes electromagnetic transitory issues with the proximity fuse the torpedo could are detonate.			

	Changes to the Plan	Creative	Useful
A	Proof of concept for reduction of product for EE, weight and sensitivity		
B	Review of design, weight and balance analysis.		
C	Review of detection thresholds (backgrounds) to differentiate by ambient/the body (metals, ice life, temperature, water pressure).		
D	Conduct a proof of concept to address potential electromagnetic interference. Conduct synthetic design analysis for potential shielding systems.			

603302

	Risks	Creative	Plausible	Black Swan?
A	Invalid integration—include outright (effect targeting) the known threats			
B	Launch mechanism (assume based mostly on submarine). Weight, speed and depth requirement. Reduced effectiveness and reliability of weapon and unit.			
C	Easily detected—mission effectiveness. Hull not magnetic. Organic, friendly fire.			
D	Cost of integration/research/training.			

	Changes to the Plan	Creative	Useful
E	Recommend other new primary detection mechanism other than magnetic		
F	Expendable boundary adjust or updates, adjust torpedo materials (softer materials)		
G	Full test on contact vessel if magnetic easily defeated (one vessels) confined. Develop a mission plan.		
H	Develop/redesign/leverage existing technologies.		

606801

	Risks	Creative	Plausible	Black Swan?
A	If the performance requirements are not detailed adequate to achieve desired combat effects satisfied with user and tested then the system is unlikely to meet expectations or be fielded/employed.			
B	Integration for two weapon systems (torpedoes and Mantic) is complex and present technical issues in establishing the two systems not knowing if they will be compatible. Additionally you have two PEO offices that need to be involved.			
C	Understanding the threat. Potential enemy capabilities. Cyber/Pseudo targets assessed by magnetic fields, decoys (ready to surface to sea).			
D	Cost & time not identified.			

	Changes to the Plan	Creative	Useful
E	Write a Capabilities Description Document (CDD) detailing the required performance of the system; obtain user concurrence signed by senior that test plan meets requirements listed in CDD; ensure user testing is conducted.		
F	Testing. Testing of system reacting an initial test of the prototypes need to be done critically to check compatibility. If compatible then the two PM offices need to integrate to have one office for this program, and keeping the two programs active for the live fire/ship and Mantic.		
G	Before we go too far down the road, we need to start to define potential recovery capabilities to deliver the re-designed Mantic torpedo with the magnetic fuse. This includes cyber experts and possibly forget, initiated by magnetic fuse to detonate the torpedo prematurely.		
H	Develop and Integrated Master Plan and Integrated Master Schedule for the upgrade in phases. Have daily/bi-weekly/ days to discuss problems. Allow for an IPR (in person) to approaches. Do early risk reduction testing. Sequester functional experts to dedicate/ establish technology. Prototype.		

604801

	Risks	Creative	Plausible	Black Swan?
A	Magnetic/sensor, redesign of sensor to utilize the torpedo magnetic/sonic depth issues			
B	Testing: time-wasted, embarrassed, depth, additional issues resolved, no budget to support, classification			
C	Operation: enemy threats changes, demand, budget, time constraints, resources			
D	Integration: needs for us, future			

	Changes to the Plan	Creative	Useful
E	Testing prior to integration		
F	Redesign to guidance system		
G	Build in additional time for MDA increase budget/services input (if more investment)		
H	Schedule Risk Assessment/Impact deliverable if major performance reduction + additional information/data/issues needed for additional review		

604802

	Risks	Creative	Plausible	Black Swan?
A	Performance: operational probe reasons may be unsatisfactory, requiring re-work or integrative efforts.			
B	Post-unsuccessful Conditions: real environments may have unfavorable results, requiring re-work.			
C	Physical design with new integration: if engineers can't physically integrate the two fuses, redesign may result in delay.			
D	Survivability: if product cannot be reproduced or maintained, it will delay/reduce Operational Capability.			

	Changes to the Plan	Creative	Useful
F	Performance of this Mantic Torpedo: One of the following possible changes to the plan could be to have a productivity performance baseline of the original torpedo to compare it with the new upgrade.		
G	Environmental Conditions: Some of the possible changes to the plan could be to test multiple test locations and extend testing to try to account for as many real-world environments as possible.		
H	Physical Design with Fuse Integration: One of the possible changes to the plan could be to look at the possibility of using a newer or custom torpedo for the integration.		
I	Survivability: One of the possible changes to the plan could be to look at safe-to-store possibilities like isolating the fuse as without a re-design, to allow for easier reproduction.		

604901

	Risks	Creative	Plausible	Black Swan?
A	There is a significant risk due to the size of fuse that they will not integrate to the Mantic torpedo.			
B	The two Mantic magnetic proximity fuse is found to be non-operational/non-functional when integrated in to mantic torpedo.			
C	The guidance system does not work as required. The torpedo misses the intended target.			
D	The spring system may need to be changed based on the anticipated location of the shot.			

	Changes to the Plan	Creative	Useful
E	We utilized a physical mock-up of a "dummy" system early in the development to be sure they integrate seamlessly together. Additionally, a tested survey will be performed for additional fuse options.		
F	Test the magnetic proximity fuse along with the contact fuse on a system integration lab and models platform to confirm operations of both fuses to ensure.		
G	Conduct testing in a systems integration lab to verify functionality of guidance systems. Review factors surveyed based on other integrative systems. Conduct testing for compatibility of fuses/ inserts.		
H	The team needs to conduct a market survey of existing ships to determine the range of vulnerable locations. Look into lessons learned from similar systems.		

604802

	Risks	Creative	Plausible	Black Swan?
A	Operators will now be required to handle and use both torpedo and mines technology. Lack of safety features in the system for handling and use (i.e in).			
B	Magnetic proximity fuse and backup contact existing mode does not perhaps within torpedo envelope.			
C	The target ship have systems that activates these magnetic signatures making the combined fuse ineffective in the magnetic mode.			
D	Software glitches.			

	Changes to the Plan	Creative	Useful
E	The program plan will now include an analysis of both mine and torpedo handling and use to ensure all concerns are addressed. Specific areas of study will include the design interfaces for arming, testing of programmers, and two training areas (ship and for mine).		
F	Allocate budget and development time to packaging magnetic proximity fuse. Develop and execute tests studies to determine potential performance trade-offs to add magnetic fusing capability.		
G	Conduct sensitivity of the magnetic sensor trigger used in a "normal" environment to verify the new combined fuse has the potential to be effective.		
H	Develop a more robust software development and test program to ensure that the new software software are ready as needed.		

The Score Sheet

Expert Panel Scoring Instructions

- Assume the role of Admiral Tyrosine.
- You will be scoring 20 sets of deliverables.
- Each deliverable has two parts– "Risk List" and "Changes to the Plan"
- The Risk List has four risks. You will be scoring each risk along two dimensions– Creative and Plausible. Use a 0 to 100 grading scale. You will also mark whether or not the each of the four risks match the four Black Swans.
- The Changes to the Plan has four changes. You will be scoring each change along two dimensions– Creative and Useful. Use a 0 to 100 grading scale.
- You do not have to justify your scores. A grade of 70 or greater should represent "quality" score.

Definitions

Creative- having or showing imagination and artistic or intellectual inventiveness.

Plausible- seemingly true, acceptable.

Useful- having practical utility.

"Black Swans": Risks that the real world/historical team missed, but in hindsight should have caught.

- Tested torpedoes without explosives in order to conserve limited test articles.
- Tested torpedoes without hitting objects to recover and reuse the torpedoes for more testing.
- The torpedoes ran too deep to sense the magnetic field of the enemy ship.
- Integration of the magnetic fuse interfered with the contact detonation mechanism causing torpedoes to bounce-off the ships without detonating when in the "contact fuse" mode.

Step-8 Of the 20 sets of deliverables, what sets do you consider to be the three you would most like to receive if you were the Admiral?

Session#:_____

Session#:_____

Session#:_____

Step-9 May I include your DAU biography in my dissertation?

YES NO (Circle Your Choice)

- Step-10
  - Print Your Name_____
  - Return this sheet to Dave Gallop by 15 March 2016
    - Scott Hall Room 116

      or
    - Mail to Dave Gallop

      6735 Surbiton Dr.

      Clifton, VA 20124-1928

THANK YOU FOR YOUR CONTRIBUTION TO SCIENCE