2008

Defining Risk Assessment Confidence Levels For Use In Project Management Communications

Gary Johnson
University of Central Florida

DEFINING RISK ASSESSMENT CONFIDENCE LEVELS FOR USE IN PROJECT MANAGEMENT COMMUNICATIONS

by

GARY L. JOHNSON
B.S. University of Central Florida, 1979
M.S. Florida Institute of Technology, 1988
M.S. University of Central Florida, 1992

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Industrial Engineering and Management Systems
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Fall Term
2008

Major Professor: Julia Pet-Armacost

ABSTRACT

A review of the literature regarding risk management and effective risk communications identified that very few researchers have addressed risk assessment confidence levels when using risk scoring methods. The focus of this research is to develop a definition of risk assessment confidence levels for use in internal project management communications and to evaluate its usefulness. This research defines risk assessment confidence level as "The degree of certainty that the likelihood or consequence score (assigned by the risk assessor) reflects reality." A specific level of confidence is defined based on the types of analyses that were conducted to determine the risk score.

A survey method was used to obtain data from a representative sample of risk assessment professionals from industry and academia to measure their opinion on the usefulness of the defined risk assessment confidence levels. The survey consisted of seven questions related to usefulness—four questions addressed the importance of stating confidence levels in risk assessments and three addressed the usableness of the proposed confidence level. Data were collected on the role and experience level of each of the respondents and the survey also included a comment section to obtain additional feedback.

The survey generated 364 respondents representing a broad variety of roles associated with decision making and risk management with experience levels from fairly new to experienced risk assessors. The survey data were analyzed by calculating the proportion of respondents who gave negative, neutral and positive responses to the survey questions. An examination of the roles of the survey respondents indicated that no single group was dominant. A non-parametric Kruskal-Wallis test generally failed to reject the hypothesis that the means of the survey response distributions were identical. There was one exception which indicated that

there are differences based on role and by inspection of the responses, it appears that decision makers, academics, and others more strongly support the need for confidence level information to reduce the difficulty in making risk based decisions in projects.

The survey responses at a confidence level of 95% have a range of errors from 3.84 to 4.97%. Based on the results of the survey, 77 – 83% of those surveyed indicated agreement that knowing the confidence the assessors have in their assessment is important and would improve a management decision. The survey showed that 60 – 86% of the respondents agreed that the confidence levels and their definitions as presented in the survey were usable. The question with the lowest agreement (60%) was related to the way in which the individual levels were defined. The ad-hoc comments provided in the survey were divided into eleven groups based on similarity of the subject of the comment and then examined for common themes. These added additional insight into the results and useful information for future research efforts.

This research validates that the use of risk assessment confidence levels is considered to be useful in project risk management. The research also identified several potential areas for future work, including determining the appropriate number of confidence levels that should be defined, refining the definition of the individual confidence level definitions, examining historical perspectives of whether the risk assessments were accurate, examining the concept of shiftability of risk assessments, further research on communication of variability of risk assessments, and research into the usefulness of risk matrices.

This research is dedicated to my wife Lisa and son Michael.

ACKNOWLEDGMENTS

My love and appreciation goes out to my parents and in-laws for being patient as my time was focused on my research.

My sincere gratitude goes out to my co-workers at the National Aeronautics and Space Administration at both the Johnson Space Center and Kennedy Space Center for continuing to support and encourage me as I carried out my research. My sincere appreciation goes to all of the professors at the University of Central Florida who have mentored me and encouraged me throughout my academic career.

Special gratitude to Dr. Jose Nunez, Mr. Jim Card, Mr. Bill Cleek, Ms. Sharon Thomas, Ms. Susan Brand, Ms. Daisy Mueller, Mr. Mike Lutomski and Mr. Alexis Hongamen for supporting the focus group.

A special thank you goes out to Dr. Pet-Armacost for the time and support she gave me. Thanks for the support and advice of all my committee members, Dr. Armacost, Dr. Hoekstra, Dr. Perera and Dr. Reilly III.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS/ABBREVIATIONS

ABS	American Bureau of Shipping
ANOVA	Analysis of Variance
EPA	Environmental Protection Agency
FMEA	Failure Mode and Effects Analysis
IRB	Institutional Review Board
NASA	National Aeronautics and Space Administration
NESC	NASA Engineering Safety Center
NRC	National Research Council
PDF	Probability Density Function
PMBOK	Project Management Book of Knowledge
PMI	Project Management Institute
QRA	Quantitative Risk Assessment

CHAPTER 1: OVERVIEW OF RESEARCH

The focus of this research is to develop a definition of risk assessment confidence levels for use in internal project management communications and to evaluate its usefulness. The specific objectives of this dissertation are:

1. Validate the need for defining risk assessment confidence levels for use in internal project management communications
2. Define confidence levels
3. Validate that the use of risk assessment confidence levels is considered to be useful in project risk management

This research defines confidence level as the degree of certainty that the likelihood or consequence score (assigned by the risk assessor) reflects reality.

A critical aspect of any project is risk management. "Risk management is a critical part of project management as unmanaged or unmitigated risks are one the primary causes of project failure." (Lyons and Skitmore 2004 p. 51) "The reality is that projects are subject to the shifting forces and constant changes due to the external factors, changing objectives and poor methods for project realization. So the process of risk and uncertainty management must be continuous, holistic and conducted in real time to be of any value to project managers." (Jaafari 2001 p. 93) The National Research Council notes that "Managing risk is one of an owner's most important functions in making any major project successful." (National Research Council 2005 p. 8) Hillson discussed this during his proposal to use a work breakdown structure as a method to communicate risk, "Successful and effective risk management requires a clear understanding of the risks faced by the project and business. This involved more than simply

listing identified risks and characterizing them by their probability of occurrence and impact on objectives. The large amount of risk data produced during the risk process must be structured to aid its comprehension and interpretation, and to allow it to be used as a basis for action." (Hillson 2003 p.95)

The focus of this research is to develop a definition of risk assessment confidence levels for use in internal project management communications and to evaluate its usefulness. In order to understand the research it is important to understand the foundations of risk and risk assessment.

What is risk? There are many definitions found in the literature:

- "Risk can be defined as the exposure to the probability that an event with adverse consequences might occur." (Ben-David and Raz 2001 p.14)
- "A potential issue that can be described in terms of likelihood of occurrence and the consequence after it occurs." (Bradt 2004 p. 2)
- "Risk is the potential for realization of unwanted negative consequences of an event" (Rowe 1977 p. 24)
- "Project risk may be defined simply as the possibility of an unintended future event with potential undesirable consequences" (National Research Council 2005 p. 13)

Project risk is important to understand and manage. "To be successful, the organization should be committed to addressing the management of risk proactively and consistently throughout the project." (PMI 2004 p. 240) From the definitions in literature there are common themes such as that risk is bad (adverse, issue, unwanted) and that it has an attribute of probability (likelihood, potential). Once a risk has been identified, risk management consists of the steps taken to either reduce risk or to mitigate risk to an acceptable level. The PMBOK (Project Management Book of Knowledge) defines project risk management as consisting of the

following six processes: (1) risk management planning; (2) risk identification; (3) qualitative risk analysis; (4) quantitative risk analysis; (5) risk response planning and (6) risk monitoring and control. (PMI 2004 p. 237) Haimes et al. (2002) provides a methodological approach or framework for the project team to follow in order to perform all of the risk management steps of identification, prioritization, assessment and management. While Haimes' approach is specifically applied to large-scale systems, it is reasonable to use such a framework on smaller scale projects.

While risk can be identified, quantified, and mitigated (controlled), none of these steps can be effective without effective **communication**. Is effective risk communication important? The National Research Council recognized that "… even though good risk communication cannot always be expected to improve a situation, poor risk communication will nearly always make it worse." (National Research Council 1989 p. 3) Ibrekk and Morgan (1987) summarized the need for effective communication when they stated that "Effective risk communication is important. Without it the most carefully performed risk analysis may be useless; the most thoughtfully devised risk management strategy may be ineffective; scarce resources and attention may be wasted." (Ibrekk and Morgan 1987 p. 519) Knuth and Connelly (1998) provide a discussion on factors that affect the receipt of information.

Table 1 provides general definitions found in risk management literature that will be used throughout this document.

Table 1 Definitions

Term	Definition
Decision	A process that starts with a pool of information from which the situation is defined, stakes are assessed, available options are considered and predictions are made. It ends with the selection of an option based on selected criteria or heuristics. (Kanner 2005 p. 314)
Risk	Risk is characterized by the combination of the probability that a program or project will experience an undesired event (some examples include a cost overrun, schedule slippage, safety mishap, health problem, malicious activities, environmental impact, failure to achieve a needed scientific or technological breakthrough or mission success criteria) and the consequences, impact, or severity of the undesired event, were it to occur. (NASA NPR8000.4 2004 p. 5)
Risk Management	Risk Management (RM) is a process wherein the program/project team is responsible for identifying, analyzing, planning, tracking, controlling, and communicating effectively the risks (and the steps being taken to handle them) both within the team and with management and stakeholders. (NASA NPR8000.4 2004 p. 5)

Chapter 2 presents a literature review to (1) to provide the reader an overview of the existing literature as it relates to risk management and effective risk communications and (2) provide the foundation to demonstrate that there exists a gap in the literature addressing risk assessment confidence levels when using risk scoring methods.

Chapter 3 provides the background for the research, the research design and the methodology for its validation.

Chapter 4 provides the results of the research validation and proposes potential areas for future research.

Chapter 5 summarizes the conclusions drawn by this dissertation.

CHAPTER 2: LITERATURE REVIEW

The focus of this research is to develop a definition of risk assessment confidence levels for use in internal project management communications and to evaluate its usefulness. The objectives of this chapter are to (1) to provide the reader an overview of the existing literature as it relates to risk management and effective risk communications and (2) provide the foundation to demonstrate that there exists a gap in the literature addressing risk assessment confidence levels when using risk scoring methods.

Project risk exists in all projects. "All projects involve some amount of risk, resulting from them being temporary endeavors aimed at achieving some unique set of predetermined time, cost and performance objectives." (Ben-David and Raz 2001 p. 14) Project risk is important to understand and manage. "To be successful, the organization should be committed to addressing the management of risk proactively and consistently throughout the project." (PMI 2004 p. 240) There are many definitions of risk found within the literature. "A Guide to the Project Management Body of Knowledge PMBOK® Guide" (PMI 2004) defines project risk "… an uncertain event or condition that, if it occurs, has a positive or negative effect on at least one project objective, such as time, cost, scope, or quality...." (PMI 2004 p. 238) The National Research Council states that "Project risk can be defined simply as the possibility of an unintended future event with potential undesirable consequences." (National Research Council 2005 p. 13) There is commonality in the definitions found in the literature in that there is an uncertain or unintended event that if it occurs, has an effect on a project attribute such as cost, scope, quality, safety. Some definitions consider risk as having undesirable consequences (National Research Council 2005) while others consider that risk can have either positive or

negative consequences (PMI 2004). Many useful definitions of risk are found in the literature such as Kaplan and Garrick (1981), PMI (2004), Gifford et al. (1979) and Moschandreas and Karuchit (2005). This research on the use of confidence levels for internal project management communications could be used for risk defined as having either positive or negative consequences. The PMBOK emphasized the importance of managing risk when it states "To be successful, the organization should be committed to addressing the management of risk proactively and consistently throughout the project." (PMI 2004 p. 240) The act of addressing risk throughout the project is risk assessment. The literature provides insights into what risk assessment is and why it is important.

Anderson and Narasimhan provide a thorough discussion on risk assessment. They state that "… "risk assessment" refers to the explicit identification of the difficulties that are likely to be encountered in achieving "successful implementation" of a project." (Anderson and Narasimhan 1979 p. 513) The term "difficulties" as used here means the differences between what was envisioned as the nominal implementation of a project or those things that occur that may be barriers to achieving the desired implementation. Anderson and Narasimhan note that "Project risk assessment assists the practitioner in prioritizing the search for ways to counteract the deterrents that may exist by focusing his attention on those factors which possess the greatest opportunity for changes and thus enhancing chance of successful implementation." (Anderson and Narasimhan 1979 p. 513) According to Anderson and Narasimhan (1979), risk assessment is needed because it is essential for the managing a successful project, formulizes the evaluation and communication of information concerning the project environment, and helps with prioritization during the search for those changes that enhance the chance of project success. Cross and Ballisio noted the importance of complete assessment "If a risk assessment is not

comprehensive or complete, then the validity of the results is questionable." (Cross and Ballesio 2003 p. 3) Kangari and Riggs (1989) examined risk assessment as it applied to the construction industry. While their work addresses the construction industry, the basics of risk assessment documented in their work is applicable across all project disciplines.

There are various methodologies for communication of the risk attributes such as likelihood and consequence that can be found in the existing literature. Cox et al. proceed to state that qualitative risk systems "… simplify risk assessments by reducing the required inputs and calculations to a manageable set of judgments, while making the rating logic transparent and easy to apply. They usually require only a few qualitative judgments as inputs, together with supporting reasoning and documentation, and they usually produce simple categorizations of risk as outputs that can be communicated relatively easily to policymakers and stakeholders." (Cox et al. 2005 p. 651) The literature uses terms like simple and manageable which would indicate that the development of a complicated system to communicate confidence levels would be quickly rejected by its intended users (project/program managers, decision makers). Risk assessment methodologies can be found throughout the literature.

Macgill and Siu present a set of questions that can be used to define a risk state. These questions are:

1. What are the expected effects?
2. Is there any doubt?
3. Are the effects acceptable?
4. Does everyone agree?

(Macgill and Siu 2005 p, 1111)

Questions 1 and 2 will require the experts to define the magnitude/impact of the risk, along with any uncertainty of that assessment. Question 3 will require the social and professional judgments of the project team members to determine the acceptability of the situation. Question 4 really concerns the degree of compatibility achieved to Question 3. Questions 1 and 2 can be defined in various terms including technical, cost, schedule, and safety. Not repeated here, but similar lists are also found elsewhere in the literature such as Kaplan and Garrick (1981) and Cross and Ballesio (2003).

Results of project risk assessment are nominally documented in categories of technical, cost and schedule risk. This was documented by Thamhain and Wilemon who surveyed 500 engineering professionals and found that more than 90% listed technical, cost and schedule as the most important factors to measure team performance. (Thamhain and Wilemon 1987) Technical risk may be defined as the risk of the project or task not meeting the customer technical or performance requirements. Cost risk may be defined as the risk of exceeding the budget at a task or project level. Schedule risk is the risk of not meeting critical project schedule milestones. Within a project there maybe one or more set of definitions for risk scores. Table 2 is an example of the risk score matrix used to score the impact, magnitude or consequence of a risk:

Table 2 Magnitude index

Magnitude	Technical risk R_t	Cost risk R_c	Schedule risk R_s
1 (Low)	No measurable impact on the systems' ability to meet requirements	No measurable impact on the budget, some transfer funds within the project may be required	Minimal impact, slight changes compensated by available program slack
3 (Minor)	Minor reduction in some requirements, the same system architectures must be selected for some components	Cost estimates exceed budget by up to 5%	Additional activities required, able to meet miss minor milestone
7 (Significant)	Significant reductions in requirements, alternative subsystem architectures must be selected	Cost estimates exceed budget by up to 50%	Program critical path affected
9 (High)	Unable to meet system requirements, alternative system architecture must be developed	Cost estimates exceed budget by greater than 50%	Cannot achieve key program milestone

(Componation et al. 2001 p. 28)

Table 3 is an example of the risk score matrix used to score the probability of a risk.

Table 3 Probability index

Probability	Description
1 (Very low probability)	Uncertainty can be mitigated utilizing standard practices
3 (Low probability)	Uncertainty can be mitigated with minimal oversight
5 (Medium probability)	Uncertainty can be mitigated but workarounds will probably be required
7 (High probability)	Uncertainty cannot be mitigated with the existing approach, but alternative approaches are available which may mitigate the risk
9 (Very high probability)	Uncertainty cannot be mitigated, no known workarounds are available to mitigate the risk

(Componation et al. 2001 p. 29)

It should be noted that the term "uncertainty" as used in Table 3 is different than used in the context of confidence levels. "Uncertainty" is also used to describe the unknown as far as the probability or likelihood of an event occurring or the unknown as far as the consequences that may result should the event occur. This research into the use of confidence levels uses the term "certainty" but in the context that it is the degree of certainty that the likelihood or consequence score as assigned by the risk assessor reflects reality. Other examples of risk communications can be found in the literature such as in Greenberg and Cramer (1991).

Table 2 and Table 3 demonstrate a simple methodology. As shown in Figure(s) 1, 2 and 3, Macgill and Siu present a representation using benchmarks of the risk state as defined by their questions which are:

1. What are the expected effects?
2. Is there any doubt?
3. Are the effects acceptable?
4. Does everyone agree?

(Macgill and Siu 2005 p. 1111)

Figure 1 summarizes these benchmark answers, including qualitative labels and alternate symbolic, numeric or color-coded representations.

Four core dimensions and associated benchmarks				Alternative representations of benchmarks			
Effects (E)	Scientific doubt (D)	Accept-ability	Social con-flict (C)	Qualitative label	Symbol	Colour	Number
Low	Certainty[a]	Yes	Consensus	Alright	□	green	0
Mixed[b]	Puzzlement/ opinion[c]	Don't know	Debate	Debatable	⊡	amber	−1
Detrimental	Uncertainty[d]	No	Conflict	Critical	■	red	−2
Unknown[e]	Ignorance[f]	Don't care	Unengaged	Super critical	■	black	−3

[a] We know: there is no doubt.
[b] There may be high detriment on ecological aspects, and low detriment, or even high benefit on financial aspects, leading to a mixed picture overall.
[c] We are puzzled.
[d] We don't know.
[e] Not considered.
[f] We don't know what we don't know.

Figure 1 Four Core Dimensions and Associated Benchmarks
(Macgill and Siu 2005 p. 1114)

Macgill and Siu then took a subset of the four core questions and created a 3-dimensional representation as shown in Figure 2.

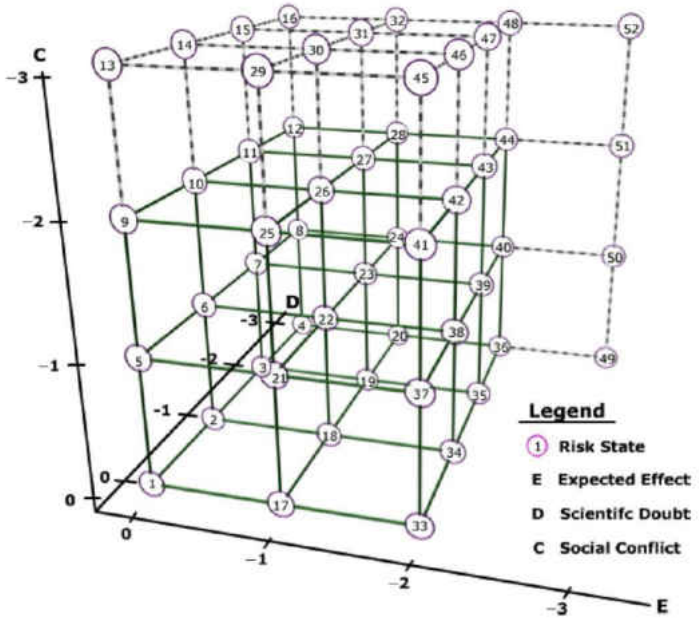


Figure 2 3-Dimensional View of Risk States
(Macgill and Siu 2005 p. 1114)

The visual representation in Figure 2 has been restricted by the ability to easily show 3-dimensions on a 2-dimensional page. Figure 3 shows examples of how this might be represented for risk states 1 and 10.

Risk state	Co-ordinates	Physical construct		Social construct	
		Expected effects (E)	Scientific doubt (D)	Social conflict (C)	Acceptability[a]
1	E0, D0, C0	☐	☐	☐	☐ ☐ ☐ ☐ ☐
10	E0, D-1, C-2	☐	⊡	■	☐ ⊡ ■ ■ ■
10	E0, D-1, C-2	☐	⊡	■	☐ ☐ ■ ☐ ■
10	E0, D-1, C-2	☐	⊡	■	⊡ ■ ☐ ⊡ ☐

There are many further combinations of constituency positions for risk state 10.

[a] The five constituencies shown here are suggested to be, from left to right: expert scientists, regulators, stakeholders, the general public, and the mass media.

Figure 3 Examples of Benchmark Risk States
(Macgill and Siu 2005 p. 1115)

From the literature it can be observed that the communication of the risk state can be simple or complicated. The focus of this research is to develop a definition of risk assessment confidence levels for use in internal project management communications and to evaluate its usefulness. The confidence level is defined as the degree of certainty that the likelihood or consequence score which has been assigned by the risk assessor reflects reality.

The Australian Standard on Risk Management recognized that the uncertainty of the risk assessment is important when it states that "Whenever possible, the confidence placed on estimates of levels of risk should be included." (Joint Technical Committee OB/7 1999b p. 13) Aven and Pitblado (1998) directly address uncertainty in the risk message and its importance to the decision maker, but they do not address how to implement that communication. "Now, following a classical interpretation of risk, the message is partly 'disturbed' by uncertainties in estimates. These uncertainties are normally very large and represent a serious weakness of the analyses, at least seen from the decision-makers point of view. While absolute errors can be a

problem in some cases, most decision-making, is based on relative interpretation of risk results, with the difference in risk between various options compared with their costs. Errors in this relative mode are generally believed to be very much less than the absolute errors." (Aven and Pitblado 1998 p. 25) Rose addresses confidence level and its relationship to data quality and quantity in what is called a "Chance adequacy matrix". This is described as "… a matrix widely used among many companies, whose origin is not known to the writer. Assuming that a geological model, or concept, is recognized, it compares (1) quantity and quality of information against (2) what the information is signifying with respect to at least minimal adequacy about the particular geologic chance factor." (Rose 2001 p. 36) Rose relates data quality to the matrix. "… in order for us to render judgments of high confidence, either encouraging or discouraging, we require considerable data of good quality. Conversely, sparse or poor-quality data frequently allow only intermediate confidence statements… geoscientists should recognize that the absence of information does not , by itself, imply a negative outcome – only that there are no data." (Rose 2001 p. 36) The Chance adequacy matrix is shown in Figure 4.

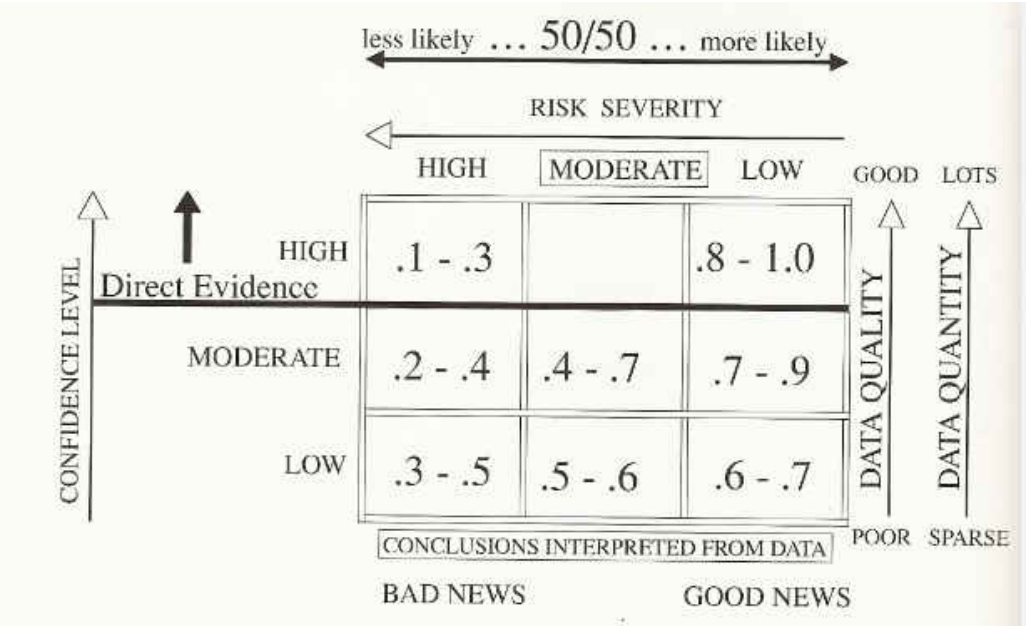


Figure 4 Chance Adequacy Matrix
(Rose 2001 p. 38)

The American Bureau of Shipping (ABS) process for risk assessment shown in Figure 5 shows the relationship between risk assessment and the reduction of uncertainty.

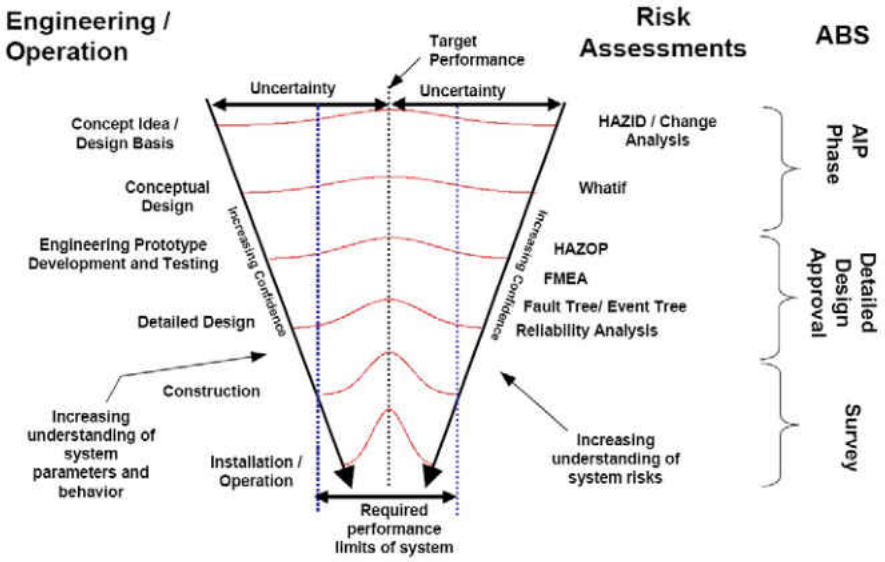


Figure 5 Concept to Operation Process
(Verzbolovskis 2004 p. 51)

What can be observed from the literature is that with an increase in the quantity of data, assuming that the data are quality data, there can be a reduction in uncertainty which increases the confidence in the analysis. Figure 6 shows the relationship between the layers of analysis that are done, the reduction of uncertainty and the resources invested. "Each layer of analysis provides more detailed and certain loss exposure information, but the resources invested in the analysis increase at each level. The filtering effect of each layer allows only key issues to move into the next more detailed level of analysis. At any point, sufficient information for decision making may be developed, and the analysis may end at that level. (All levels of analysis will not be performed for every issue that arises). In fact, most issues will probably be resolved through risk/reliability screening analyses or broadly focused, detailed analyses." (American Bureau of Shipping 2000 p. 50)

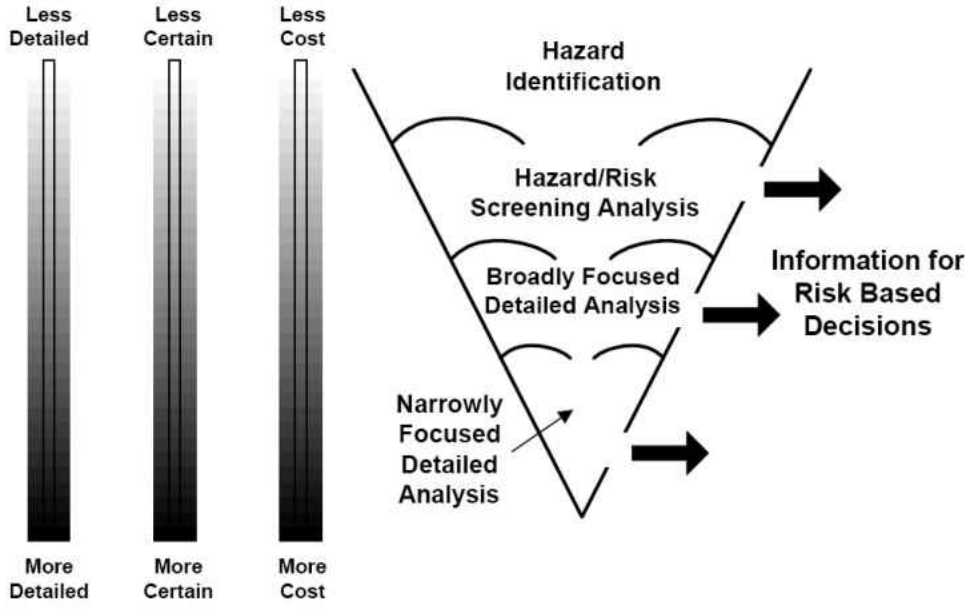


Figure 6 Risk Analysis Levels
(American Bureau of Shipping 2000 p. 50)

The literature contains different approaches to risk assessment and risk management. As an example, Basu and Lee (2006) proposed a two-level approach consisting of a Level 1 simplified approach and Level 2 which is an advanced approach. Level 1 is a semi-quantitative risk assessment and ranks the criticality of the structure, components and assemblies. The advantage of limited assessment in Level 1 is the use of resources to focus on the structure that needs attention. Level 2 uses more rigorous quantitative assessment on the areas identified in Level 1 as needing additional attention.

It is important for the decision maker to understand the confidence in the assessment in order for them to make a decision on whether to expend more resources to increase the confidence level. Long and Fishhoff made the observation that "Barring systematic bias, the uncertainty about a risk will decrease as a function of the resources invested in learning about it." (Long and Fischhoff 2000 p. 341) Within the literature, confidence level is also called "uncertainty".

Thompson (2002) noted that "Uncertainty arises from our lack of perfect knowledge, and it may be related to the model used to characterize the risk, the parameters used to provide values for the model, or both. In some cases, we can reduce uncertainty by obtaining better information, but this may not always be possible. Uncertainty implies that we might make a nonoptimal choice because we may expect one outcome but something quite different may actually occur." (Thompson 2002 p. 648) Thompson and Bloom noted that "After making the decision and preparing their positions, risk managers also felt strongly about having risk assessors inform them about uncertainties and vulnerability that affect their case." (Thompson and Bloom 2000 p. 337) In addition, they conclude that "Overall, the risk managers indicated their interests in hearing both qualitative and quantitative information about multiple aspects of

the decision. This work suggests that to inform risk managers better, risk assessors must also appreciate the broader context of the decisions to be made, and they must convey how uncertainties and weaknesses in the assessment may influence stakeholder perceptions and the effectiveness of different risk management options." (Thompson and Bloom 2000 p. 347) The literature supports this research in that the communication of the confidence level or uncertainty associated with the risk assessment is needed by the decision makers in order to choose the correct option for mitigating the risk. Finkel (1994) and Banon (1994) provide additional insight into uncertainty. A view on the need for uncertainty as an essential element to lower risk is given by Emblemsvåg and Kjølstad (2002) who point out that "From the Law of Incompatibility we understand that there are limits to how precise decision support both can and should be (to avoid deception), due to the inherent uncertainty caused by complexity. By increasing the uncertainty in analyses and other decision support material to better reflect the true and inherent uncertainty will lower the actual risk." (Emblemsvag and Kjolstad 2002 p. 844) The literature does provide insight into some of the tools that are available to provide both qualitative and quantitative information which can be used to increase the confidence of the risk assessor or decrease the uncertainty of the assessment. Figure 7 shows some of the risk assessment tools as described by the National Research Council.

Tool	Characteristics
Two-dimensional impact/ probability	Qualitative, simple to use and most frequently used, can be expanded to three or more dimensions, and can be combined with FMEA
Pareto diagram	Simple qualitative method for prioritizing risk elements
Failure modes and effects analysis (FMEA)	Qualitative, used for initial screening only, effective in a team environment
Project Definition Rating Index	Qualitative, used in front-end project planning, effective in a team environment
Multivariate statistical model	Quantitative, requires historical database
Event tree	Quantitative, rarely used for risk analysis
System dynamics model	Both qualitative and quantitative, rarely used but effective, requires skilled modelers
Sensitivity analysis	Quantitative, useful regardless of which other process used, useful in absence of hard data
Project simulation	Both qualitative and quantitative, useful for team building, expensive to implement
Stochastic simulation	Quantitative, frequently used, often misused, so limitations must be made clear
Additive model	Quantitative, can be adjusted as project progresses

Figure 7 Risk Assessment Tools
(National Research Council 2005 p. 39)

The Australian Standard on Risk Management provides additional information on tools that can be used to increase the confidence the risk assessor has in the assessment:

a) "Past records;

b) Relevant experience;

c) Industry practice and experience;

d) Relevant published literature;

e) Test marketing and market research;

f) Experiments and prototypes;

g) Economic, engineering or other models;

h) Specialist and expert judgments."

(Joint Technical Committee OB/7 1999b p. 13)

Techniques to be considered are:

i. "structured interviews with experts in the area of interest;

ii. use of multi-disciplinary groups of experts;

iii. individual evaluations using questionnaires;

iv. use of computer and other modeling; and

v. use of fault trees and event trees."

(Joint Technical Committee OB/7 1999b p. 13)

Both the lists provided by the National Research Council and in the Australian Standard provided examples of some of the attributes that can define what confidence levels are. There are other examples such as described by Grose (1987), Apostolakis (2004), Carbone and Tippett (2004), Dey (2002) and Moynihan et al. (2002).

The following set of questions examine the added value that the analysis will provide to the decision process based on guidelines issued by the Environmental Protection Agency (EPA) (2004):

- Will the quantitative analysis improve the risk assessment?
- What are the major sources of uncertainty?
- Are there time and resources for a complex analysis?
- Does this project warrant this level of effort?
- Will a quantitative estimate of uncertainty improve the decision? How will the uncertainty analysis affect the regulatory decision?
- How available are the skills and experience needed to perform the analysis?
- Have the weaknesses and strengths of the methods involved been evaluated?

- How will the uncertainty analysis be communicated to the public and decision makers?

When looking at which tools to use, consideration should be given as identified by Pfleeger (2000) are:

- Sample size – "… error and unreliability are underestimated when sample size is small. Psychological research on risk perception tells us that "both scientists and lay people may underestimate the error and unreliability in small samples of data, particularly when the results are consistent with preconceived, emotion-based beliefs" (Whittemore 1983 p. 28)" (Pfleeger 2000 p. 270)
- Study design – "A more important concern is our reluctance to examine the study designs that lead to quantitative risk assessments. …In the software realm, we are rarely asked to present the studies that support our risk assessments, let alone their underlying designs and data sets. And because we do not always publish or make available the complete set of study documentation, our studies are almost impossible to evaluate or replicate." (Pfleeger 2000 p. 270)
    - Relevance and quality of data – "One aspect of study design quality that is often ignored is the relevance and quality of the data being studied." (Pfleeger 2000 p. 270)
    - Use of data from similar studies – "We often make similar mistakes, using data from other studies instead of generating our own. … The point is not that reusing data is bad. Rather, the point is that you have stronger evidence if you observe a phenomenon *n* different times than if you

observe it *m* times (for m less than n) and then reanalyze data from the *m* cases until you build up *n* of them." (Pfleeger 2000 p. 270)

- Objectivity – "Independent confirmation of the results of a technique will increase the likelihood that the technique will perform well the next time. That is, objective assessment lowers the risk of failure." (Pfleeger 2000 p. 270)
- Scale – "… we should always ask ourselves how risks observed or assessed on small project scale up to medium or large projects." (Pfleeger 2000 p. 271)

Andrews et al. reported on the work conducted by the New Jersey Department of Environmental Protection starting in 1998. Of interest to this research was that in surveying environmental threats, the assessors were asked for their "level of confidence" in their assessment of the impacts of the threats. Rated using a low, medium or high confidence level, these were defined as:

- "Low Uncertainty: Impact estimate is quantitative and well documented. Scientific consensus exists on estimating approach… It is highly probable that the reported score is correct (68% or better, i.e., one standard deviation).
- Medium Uncertainty: Some documentation exists. A literature relying on this estimating approach exists... We are confident that, if scores above are wrong, they are, on balance, only off by one ordinal (e.g., High vs. Medium). There is at least a 50% probability that the reported score is correct (even odds or better).
- High Uncertainty: Impact estimated is qualitative and poorly documented. No scientific consensus exists on estimating approach… Scores above are, on

> balance, quite arbitrary, and could be off by more than one (High vs. Low). It is not more probable that the reported score is correct than that a lower or higher score is correct, so the probability that the reported score is correct is about 34%." (Andrews et al. 2004 p. 1329)

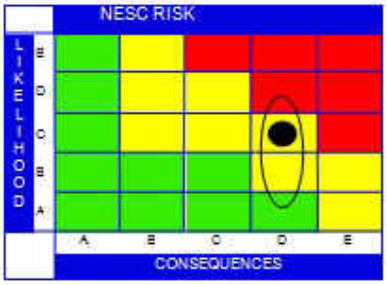
Of interest is that the case noted by Andrews et al. has similarities and differences from the research in this dissertation. The work of Andrews et al., has a similarity in that it documents the use of "low," "medium" and "high" definitions for levels of confidence that were used in a survey of environmental threats given by the New Jersey Department of Environmental Protection. (Andrews et al. 2004) What Andrews et al., do not provide is the research basis for demonstrating that the use of these levels is deemed useful by the decision makers. In addition, the work by Andrews et al., was limited to a small population having concerns with New Jersey Environmental Protection, again for both internal and external communications. There are other approaches available in the literature such as that proposed by Moses and Malone Jr. (2005).

Moses and Malone Jr. (2005) presented the results of a study conducted for NASA's Engineering and Safety Center that focused on the development of a risk assessment matrix. Within the presentation and supplemental materials handed out at the 2005 Project Management Challenge Moses and Malone presented the Figure 8 as "proposed enhancements":

Proposed Enhancements - Applicable to Traditional Risk Matrix as well as NARMA

Capture and Communicate Uncertainty

NESC has expressed their concern that the current 5 x 5 matrix structure does not completely communicate the uncertainty of the matrix placement of an assessed issue. It is recommended that methods for assessing, incorporating and communicating this uncertainty be expanded and taken to the next level. To the right is one proposed method of capturing and communicating uncertainty.

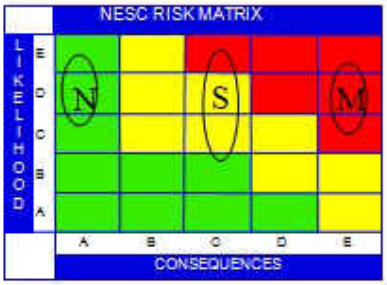


Uncertainty Level	Uncertainty % of Estimate	Score	Criteria for System	Score	Criteria for Assessor
5	200	E	No data available	E	No experience
4	100	D	system remotely similar	D	Little Experience
3	50	C	Little Data Similar System	C	Some Experience
2	25	B	Data Similar Equipment	B	Expert Similar system
1	10	A	Data on Exact System	A	Expert in System

Establishing a predetermined level of uncertainty as a starting point for capturing the uncertainty of an initial assessment is proposed as one tool by which an individual can begin the task of developing uncertainty bounds

The current method of communicating the consequence value of the issue within the 5 x 5 matrix involves selecting the attribute (SHE, MS or NV) with the highest value. A more mature method of communicating information about the issue would involve using descriptors for each attribute (S for SHE, M for MS and N for NV) and placing them in their respective consequence area in the 5 x 5 matrix.

Notice in the figure below that the ellipses enclosing the letters N, S, and M are of different sizes. This is a graphical representation of the fact that there are different levels of uncertainty bounding the values for N, S, and M.



The National Aeronautics and Space Administration, Langley Research Center, Engineering and Safety Center under Award NNL04AA29T provided funding for this study. Any opinions, findings, or conclusions expressed here are those of the authors and do not necessarily reflect the views of the Administration.

Figure 8 Proposed Enhancements to Risk Matrix
(Malone and Moses 2005)

In the work of Moses and Malone (2005), the objective is to have the risk assessor identify the range of values that they believe the likelihood and consequence may take. This approach does not address the confidence that the assessor has in their assessment as they could have a low confidence in the entire range that is provided to the decision maker, or a low

confidence in one end of the indicated range with a higher confidence in the upper end of the range etc. My view of the approach given by Moses and Malone (2005) is that it is more an indicator of the variability of the assessment scores for likelihood and consequence than an indicator of how confident the assessor is in their assessment. Moses and Malone (2005) did propose for additional investigation "Uncertainty Levels" as shown in Figure 9. Again, this seems more at defining levels of variability versus the confidence the assessor has in the assessment they are providing to the decision maker.

Uncertainty Level	Uncertainty % of Estimate	Score	Criteria for System	Score	Criteria for Assessor
5	200	E	No data available	E	No experience
4	100	D	system remotely similar	D	Little Experience
3	50	C	Little Data Similar System	C	Some Experience
2	25	B	Data Similar Equipment	B	Expert Similar system
1	10	A	Data on Exact System	A	Expert in System

Figure 9 Uncertainty Levels
(Malone and Moses 2005 p. 15)

The objective of all of these approaches is to communicate risk to the decision maker.

Enrico Zio provided the following definition of risk communication "The objective of risk communication is to provide people with a basis for making an informed decision". (Zio 1997 p. 32) Kahlor et al. stated that "The goal of most risk information campaigns is to help people understand risks, make wise choices, and develop stable and beneficial changes in their risk-related behaviors." (Kahlor et al. 2003 p. 355) Bernstein relates a simple story that summarizes the affect of communication and uncertainty on decision making:

> "… we sometimes have too little information to apply the laws of probability
>
> Once at a professional investment conference, a friend passed me a note that read as follows:
>
> The information you have is not the information you want.
>
> The information you want is not the information you need.
>
> The information you need is not the information you can obtain.
>
> The information you can obtain costs more than you want to pay.
>
> We can assemble big pieces of information and little pieces, but we can never get all the pieces together. We never know for sure how good our sample is. That uncertainty is what makes arriving at judgments so difficult and acting on them so risky." (Bernstein 1996 p. 202)

The literature is an abundant source of research on improving communication as can be found in Parrott et al. (2005), Thompson and Bloom (2000), Zmud (1978), Leiss (1996), Viscusi and Zeckhauser (1996), Covello et al. (1986), Atman et al. (1994), Aven and Vinnem (2007), Bostrom et al. (1994), Edwards et al. (2001) and Ahlberg R.W. (1999). One of the more relevant studies in the literature was that done by Thompson and Bloom (2000).

Thompson and Bloom reported on a two-phase study that looked at communication of risks between the management level and the technical assessors. (Thompson and Bloom 2000) The first phase of the study consisted of telephone interviews with Environmental Protection Agency managers from across the country. The goal of the first phase was to "understand how each risk manager used risk information in making regulatory decisions about hazardous air pollutants, the interviews followed a topic guide that addressed the following five questions.

1. What processes do risk managers follow in making policy decisions?
2. How do risk managers view their roles in the decision making process compared to the risk assessor's role?
3. Regarding content and format, what risk information do risk managers find most valuable when making decisions?
4. What additional information would risk managers find useful?
5. What techniques have risk mangers found effective for communicating environmental risk to the public?"

(Thompson and Bloom 2000 pp. 334 - 335)

Some of the significant results as identified in the study were:

1. Risk managers trust that the technical assessors of risk know what they are doing.
2. Risk managers felt like there was information that was not communicated that they would like to have received.
3. At the national and non-national level there were different concerns which were reflected by the differences in the nature of the decisions that were being made at that specific level.
4. Communication to the public about risks is challenging.

The second phase of the study consisted of using focus groups to examine samples of risk assessment briefing materials and then getting feedback about their utility. The focus groups had opinions on which materials presented the necessary data and suggested improvements to more clearly communicate the risk and its attributes. The study concluded that:

1. Risk managers desire to see both qualitative and quantitative attributes about the various aspects of the decision at hand.
2. The risk assessors should be aware of the broader view of the decisions to be made.

The risk assessor should communicate both uncertainty and weaknesses of the assessment as this may influence the risk manager's decision.

Conrow and Fredrickson (1996) described common deficiencies in risk management after observing several DoD programs. These are:

"First Deficiency. The risk management process is often weakly structured or 'ad hoc' for both the government and contractors. There may be no clearly delineated mechanism in-place for managing program risk (e.g., organizational responsibilities, analyses, products, etc.), or if a risk management process exists, it may be present on paper only.

Second Deficiency. The risk assessment portion of the risk management process is often too subjective and not adequately documented.

- The prescribed risk assessment categories may be overly broad (e.g., management, technical), leading to difficulty in evaluating results and implementing a viable, measurable risk mitigation strategy.
- A weak risk assessment methodology may be used, which introduces considerable doubt as to the accuracy and value of the results for senior management use.
- Ordinal risk assessment scales are often incorrectly applied. Mathematical operations cannot be applied to scores obtained from uncalibrated ordinal risk assessment scales. Risk values generated by mathematical operations are generally meaningless and may hide true risk issues. (Conrow 1995)

- The risk assessment results may be summarized into broad categories (e.g., low, medium, and high) without sufficient backup to understand the nature of the risk present.
- The government and contractors may use different, incompatible risk assessment methodologies making comparing results difficult, if not impossible.

Third Deficiency. The emphasis of the risk assessment process is generally on the uncertainty associated with a specific event occurring, with less attention given to the consequence of the event occurring. Risk is often inaccurately referenced as only the uncertainty term. However, it is the product of the uncertainty and consequence terms that yields risk. (Defense Systems Management College 1990; Defense Systems Management College 1989) In addition, both the uncertainty and consequence terms require evaluation and tracking over time.

Fourth Deficiency. Program risk assessments and mitigation plans are often unlinked. In addition, they may be prepared on an as-needed basis with limited tracking against key program milestones." (Conrow and Fredrickson 1996 pp. 7 - 8)

Many definitions of risk were found in the literature, but there was commonality found in the basics of their definition. The use of the term uncertainty was found in different contexts throughout the literature. The literature review reveals that risk is important to understand and manage in order to make informed decisions and manage a project successfully. The literature demonstrates that risk communication can be simple or complicated, but states that it should be simple and manageable. As a result of the literature review it was clear that there was a need for the risk assessor to communicate the confidence in their assessment to the decision maker. The literature identified that knowledge in the assessor's confidence enabled the decision maker to expend more resources to increase the confidence level if desired. A review of the literature

regarding risk management and effective risk communications identified that very few researchers have addressed risk assessment confidence levels when using risk scoring methods. What was missing in the body of knowledge was a definition of what a confidence level is, what the confidence levels might be and any data to show the usefulness of confidence levels that could be backed up with statistical data. Table 4 summarizes the literature review by correlating the reference materials to the topic areas.

Table 4 Literature Review Summary

Reference \ Topic	Risk Definition	Communication	Risk Analysis	Risk Assessment	Risk Management	Confidence Level Definition	Assessment Confidence	Usefulness Acceptability Study	Uncertainty	Variability	Risk Matrix/ Risk Scoring
(Ahlberg 1999)		X		X							
(American Bureau of Shipping 2000)	X		X	X					X		X
(Anderson and Narasimhan 1979)				X							
(Andrews et al. 2004)						X	X		X		
(Apostolakis 2004)				X							
(Atman et al. 1994)		X									
(Aven and Pitblado 1998)		X	X	X					X		
(Aven and Vinnem 2007)	X	X			X						
(Banon 1994)				X					X		
(Basu and Lee 2006)				X	X						X
(Ben-David and Raz 2001)	X		X	X	X						
(Bernstein 1996)	X	X							X		
(Bostrom et al. 1994)		X									
(Bradt 2004)	X	X			X						X
(Brown 1990)				X			X		X		
(Carbone and Tippett 2004)			X	X	X						X
(Componation et al. 2001)	X			X	X						X
(Conrow 1995)				X	X						
(Conrow and Fredrickson 1996)					X						
(Covello et al. 1986)		X									
(Cox, Jr. 2008)				X							X
(Cox et al. 2005)				X							X
(Cross and Ballesio 2003)	X			X							
(Dey 2002)	X		X	X	X						X
(Edwards et al. 2001)		X									
(Emblemsvag and Kjolstad 2002)	X			X					X		
(EPA 2004)			X	X	X				X	X	
(Finkel 1994)					X				X		
(Gifford et al. 1979)	X								X		
(Greenberg and Cramer 1991)		X									X
(Grose 1987)	X	X	X	X	X				X		X
(Haimes et al. 2002)				X	X						X
(Hassenzahl 2006)									X		
(Hillson 2003)		X			X						
(Ibrekk and Morgan 1987)		X							X		
(Jaafari 2001)	X				X				X		
(Jardine and Hrudey 1997)	X	X							X		
(Joint Technical Committee OB/7 1999a)	X	X	X	X	X		X				X
(Kahlor et al. 2003)		X									
(Kangari and Riggs 1989)			X	X	X				X		
(Kaplan and Garrick 1981)	X			X					X		
(Klinke and Renn 2001)	X		X	X	X		X		X		X
(Knuth and Connelly 1998)		X									
(Leiss 1996)		X									
(Long and Fischhoff 2000)		X							X		X
(Lyons and Skitmore 2004)					X						
(Macgill and Siu 2005)	X		X	X	X		X		X		X
(Malone and Moses 2005)		X		X					X	X	X
(McNeill and Freiberger 1993)									X		
(Moschandreas and Karuchit 2005)	X	X			X				X	X	
(NASA NPR8000.4 2004)	X	X	X	X	X						X
(National Research Council 1989)		X			X						
(National Research Council 2005)	X	X	X	X	X				X		X

Table 4 Literature Review Summary (con't)

Reference \ Topic	Risk Definition	Communication	Risk Analysis	Risk Assessment	Risk Management	Confidence Level Definition	Assessment Confidence	Usefulness Acceptability Study Usefulness	Uncertainty	Variability	Risk Matrix/ Risk Scoring
(Parrott et al. 2005)		X									
(Pfleeger 2000)	X		X	X	X						
(PMI 2004)	X	X	X	X	X						X
(Rose 2001)			X	X			X		X		X
(Rowe 1977)	X										
(Thamhain and Wilemon 1987)				X							
(Thompson 2002)		X			X				X	X	
(Thompson and Bloom 2000)		X		X	X						
(Verzbolovskis 2004)			X	X	X				X		X
(Viscusi and Zeckhauser 1996)		X									
(Whittemore 1983)			X	X							
(Zio 1997)		X	X	X	X						
(Zmud 1978)		X									

CHAPTER 3: RESEARCH DESIGN AND METHODOLOGY

This chapter provides background for the research, research design and methodology for its validation. Specifically this chapter:

- Validates the need for defining risk assessment confidence levels for use in internal project management communications from the literature research
- Defines the proposed confidence levels and their communication
- Defines the methodology for validation of the importance and useableness of the proposed confidence levels

Validation of Need

Within the existing literature there is a consistent thread that states that 1) it is important for the decision maker to understand the uncertainties associated with the decision at hand and 2) specifically related to risk management that the uncertainty of the likelihood and consequence of a risk, as defined during the risk assessment phase, is important to a decision maker. Thompson and Bloom identified that uncertainty results from imperfection in our knowledge which can sometimes be reduced through better information. They also examine the needs of the decision maker which included having information about "…uncertainties and vulnerability …" (Thompson and Bloom 2000 p. 337) that affect their case. This information also needs to capture "… multiple aspects of the decision…" . (Thompson and Bloom 2000 p. 347) Andrews, Hassenzahl and Johnson (2004) discovered that decision makers are willing to make decisions with a degree of uncertainty, imprecision and ignorance due to the expense and time required for obtaining better information. This does not mean that the decision maker wants to avoid

knowing that there is uncertainty, imprecision and ignorance in the information being provided, only that they desire to make the decision with the information at hand and not pursue additional information. McNeill and Freiberger described the need by decision makers for a balance between meaningfulness and precision with a clear understanding relative to uncertainty. They also note that "… it is important that risk management approaches do not appear more reliable than they are because then decision makers can be led to accept decisions they normally would reject." (McNeill and Freiberger 1993 p. 846)

By showing a "confidence level," the risk assessor can mitigate the problem pointed out by Hassenzahl. Hassenzahl pointed out that the decision maker may be expecting "… precise numerical calculations…" (Hassenzahl 2006 p. 256) and unless told otherwise by the risk assessor, the risk assessment may be interpreted as accurate depictions of the risk. The Australian Standard on Risk Management recognized the importance of defining what it called "confidence" in providing a risk assessment to the decision maker when it stated "Whenever possible, the confidence placed on estimates of levels of risk should be included." (Joint Technical Committee OB/7 1999b p. 13)

Jardine and Hrudey observed that "Overconfidence by scientists in their ability to judge probability and characterize uncertainty has been demonstrated and must be considered in communicating risk. Confidence in risk assessments may be grounded in frequency-based probability estimates. However, if a particular risk assessment must rely primarily on judgment and inference (i.e., subjective probabilities) attaching the confidence associated with frequency-based probability estimates would not be justified. Risk managers and communicators must

recognize these distinctions so they can accurately convey when they are relying on subjective probabilities." (Jardine and Hrudey 1997 p. 492)

The use of confidence levels during communication of the risk assessment complies with the five attributes of successful risk assessment that are proposed by Cross and Ballesio. These five attributes are:

- "Completeness/Comprehensiveness
- Consistency
- Tractability/Documentation
- Credibility/Realism
- Efficiency"

(Cross and Ballesio 2003 p. 3)

The use of confidence levels provides an indicator of the completeness and comprehensiveness of the risk assessment. Cross and Ballesio propose that successful risk assessment "… identifies all of the probabilistically significant risk contributors." (Cross and Ballesio 2003 p. 3) Without the knowledge of the confidence in the risk assessment, the decision maker may assume that a sophisticated analysis has been completed by the risk assessor. The use of confidence levels provides an indicator of consistency between risk assessments. Cross and Ballesio note that "It is worth stressing that one of the main values of risk assessment is the ranking or comparative values of risk contributors. If different parts of the risk assessment are developed inconsistently, such comparisons would be invalid." (Cross and Ballesio 2003 p. 3) The use of confidence levels provides an indicator of the consistency in the analysis between risk assessments. When a decision maker is presented several risk assessments, the confidence level

provides feedback on how extensive or complete the analysis was on each risk. At a high level, the use of confidence levels documents and provides the decision maker with insight into the steps of analysis that were or were not taken, the levels of approximation that were performed. The use of confidence levels provides an indicator of how credible and realistic both the assessor and the assessment are. Without the use of confidence levels, unless the decision maker probes the risk assessment further it may be assumed that the risk assessor is presenting the risk assessment after performing extensive analysis; while the reality is that the assessment is based on little or no real data and is completely subjective in nature. Confidence levels provide an indicator into how conservative the risk assessment methods and assumptions are.

The use of confidence levels also provides a tool for the efficient allocation and use of resources. The decision maker can use the confidence level as an indicator of how important and substantial the risk is and what resources are needed to understand the risk better before committing substantial resources to its mitigation. A risk that has been determined to be low consequence and low likelihood based on no analysis may in reality be of a much higher consequence and/or likelihood when further analysis is accomplished. Without some indictor in the level of analysis, the decision maker may make assumptions and allocate resources to risks which while they have high consequence and/or likelihood, under further inspection and analysis are determined to not be significant.

The American Bureau of Shipping (ABS) states the goal of risk analysis to be the provision of information that "… helps stakeholders make more informed decisions …" (American Bureau of Shipping 2000 p. 49) ABS recognizes a goal to minimize the effort spent on analysis by starting at a high level (less detailed analysis) and then perform more detailed analysis where it benefits significantly the needs of the decision maker. From this, we can derive

that the decision maker needs to know what level of analysis has been done. Figure 4 in Chapter 2 showed a pictorial of the relationship between the level of analysis, certainty and cost.

Rose presents what is called a "chance adequacy matrix". This is described as "… a matrix widely used among many companies, whose origin is not known to the writer. Assuming that a geological model, or concept, is recognized, it compares (1) quantity and quality of information against (2) what the information is signifying with respect to at least minimal adequacy about the particular geologic chance factor." (Rose 2001 p. 36) Rose relates data quality to the matrix. "… in order for us to render judgments of high confidence, either encouraging or discouraging, we require considerable data of good quality. Conversely, sparse or poor-quality data frequently allow only intermediate confidence statements… geoscientists should recognize that the absence of information does not , by itself, imply a negative outcome – only that there are no data." (Rose 2001 p. 36) Rose graphically showed the relationship of data quality and confidence levels in "the chance adequacy matrix" shown in Figure 4 of Chapter 2.

If a risk assessor determines the probability and consequence associated with a risk, "… the degree of reliability associated with the assessment of each component is called certainty of assessment." (Klinke and Renn 2001 p. 161) Certainty of assessment is based on either a qualified or quantified evaluation which Klinke and Renn would rank as extremely low on one end of the spectrum and extremely high on the other end. The nature of the uncertainty when the certainty of the assessment is low can be characterized in "… terms of:

- statistical confidence intervals, which can be calculated on the basis of event distributions or expert opinions;
- remaining fuzzy uncertainties (identifiable, but not calculable); and
- pure ignorance."

(Klinke and Renn 2001 p. 161)

Klinke and Renn state that when the characterization involves statistical or fuzzy uncertainties, there is often an issue with being able to distinguish between them. In this case, the analysis tends to merge these into one characterization and a subjective assessment will include a range of probabilities within the assessment. Klinke and Renn go on to state that "We found it useful, however, to have a distinction between confidence intervals based on empirical distributions or data-rich simulations and informed guesses of potential impacts based on intuition, personal experiences, or scenario brainstorming." (Klinke and Renn 2001 p. 162)

Graphical Representation of Confidence Levels

In order to understand the relationship among likelihood, consequence and confidence levels it is necessary to create a graphical representation of the relationships. The first step is to display a normal 3x3 risk matrix as shown in Figure 10. This risk matrix could also be a 5x5 or other multi-level risk matrix. The 3x3 was chosen in order to simplify the representation of confidence levels which will require display of a 3-d graphical representation. For the purposes of this discussion, the definitions of each level (1, 2, and 3) for likelihood and consequence are not essential.

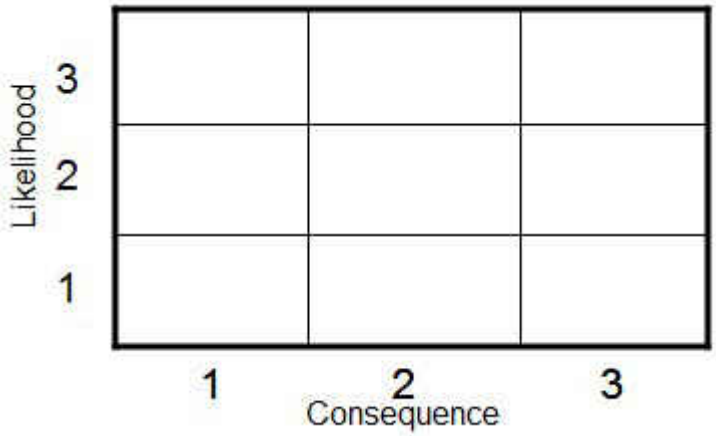


Figure 10 3x3 Risk Matrix

The risk assessor will determine a likelihood and consequence of the identified risk based on whatever definitions are used for that specific project. This may be represented as shown in Figure 11. In this example, a likelihood of 3 and a consequence of 2 have been determined by the assessor.

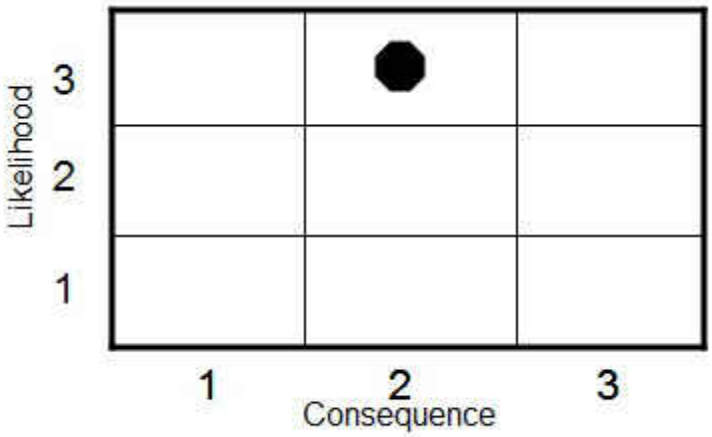


Figure 11 Risk LxC

What is not being communicated to the project manager or other audience is any indication of what resources and effort have been spent on determining this likelihood and consequence score.

The ranges of possibilities are from this being a guess based on nothing but an idea that the risk assessor has, to being the result of an extensive qualitative analysis. How much confidence should the project manager place on the risk assessor's assessment? Without some indicator of how much confidence the risk assessor has in the assessment, the project manager will have to use his/her personal knowledge of the risk assessor, other information that he/she has about the risk being presented, or query the risk assessor about the level of analysis.

This situation can be mitigated by having the risk assessor present confidence levels for the likelihood and confidence scores. This can be conceptualized by thinking of the risk matrix as a 3-d object with confidence level being the 3rd dimension. This is shown in Figure 12.

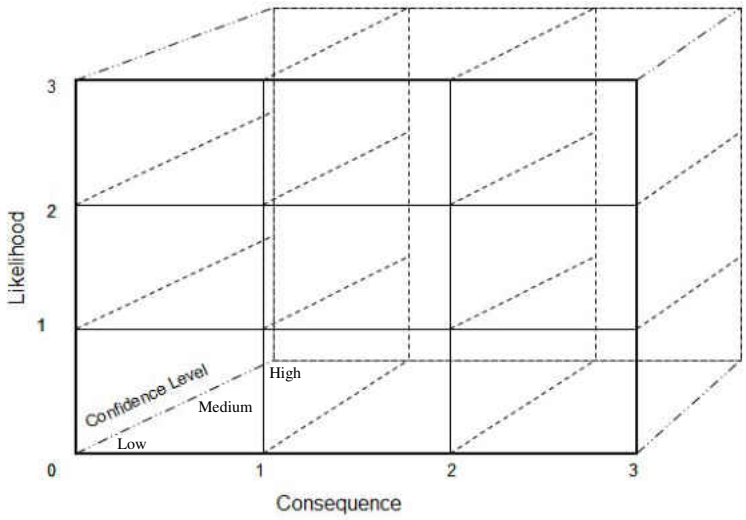


Figure 12 3-D Representation of Risk Matrix

Think of the front of Figure 12 as the nominal LxC matrix. Using the Confidence Level as the 3rd dimension the likelihood (L) and consequence (C) must be thought of separately. Location of the likelihood score would consist of locating it in the matrix at a point consisting of a likelihood (L) score of 1, 2, or 3 and then going back into the 3rd dimension to the confidence level score of low, medium or high. Similarly, the consequence (C) score of 1, 2 or 3 is applied

and then again going back into the 3rd dimension to the confidence level score of low, medium or high for the confidence score. The visualization requires that the likelihood and consequence be thought of as two different points within the 3-d matrix. Applying the risk assessor's likelihood and consequence to the 3-dimensional view, the reader can now begin to visualize that likelihood and consequence have confidence levels.

Confidence Level Components

Everyone has heard the term "garbage in – garbage out". This phrase can be applied to confidence levels by examining three components that determine the level of confidence in the risk analysis and what determines whether the inputs into the risk analysis are garbage or useful.

The three components that determine confidence level are:

- Analysis methodology
- Model quality/fidelity
- Knowledge

As shown in the literature research, analysis methodology can initially be broken down into the categories of qualitative and quantitative. Examples of quantitative methodologies are: probability analysis, Monte Carlo simulation, event or fault tree analysis, modeling, simulations, experiments, prototypes, testing, failure mode and effects analysis (FMEA), quantitative risk analysis (QRA). Examples of qualitative methodologies are: subjective assessment, relevant experience, relevant knowledge, expert opinion/judgement, peer review, or observed previously.

The quality and fidelity of any analytical models that are available or generated are dictated by several factors. These factors are:

- Has the model been verified? Verification checks that the model programming meets the specifications and that there are no errors or bugs in the model.
- Are the models validated? An analytical model can be validated by having real data showing that the analytical model is capable of predicting the output behavior of a system across a reasonable range of inputs.
- Are the input functions of the model known based on past experience or are they assumptions on the part of the model designer?

Knowledge can be defined as in the Oxford English Dictionary's definition of "The fact of knowing a thing, state, etc., or (in general sense) a person; acquaintance; familiarity gained by experience." (Oxford English Dictionary Online 1989) The important aspect of this definition as it relates to the definition of confidence levels is that the risk assessor is able to base a risk assessment on experience. A situation or potential situation (or one similar to it) may have never been experienced before (no knowledge), may have been experienced with a different set of parameters (applicable knowledge) or have been experienced before (relevant knowledge or historical basis).

Proposed Confidence Levels

To define the proposed confidence levels it is necessary to understand the methodologies available that, when accomplished, increase the confidence that the likelihood or consequence score provided to the decision maker reflects reality.

At a very high level, Grose provides a list of traditional methods to reduce or eliminate what was called "ignorance". These methods are:

- “analysis,
- synthesis,
- simulation, or
- test”

(Grose 1987 p. 28)

The shipbuilding industry provides some insights into the use of different levels of analysis that may be done based on the sophistication of the end result that is desired. To reduce the amount of data that is needed, a two tiered-approach is used where Level 1 uses a semi-quantitative assessment of risk to identify the areas requiring further in-depth analysis. Level 2 then focuses on the areas identified in Level 1 and perform additional, more sophisticated analyses. A defining difference between Level 1 and Level 2 is that Level 1 does not require obtaining additional information other than what would be normally available from the design and the accompanying analysis of the vessel under consideration. The Level 1 assessment can be characterized as more subjective than the Level 2 assessment.

In designing confidence levels, it is beneficial to revisit Figure 5 in Chapter 2 which is a pictorial of the process from design concept to operation used by ABS. Using Figure 5 as a starting point we can now convert this concept into confidence levels for risk assessment by taking slices of time at various points. Once the general outline of the confidence levels has been generated, it can then be converted into different scales (1 - 3, 1 - 5, low/medium/high etc.). Different confidence levels will be generated for likelihood and for consequence.

General Confidence Level Definition

It is useful to define the general definition of what a confidence level is before development of a more detailed definition. For the purposes of this research, a confidence level, as it relates to risk assessment is defined as "***The degree of certainty that the likelihood or consequence score (assigned by the risk assessor) reflects reality***".

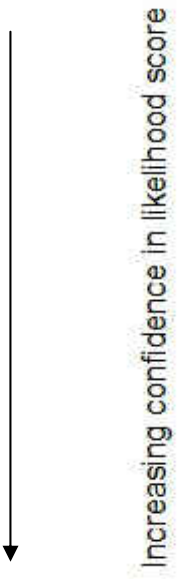
Likelihood Confidence Levels

Defining the confidence in assessed likelihood ranges from making a guess, to certainty because the event has been observed. Table 5 shows various techniques or events that increase our confidence in the assessed likelihood score.

Table 5 Likelihood Confidence Levels

Guess
Subjective Assessment
Relevant Experience
Applicable Knowledge
Expert Opinion/Judgement
Peer Review
Observed previously
Probability Analysis
Monte Carlo Simulation
Event or Fault Tree Analysis
Analysis
Synthesis
Historical Basis or relevant knowledge
Modeling
Simulations
Experiments
Prototypes
Testing

No confidence in likelihood score



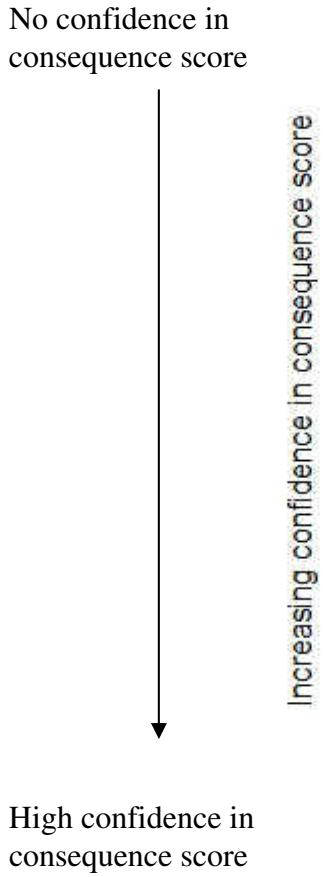
High confidence in likelihood score

Consequence Confidence Levels

Defining the confidence in assessed consequences ranges from making a guess, to certainty because the event has been observed. Table 6 shows various techniques or events that increase our confidence in the assessed consequence score.

Table 6 Consequence Confidence Levels

Guess
Subjective Assessment
Relevant Experience
Applicable Knowledge
Expert Opinion/Judgement
Observed previously
Scenario Brainstorming
Peer Review
Monte Carlo Simulation
Event or Fault Tree Analysis
Analysis
Synthesis
FMEA
QRA
Historical Basis or relevant knowledge
Modeling
Simulations
Experiments
Prototypes
Testing



Communicating Confidence Levels

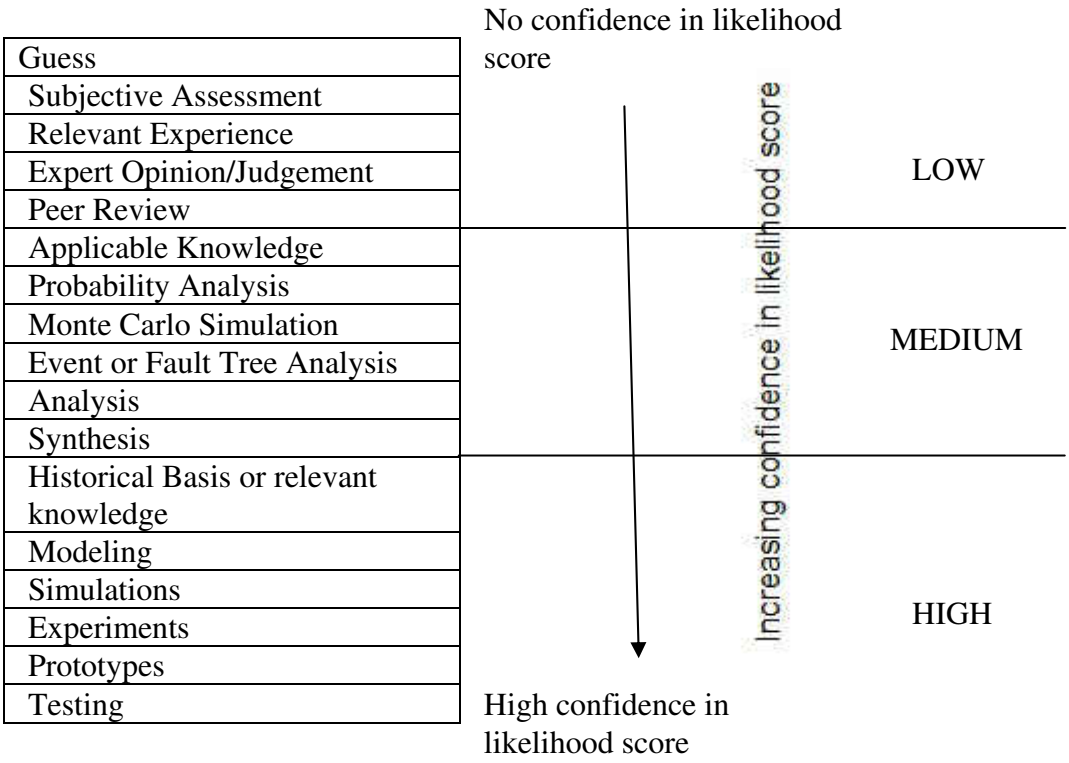
The confidence levels as shown in Table 5 and Table 6 are not practical for communication of the confidence level to the decision maker, but form the foundation on which to build an effective method of communication. This can be accomplished by converting the details of the tasks which determine confidence level into different scales (1 - 3, 1 - 5, low/medium/high etc.). For the purposes of this research an example of conversion into low,

medium and high will be used. The same methodology can be used to convert to a different numeric or alphanumeric scale.

STEP 1:

Using Table 5 and Table 6, determine the points in the tasks which increase the confidence level at which there is a significant increase in resources required to increase the confidence level. The objective is to break each table into three sections labeled as "low," "medium" and "high".

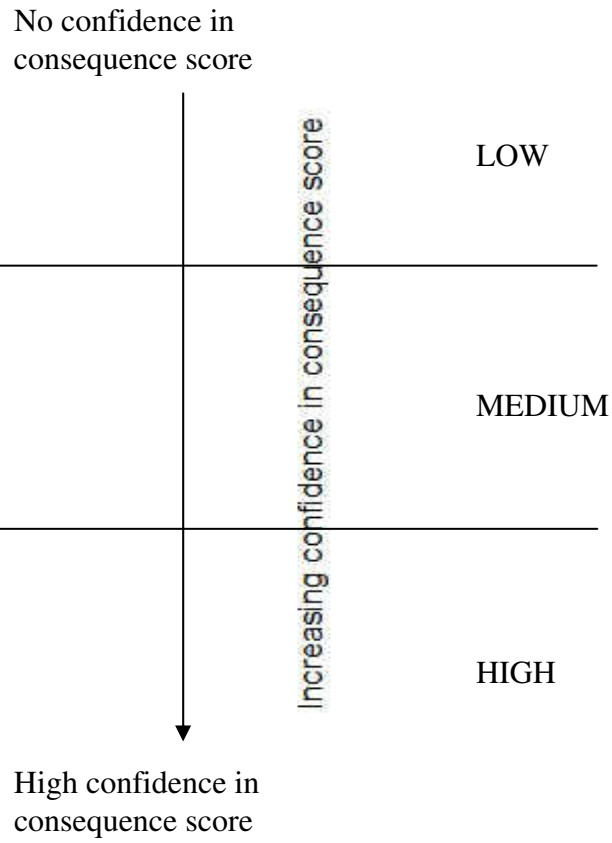
Table 7 Likelihood Confidence Levels Broken Into Levels



The same approach is applied to the consequence confidence levels that were shown in Table 6.

Table 8 Consequence Confidence Levels Broken Into Levels

Guess
Subjective Assessment
Relevant Experience
Expert Opinion/Judgement
Scenario Brainstorming
Peer Review
Applicable Knowledge
Monte Carlo Simulation
Event or Fault Tree Analysis
Analysis
Synthesis
FMEA
QRA
Historical Basis or relevant knowledge
Modeling
Simulations
Experiments
Prototypes
Testing



STEP 2:

Using Table 7 and Table 8, define a brief description of the confidence at each level.

Likelihood:

Low Confidence Level:	Risk assessment is based on subjective opinion or relevant past experiences.
Medium Confidence Level:	Risk assessment is based on similar conditions being observed previously and/or qualitative analysis. Qualitative analysis is based on unverified models and/or data.
High Confidence Level:	Risk assessment is based on testing, high fidelity modeling or simulation, use of prototype or experiments. Qualitative analysis is based on verified models. Quantitative assessment is based on an historical basis and/or data.

Consequence:

Low Confidence Level:	Risk assessment is based on subjective opinion or relevant past experiences.
Medium Confidence Level:	Risk assessment is based on similar conditions being observed previously and/or qualitative analysis. Qualitative analysis is based on unverified models and/or data.
High Confidence Level:	Risk assessment is based on testing, high fidelity modeling or simulation, use of prototype or experiments. Qualitative analysis is based on verified models and/or data. Quantitative assessment is based on an historical basis.

After looking at likelihood and consequence separately, it is evident that the definitions are the same or similar, so for simplicity could be combined into one set of definitions:

Confidence Level Definitions

Low Confidence Level:	Risk assessment is based on subjective opinion or relevant past experiences.
Medium Confidence Level:	Risk assessment is based on similar conditions being observed previously and/or analysis. Analysis is based on unverified models and/or data.
High Confidence Level:	Risk assessment is based on testing, high fidelity modeling or simulation, use of prototype or experiments. Analysis is based on verified models and/or data. Assessment is based on an historical basis.

STEP 3:

Determine options for communication of both the likelihood and consequence confidence levels. Two options should be considered. First is a textual communication and second is graphical. These can be tailored, depending on the project specific methodologies used for risk communication.

Textual communication:

Given that likelihood and consequence may be communicated in several forms, below are examples using a 3x3 risk matrix as the basis (1 – Low likelihood or consequence, 2 – Medium likelihood or consequence, 3 – High likelihood or consequence):

Example 1a:
Likelihood – 2, Consequence – 3;

Example 2a:
Likelihood – 2, Consequence – C (2), S (2), T (3)

In example 2a, a score has been given to each element of consequence being assessed. C is for cost, S is for schedule and T is for technical. Most risk systems take the highest score of the elements being assessed and use that for the overall consequence score.

Text methods for communicating the confidence level (C_L) are:

Example 1b:
Likelihood – 2, Consequence – 3
C_L : M H

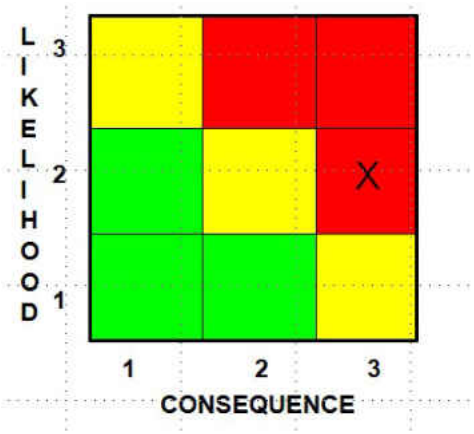
Example 2b:
Likelihood – 2, Consequence – C (2), S (2), T (3)
C_L: M L L H

Example 1b communicates that the risk assessor has observed similar conditions or performed analysis which used unverified models and/or data to provide a 2 on a scale of 1 to 3 (1 - low, 2 – medium, 3 – high) as the likelihood of the risk occurring. This is indicated to the decision maker by the M for medium confidence level accompanying the likelihood score of 2. The risk assessor has more confidence in the consequence score being communicated to the decision maker. This would be due to testing, or high fidelity testing/simulation or the use of prototypes or experiments. In this case there are verified models and/or data to backup the assessment, as well as historical experience to rely upon.

Example 2b is similar to Example 1b except it shows the confidence level in the assessment which has three (3) components which are cost, schedule and technical. In this specific example the assessor is not as confident in the scores assigned for cost and schedule, but has high confidence in the score assigned for the technical score.

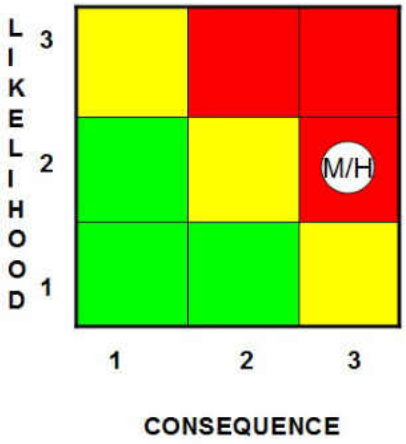
Graphical communication:

Given that likelihood and consequence may be communicated in several forms, below demonstrates how example 1a (Likelihood – 2, Consequence – 3) can be graphically presented using a 3x3 risk matrix as the basis (1 – Low likelihood or consequence, 2 – Medium likelihood or consequence, 3 – High likelihood or consequence):



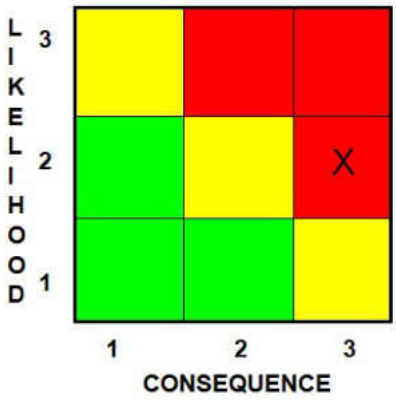
In the graphical version of Example 1a above, the X represents a likelihood score of 2 and a consequence score of 3.

Using the 3x3 graphical matrix, confidence level for Example 1b can be shown as:



In the graphical version of Example 1b above, the circle represents a likelihood score of 2 with a medium confidence level and a consequence score of 3 with a high confidence level.

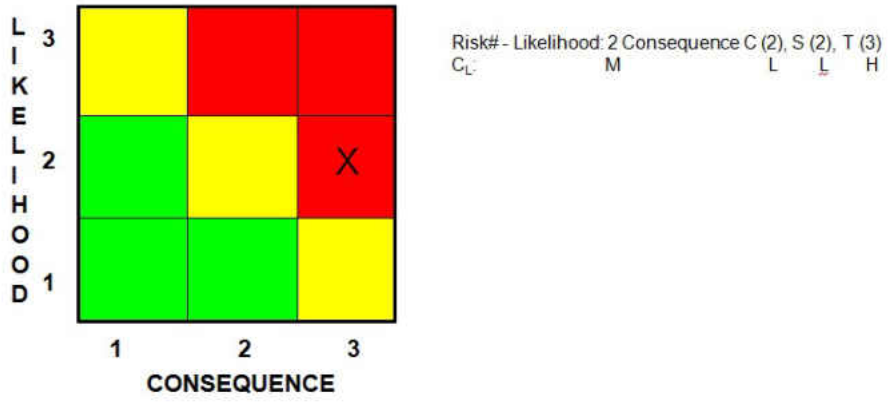
Example 2a can be shown as:



Risk# - Likelihood: 2 Consequence C (2), S (2), T (3)

In the graphical version of Example 2a above, the X indicates a likelihood score of 2 and a consequence score of 3. The accompanying text indicates that the likelihood has a score of 2 and the consequence score has three components. The three components and their individual scores are cost as a 2, schedule as a 2 and technical as a 3. The overall consequence score of 3 is determined by taking the highest of the cost, schedule and technical consequence scores.

Example 2b can be shown in a similar fashion:



In the graphical version of Example 2b above, the X indicates a likelihood score of 2 and a consequence score of 3. The accompanying text indicates that the likelihood has a score of 2 and the consequence score has three components. The three components and their individual scores are cost as a 2, schedule as a 2 and technical as a 3. In this case, a confidence level has been provided for each area of assessment. The likelihood has a medium confidence level. Cost and schedule both have low confidence levels, while technical has a high confidence level. The overall consequence score of 3 is determined by taking the highest of the cost, schedule and technical consequence scores.

There are multiple ways, both textual and graphical to communicate likelihood, consequence and confidence level. In order to focus the research survey on the importance of confidence levels, the simplest method is using text to communicate the concept.

Research Process

To validate the hypothesis that confidence levels are important, it is essential to demonstrate both importance and useableness of the results. The research process is shown in Figure 13. The literature review (Figure 13, Step 1) was conducted to provide an understanding of whether the research topic has been previously covered in the literature and to provide the foundation for achieving the research objectives (Figure 13, Step 2). After development of the hypothesis from the research, a draft survey (Figure 13, Step 3) was developed to serve as the instrument for the validation of the research. In order to validate the survey itself, a focus group was used (Figure 13, Step 4). The focus group reviewed the draft survey and provided feedback to clarify the survey instrument so that usable data would be collected. After incorporation of the modifications based on the focus group feedback the draft survey was modified (Figure 13, Step 5), resulting in a pilot survey (Figure 13, Step 6) which was deployed. The use of a pilot survey enabled verification that the survey was clear and met its intended purpose. In addition, the pilot survey provided an opportunity to verify the analysis approach. After receiving the results from the pilot survey, modifications were made to the survey (Figure 13, Step 7) before deploying the final survey (Figure 13, Step 8) to validate the research results. The survey was then deployed across various industry disciplines and academia. The data that were gathered (Figure 13, Step 9) from the survey were analyzed (Figure 13, Step 10) and the results used to formulate research conclusions (Figure 13, Step 11).

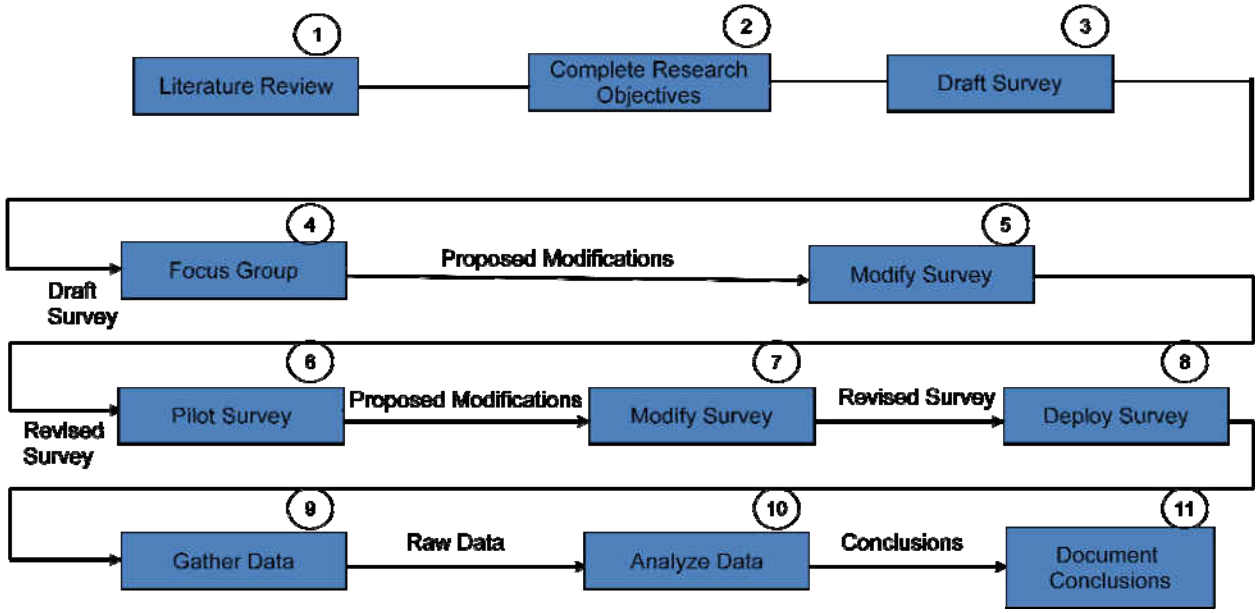


Figure 13 Research Process

Research Procedures

This section enumerates the steps used to gather and analyze data during the research process shown in Figure 13. The "participants" section describes the sources and number of participants in the focus group(s) and survey. The "instruments" section describes the measures used to gather data, their appropriateness and how these measures answer the research question(s). The "procedures" section details the steps taken to contact the participants that are described in the participants section, obtain their cooperation, and then deploy the instruments. The "analysis' section describes the approach to determining how the data are analyzed.

Focus Group

The focus group was used to review the survey before its deployment in order to verify the survey was clear, met its intended purpose and to identify anything that might hinder the survey participants from starting or continuing to take the survey. The focus group feedback was used to modify the survey instrument to maximize its effectiveness in collecting the required data. The draft survey content is found in Appendix A.

Focus Group Participants

It was important that the focus group represented different perspectives on the survey in order to receive balanced feedback. It was also important that the focus group not be so large as to be unmanageable but also not be too small that the balance of perspectives was not achieved. The focus group consisted of eight individuals. These individuals represented a variety of experience in aerospace, industry, and academics. A breakdown of the relevant types of experience as well as total years experience of the focus group is shown in Table 9.

Table 9 Focus Group Experience

Type Experience	#1	#2	#3	#4	#5	#6	#7	#8
Risk Management	X	X		X	X			
Program Management		X			X			
Project Management	X	X	X				X	X
Project/Program Management				X		X		
Project Analysis				X				
Project Integration							X	X
Scheduling							X	
Engineering	X		X					
Operations	X				X			X
Manufacturing							X	
Academia							X	X
Total Years Experience	18	34	12	31	7	12	17	20

Since it was not possible to get all of the participants in the same physical location, the focus group meeting was conducted using a combination of a meeting held at the Johnson Space Center in Houston, TX with additional participation by telecon. At the Johnson Space Center there were three (3) participants in addition to the meeting facilitator. By telecom there were four (4) participants. An additional participant was not able to support the meeting time/date and provided comments by e-mail.

Focus Group Instruments

The instruments used during the focus group consist of a worksheet, instructions, and the survey instrument.

Focus Group Procedures

In preparation for the focus group, research was done on techniques that could be applied during the focus group. Aldag and Kuzuhara (2002) provided some appropriate guidance that the session "… seeks to create the right atmosphere for relaxed, spontaneous thinking. A small group of employees is brought together, presented with the problem and told to follow four rules:

- Don't criticize any ideas. This creates a climate of psychological safety, reducing inhibitions.
- Freewheel. Any idea, no matter how wild, is fine.
- Try to come up with as many ideas as possible. The more ideas, the better.
- Try to combine and improve. "Hitchhiking" on others' ideas may create a chain of inspiration." (Aldag and Kuzuhara 2002 p. 584)

The procedures used for the focus group were intended to prescribe a consistent process between the focus group participants. This was important due to the participants being physically separated during both the preparation and participation phases of the focus group activities.

A worksheet was created to provide a tool for the focus group participants as they initially examined the survey instrument and to maintain consistency in the inquiry during the focus group meeting. The worksheet is included in Appendix B.

The worksheet divided the focus group review of the survey instrument into sections:

- Introductory e-mail
- Introductory text
- Survey core questions and text descriptions
- Demographics
- Thank you statement

Each section had a short list of questions that the focus group participants were asked to review after looking at that particular section of the survey instrument. This procedurally minimized the chances that the focus group participant's initial impression of the survey instrument was influenced by the questions. The questions in the worksheet were:

<u>Introductory E-Mail:</u>

1. Is the e-mail clear as to what is being asked of the recipient?
2. Would recipients be more likely to take or not take the survey based on the e-mail?
3. Is the introductory text clear as to what is being asked of the recipient?
4. Would recipients be more likely to continue or not continue the survey based on the introductory text?

Introductory Text:

1. Is the survey material clear?
2. Is the survey material sufficient?

Survey Core Questions and Text Descriptions:

1. Is the survey material clear?
2. Is the survey material sufficient?
3. Is there redundancy in the questions?
4. How do you feel about the questions?
5. Are there any negative feelings about the questions?
6. Does the order of the questions make a difference?

Demographics:

1. Is the survey material clear?
2. Is the survey material sufficient?

Thank You Statement:

1. Is the survey material sufficient?

Instructions were sent out three (3) days prior to the focus group meeting and one (1) day prior to the worksheet and survey instrument being sent to the focus group participants. Critical components of the instructions were:

- That each participant would receive the survey invitation by e-mail.
- Additionally, each participant would be sent the survey as a pdf file, only for reference if needed since the survey was not designed for someone to go back to look at previous material. Also, the pdf file would be available for reference at the focus group meeting.

- Emphasis that the participant should not use the pdf copy for the initial review since it was important to duplicate how a survey participant will see the survey.
- Upon receipt of the survey invitation, the participant should read each section of the survey and then look at the worksheet questions related to that section and take notes on how they would answer those questions.
- For the focus group activity, the worksheet would outline the questions asked by the facilitator. The focus group participants would be asked for verbal feedback and any group discussion that may be generated. The facilitator was not to engage in two-way dialog, unless it would be to clarify what was said by the focus group participants.
- The purpose of the focus group was emphasized as being to comment on the clarity and understanding of the survey itself, not discussion on what the survey is asking the survey population to rate as agree, disagree, etc.

The survey instrument was sent out to the focus group participants prior to the focus group activity exactly as it would be transmitted to survey participants. This consisted of an invitational e-mail with a link to the on-line survey instrument.

Focus Group Results

The main results of the focus group were:

- Initially the survey was predicted to take five (5) minutes which is an important number since it is quoted in both the introductory e-mail as well as the

- The focus group submitted proposed wording changes to every area of the survey (introductory e-mail, consent statement etc.) These changes really focused on clarifying the text and reducing the text length. Part of the focus here was to reduce the amount of reading by a potential survey participant as they read the introductory e-mail so they commit to taking the survey and then the actual survey itself to reduce the time the participant has to spend, hopefully allowing more participants to actual complete the survey.
- Adding a question with negative phrasing in order to be able to verify that survey participants actually understand the question and not necessarily scoring all the questions the same.
- Addition of an offer in the introductory e-mail to share the results of the survey with the participant. This aided with the question a potential survey participant may ask *What is in this for me?*
- There was a discussion of adding ad-hoc text blocks with each question. This input had to be rejected only because the survey tool did not have this capability. There was one ad-hoc text block for additional comments at the end of the questions.

- The focus group provided an input that one of the demographic questions related to the role the participant felt most experience in for as related to risk management should allow multiple answers.

Institutional Review Board

The University of Central Florida Institutional Review Board approval is shown in Appendix D.

Survey Design

The survey measures the perceived usefulness of confidence levels for use in internal project risk management communications. Usefulness can be defined through two dimensions which are labeled as *importance* and *usableness*. Larcker and Lessig provide definitions of these dimensions which can be applied to the design of the survey. Perceived importance is "… related to whether the information is relevant, informative, meaningful, important, helpful, or significant." (Larcker and Lessig 1980 p. 123) Perceived usableness is "… related to whether the information format is unambiguous, clear, or readable." (Larcker and Lessig 1980 p. 123) Further definitions of these dimensions are provided by Larcker and Lessig as they relate to their study, but they can be applied here also. "… perceived importance will refer to the quality that causes a particular information set to acquire relevance to the decision maker. If the information items are a necessary input for task accomplishment, then this quality will tend to increase the perceived usefulness of the set. Additional characteristics that are related to importance include the significance, meaning, value, and instrumentality attributed to the information set by the

decision maker." (Larcker and Lessig 1980 p. 123) "Perceived usableness will refer to the information quality that allows a decision maker to utilize the set as an input for problem solution. If the information is clearly presented and readable or can be easily transformed, then this quality will tend to increase perceived usefulness." (Larcker and Lessig 1980 p. 123) Larcker and Lessig constructed six statements or questions to measure perceived usefulness. These are shown in Figure 14.

Statements (or Questions) Used in Measurement of the Perceived Usefulness Construct

	completely disagree			neutral			completely agree
A. It would be extremely difficult to complete a specification decision without at least the information presented.	1	2	3	4	5	6	7
B. Extremely complex recalculations or adjustments are necessary in order to use the information presented to complete a specific decision.	1	2	3	4	5	6	7
C. The information presented is sufficient to complete a specific decision.	1	2	3	4	5	6	7
	none			about half			all
D. What portion of the information presented is in the correct form for completion of a specific decision?	1	2	3	4	5	6	7
E. What portion of the information presented is interpretable, without any recalculation or adjustment for the completion of a specific decision?	1	2	3	4	5	6	7
F. What portion of the information presented is essential for or instrumental in completing a specific decision?	1	2	3	4	5	6	7

Figure 14 Measurement of Perceived Usefulness
(Larcker and Lessig 1980 p. 125)

The design of the statements/questions in Figure 14 are designed such that A, C and F are all related to the perceived importance and B, D and E are related to the perceived usableness. The statements and questions are designed so that the sum of the standardized responses to A, C and F can be used as an index of the perceived importance, while the sum of the standardized responses to B, D and E can be used as an index of the perceived usableness. While the concept of importance and usableness designed by Larcher and Lessig are useful to this specific research, the wording had to be modified to fit the hypothesis being tested:

1. It is difficult to complete a risk management decision if you do not know the risk assessor's confidence in the score assessment.
2. The definition of "risk confidence level" is understandable.
3. Providing confidence levels would not improve management decisions.
4. The confidence level definitions are meaningful to a decision maker.
5. The confidence level definitions adequately describe differences in the score assessment confidence.
6. The risk assessor's confidence in the score assessment is important in making a risk management decision.
7. The risk confidence level, when supplementing the likelihood and consequence scoring, is useful in making a decision.

This research used Questions 1, 3, 4 and 6 to obtain data on the *importance* of the use of confidence levels by the decision maker. Questions 2, 5 and 7 are used to obtain data on the *usableness* of the confidence levels as they have been presented in the research hypothesis.

Instead of the 7 point scale used by Larcker and Lessig a 5 point scale which is verbally anchored at the ends and midpoint is used. This decision is based on the work of Lissitz and Lissitz and Green (1975) and Cox III (1980). Research by Lissitz and Green states that there is a "… definite leveling off in the increase in reliability after 5 scale points" and "… strong support

for rejection of 7 scale points as an optimal number." (Lissitz and Green 1975 p. 13) Cox pointed out that "… five alternatives seem adequate …" (Cox III 1980 p. 420) and that "… a scale with the optimal number of response alternatives is refined enough to be capable of transmitting most of the information available from respondents without being so refined that it simply encourages response error." (Cox III 1980 p. 408) What is important in designing the survey is to be able to get the information required. The use of the 5 point enables determination of whether the respondents agree or disagree that confidence levels are useful in internal project management communications.

Research was conducted to aid in the design of the survey. The research of Porter and Whitcomb (2003) provides some guidance on improving web-based survey response. An increased response rate was obtained when:

- A statement of scarcity of participation was included
- A statement that the participant was part of a small group asked to participate
- A deadline as to when the survey website would be shutdown was included

Dillman, Tortora and Bowker (1998) explored some principles that can be used during the design of the survey which they believe will increase survey response. These principles are not tested and are offered by the authors for future research, but they are worth noting for consideration. These principles are:

- "Introduce the web questionnaire with a welcome screen that is motivational, emphasizes the ease of responding, and instructs respondents on the action needed for proceeding to the next page."
- "Begin the web questionnaire with a question that is fully visible on the first screen of the questionnaire, and will be easily comprehended and answered by all respondents."

- "Present each question in a conventional format similar to that normally used on paper questionnaires."
- "Limit line length to decrease the likelihood of a long line of prose being allowed to extend across the screen of the respondent's browser"
- "Provide specific instructions on how to take each necessary computer action for responding to the questionnaire."
- "Provide computer operation instructions as part of each question where the action is to be taken, not in a separate section prior to the beginning of the questionnaire."
- "Do not require respondents to provide an answer to each question before being allowed to answer any subsequent ones."
- "Construct web questionnaires so that they scroll from question to question unless order effects are a major concern, large number of questions must be skipped, and/or a mixed-mode survey is being done for which telephone interview and web results will be combined."
- "When the number of answer choices exceeds the number that can be displayed on one screen, consider double-banking with appropriate navigational instructions being added."
- "Use graphical symbols or words that convey a sense of where the respondent is in the completion progress, but avoid ones that require advanced programming."
- "Be cautious about using question structures that have known measurement problems on paper questionnaires, e.g., check-all-that-apply and open-ended questions."

(Dillman et al. 1998 pp. 7-13)

Survey Structure

The survey consisted of several components:

- Introductory E-mail
- Consent Statement
- Background
- Example
- Survey Questions
- Demographics

Introductory E-mail

The introductory e-mail is an important component to the survey because it is the key as to whether a survey participant continues on to do the survey or deletes the e-mail. It is also essential that the e-mail follow the rules as defined by the Institutional Review Board (IRB).

The text of the introductory e-mail was:

> "Hi,
> My name is Gary L. Johnson; I am a doctoral student in the Industrial Engineering & Management Systems department at the University of Central Florida. The purpose of this survey is to examine the usefulness of "confidence levels" in risk management as it pertains to internal project team communications. This research will result in a tool to improve risk communications. You have been selected to participate in this research due to your extensive knowledge and experience with project management. Your answers will be kept confidential.
>
> This survey will take less than 7 minutes to complete. Research results will be made available. Please do the survey by May 20th.
>
> This research is being conducted under the supervision of Dr. Pet-Armacost (jpetarma@mail.ucf.edu). There is no direct benefit in participation (including compensation) and no penalty for not participating. You will not be required to answer any questions you

do not want to. You must be 18 years of age or older to participate.

Thank you in advance, Gary Johnson, University of Central Florida

Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (IRB). Questions or concerns about research participants' rights may be directed to UCF Institutional Review Board Office at the University of Central Florida, Office of Research and Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246. The phone numbers are 407-823-2901 or 407-882-2276. Survey responses are linked to the e-mail addresses in the Question-Pro system, limited in access to the survey administrator and protected by a user-id/password protocol. Data will be deleted from the server on 9/3/08."

The e-mail used for the survey is shown in Appendix C. The IRB approval letter is in Appendix D.

Consent Statement

The consent statement is necessary because "One of the main ethical responsibilities of a principal investigator is to ensure that potential participants have been provided with all the information they might reasonably need to know regarding a research project in order to make an informed decision whether or not to participate." (University of Central Florida Institutional Review Board 2008 p. 6)

The consent statement has four key components:

1. Purpose of the survey
2. Commitment of the survey participant as being 7 minutes of their time
3. Offer to share in the survey results
4. Institutional Review Board compliance

The text of the consent statement was:

> "Dear Survey Participant:
>
> The purpose of this survey is to examine the usefulness of "confidence levels: in risk management as it pertains to internal project team communications. This research will result in a tool to improve risk communications.
>
> This survey will take less than 7 minutes to complete. Research results will be made available to you at Research Results.
>
> This research is being conducted under the supervision of Dr. Pet-Armacost (jpetarma@mail.ucf.edu). There is no direct benefit in participation (including compensation) and no penalty for not participating. You will not be required to answer any questions you do not wish to. You must be 18 years of age or older to participate.
>
> The results of this study may be published, however your survey responses will be strictly confidential and data from this research will be reported only in the aggregate. If you have any questions at any time about the survey or procedures, you may contact Gary L. Johnson at ga911237@ucf.edu. Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (IRB). Questions or concerns about research participants' rights may be directed to UCF Institutional Review Board Office at the University of Central Florida, Office of Research and Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246. The phone numbers are 407-823-2901 or 407-882-2276. Survey responses are linked to the e-mail addresses in the Question-Pro system, limited in access to the survey administrator and protected by a user-id/password protocol. Data will be deleted from the server on 9/3/08.
>
> Thank you very much for you time and support. Please start with the survey now by clicking on the **Continue** button below."

The consent statement that was used with the survey is shown in Appendix C.

Background

The background section of the survey serves to provide the survey participants with the basis for the survey. The background section provided:

- Risk confidence level definition
- Limitation of the definition to internal communications
- Definitions for low, medium and high risk confidence levels, as used in the survey

The text of the background section was:

> "CONFIDENCE LEVEL – BACKGROUND
> Risk confidence levels are defined as "***The degree of certainty tha the likelihood or consequence score (assigned by the risk assessor) reflects reality***".
>
> Internal communications are defined as those between program or project team members (i.e. communications to the public or shareholders would be considered external communication). The use of risk confidence level for the purposes of this survey is LIMITED TO INTERNAL COMMUNICATIONS only.
>
> **The risk confidence level is used in conjunction with the likelihood and consequence assessment and is intended to provide the program or project team members a sense of the risk assessor's confidence in the likelihood and consequence assessment.**
>
> Confidence level definitions are intended to be tailored for each project based on how that project defines likelihood and consequence (3x3 matrix, 5x5 matrix, 1-5, Low-Med-High, etc.) For the purpose of this survey the confidence levels are defined for a 3x3 (Likelihood x Consequence) risk matrix using magnitudes fo Low-Medium-High.
>
> **Risk Assessor's Confidence Level Definitions:**
> **Low: Risk assessment is based on subjective opinion or relative past experiences.**

Medium: Risk assessment is based on similar conditions being observed previously and/or analysis. Analysis is based on unverified models and/or data.

High: Risk assessment is based on testing, high fidelity modeling or simulation, use of prototype or experiments. Analysis is based on verified models and/or data. Assessment is based on an historical basis."

The background section as it appeared in the survey is shown in Appendix C.

Example

In order to clarify the use of confidence levels for the survey participants, an example of its use was included in the survey.

"**EXAMPLE**

The risk assessor presents the following risk assessment to the project team:

Likelihood score of 2 with a confidence level of "high"

Consequence score of 3 with a confidence level of "low"

In the back-up data you would find tht the high confidence level for the likelihood of 2 is based on the use of proven high fidelity simulations using verified models. The low confidence level for the consequence is based on the opinion of the risk assessor and technical experts because even thought the events which trigger the risk are well modeled, the resulting effects are not modeled or known with certainty.

The project team will now use this information to decide if additional resources should be allocated to raise the confidence in the consequence score to medium to high."

The example that was provided in the survey is shown in Appendix C.

Survey Questions

There are seven (7) questions with a 5-point Likert scale using Strongly Disagree, Disagree, Undecided, Agree and Strongly Agree as the choices for rating each question. The seven questions were:

> "It is difficult to complete a risk management decision if you do not know the risk assessor's confidence in the score assessment.
>
> The definition of "risk confidence level" is understandable.
>
> Providing confidence levels would not improve management decisions.
>
> The confidence level definitions are meaningful to a decision maker.
>
> The confidence level definitions adequately describe differences in the score assessment confidence.
>
> The risk assessor's confidence in the score assessment is important in making a risk management decision.
>
> The risk confidence level when supplementing the likelihood and consequence scoring is useful in making a decision."

In addition to the survey questions, a free text field was available for the respondent to enter ad-hoc comments. The survey questions are shown in shown in Appendix C.

Demographics

Two demographics were collected as part of the survey. The demographics were the role of the survey participant and duration of that role.

The demographic for role was presented at:

"In which role(s) do you have the most experience as related to risk management?
Program/Project Manager
Risk Manager
Risk Assessor
Decision Maker
Academic
Other "

The "other" choice also had a free text field so that the respondent could add additional information.

The demographic for experience was presented as:

"What was the total duration of your roles related to risk management?
Less than 1 year
>= 1 year and < 5 years
>= 5 years and < 10 years
>= 10 years

The two demographic questions that were requested of the survey participants are shown in Appendix C.

Survey Participants

Participants in the survey included industry representatives and academia. The e-mail distribution was consolidated from multiple sources including literature, conferences and professional organizations. E-mail addresses representing a variety of authors was collected from literature during this research wherever available. This includes not only the references included in this dissertation, but also additional references that were reviewed, but not used. E-mail addresses from conference presenters and personal contacts were collected from several

conferences including the NASA Risk Management Conferences, the Project Management Challenge Conferences, Society for Risk Analysis Conferences, and Space Systems Engineering and Risk Management Conference. E-mail addresses were also collected through participation in several professional organizations including the Society for Risk Analysis, Institute of Industrial Engineers, Project Management Institute, and the American Bureau of Shipping.

A confidence level of 95% and a margin of error of 5% were desired. Since the population is very large (greater than 100,000), the standard sample size of 384 (National Statistical Service 2008) is desired. The population is assumed to be worldwide job functions that include program/project managers, decision makers within projects, risk assessors, and risk managers. Based on the work of Porter and Whitcomb (Porter and Whitcomb 2003) a web-based survey may obtain a response rate of 17 – 20%. Based on these calculations the minimum numbers of surveys that must be sent out were between 1,920 and 2,259. When the pilot survey was executed, the response rate was 8%. This accounted for e-mails that bounced back and e-mails that were either not received (due to SPAM filtering) or received but not viewed. Based on the observed response rate from the pilot survey, a minimum of 4,800 surveys were required to be sent.

Survey Results

The survey results were analyzed for the perceived usefulness of the confidence level definition and communication based on the dimensions of *importance* and *usableness*. (Larcker and Lessig 1980 p. 123) In addition two other factors were examined as to how they affected the survey results:

- The role of the survey participant was examined to so that insight of the role can be compared to the responses for a better understanding of the responses.
- The experience of the survey participant was examined so that insight of the experience of the survey participant can be compared to the responses for a better understanding of the responses.

Using the work of Larcher and Lessig (1980), the survey questions were divided into those applicable to the understanding of the survey participant's perceived usefulness through an understanding of importance and through an understanding of usableness. The questions related to these dimensions are as follows

Importance

1. It is difficult to complete a risk management decision if you do not know the risk assessor's confidence in the score assessment.
3. Providing confidence levels would not improve management decisions.
4. The confidence level definitions are meaningful to a decision maker.
6. The risk assessor's confidence in the score assessment is important in making a risk management decision.

Usableness

2. The definition of "risk confidence level" is understandable.
5. The confidence level definitions adequately describe differences in the score assessment confidence.
7. The risk confidence level, when supplementing the likelihood and consequence scoring, is useful in making a decision.

The analysis of the survey results consisted of several parts that were then combined to formulate conclusions about the research. Survey responses will be converted to numerical scores for the analysis. For the survey questions, the Likert scale is converted as follows:

Strongly Disagree – 1
Disagree – 2
Undecided – 3
Agree – 4
Strongly Agree – 5

The first analysis step was to present some overall statistics about the survey. These statistics included:

1. Response Rate
2. Participation by Role
3. Participation by Experience Level
4. Confidence Level

Prior to any further analysis, one modification to the survey data was performed. Since Question 3 is negatively phrased was necessary to reverse score it prior to transferring the data to Minitab or Excel. Reverse scoring is done by converting Likert scale scores of 5 to 1, 4 to 2, 2, to 4 and 1 to 5. The data analysis step was used to calculate different descriptive views of the data including:

1. High level view of the survey results which summarized the responses for each question based on the percentage of responses that were negative, neutral or positive.
2. Descriptive statistics for each question.
3. Testing for equivalent means of the distribution based on role and experience. This was done using a Kruskal-Wallis non-parametric test.
4. Analysis of ad-hoc comments provided to the survey

The results of the survey analysis are shown in Chapter 4.

CHAPTER 4: RESULTS

The objectives of this dissertation are:

1. Validate the need for defining risk assessment confidence levels for use in internal project management communications
2. Define confidence levels
3. Validate that the use of risk assessment confidence levels is considered to be useful in project risk management

This chapter provides the results of the research and proposes potential areas for future research.

Overall Survey Statistics

Overall survey statistics were calculated based on the survey results. These statistics consist of the response rate, the participation in the survey by the role and experience level indicated by the survey participant.

Response Rate

The number of surveys sent to participants from industry representatives and academia were 158 for the pilot survey and 5,302 for the primary survey.

Normally the response rate is calculated using:

$$\text{Response Rate} = \frac{\text{Number of Responses Received}}{\text{Number of Surveys Sent Out}}$$

The goal of this research was to obtain a 17 - 20% response rate as noted by Porter and Whitcomb as nominal for web-based surveys. (Porter and Whitcomb 2003 p. 585) Participation statistics gathered during the survey are shown in Figure 15:

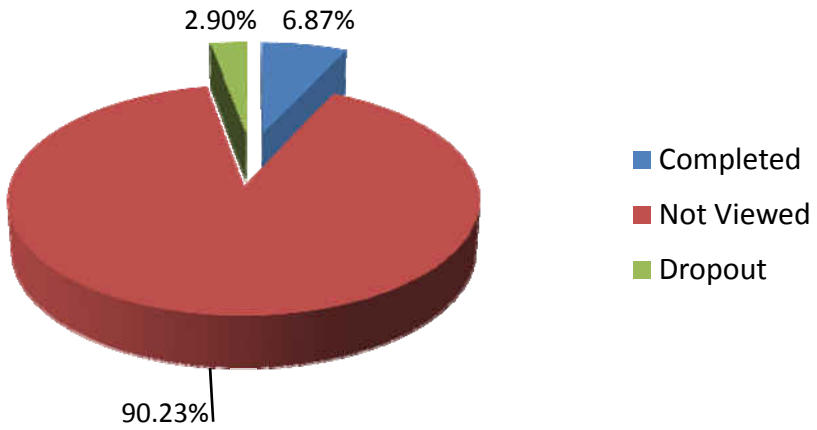


Figure 15 Overall Survey Statistics

Surveys that are actually known to have arrived in the participants in box are indicated by "completed," "viewed – not completed" and "drop out" status. These statistics are shown in Figure 16.

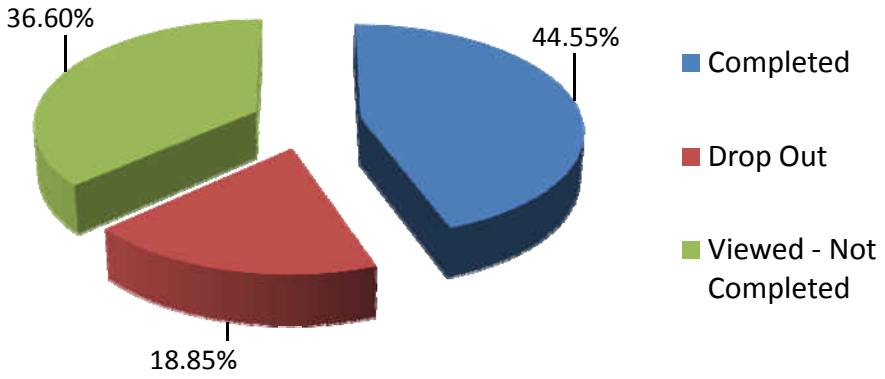


Figure 16 Active Survey Statistics

It is important to note that the use of the term "Number of Surveys Sent Out" needs to be changed in the light of the 2008 environment with SPAM and the use of SPAM filters and firewalls. As can be seen from the survey statistics, many of the e-mails that were sent were trapped by SPAM filters and firewalls as noted by "Not Viewed". For the purposes of this research, the measure of responses received is calculated as the total of those surveys that were "Viewed – Not Completed" and "Completed". This accounts for the number of e-mails that were intercepted by the SPAM filters and firewalls. Using this as the basis, we can then calculate:

Response Rate =	54.90%

In addition, completion rate can be calculated as those surveys that were completed of those that were started. Using this as the basis, we can then calculate:

Completion Rate =	70.27%

These calculations are based on the following data set:

Surveys Sent	Surveys Complete	Not Viewed	Viewed - Not Completed	Started	Dropout	
5302	364	4784	299	518	154	

Where:
Surveys Sent, Surveys Complete, Started, Viewed – Not Completed come from the survey tool
Not Viewed = Surveys Sent - Started
Dropout = Started - Surveys Complete

The survey generated a pool of 364 responses. A review of the data did not find any rationale to exclude any of the data from the completed survey responses.

Participants by Role

Of the 364 responses, the number and percentage of participants by the role in which they had the most experience, as related to risk management is shown in Table 10. The survey tool allowed multiple selections for role. For analysis, the role was reduced to one role based on the most significant role where "Program/Project Manager" was most significant and the remaining hierarchy was "Risk Manager", "Risk Assessor", "Decision Maker", "Academic" and "Other."

Table 10 Participants by Role

Role	Number	Percentage
Program/Project Manager	181	50%
Risk Manager	34	9%
Risk Assessor	70	19%
Decision Maker	29	8%
Academic	34	9%
Other	16	4%

This data is shown graphically in Figure 17.

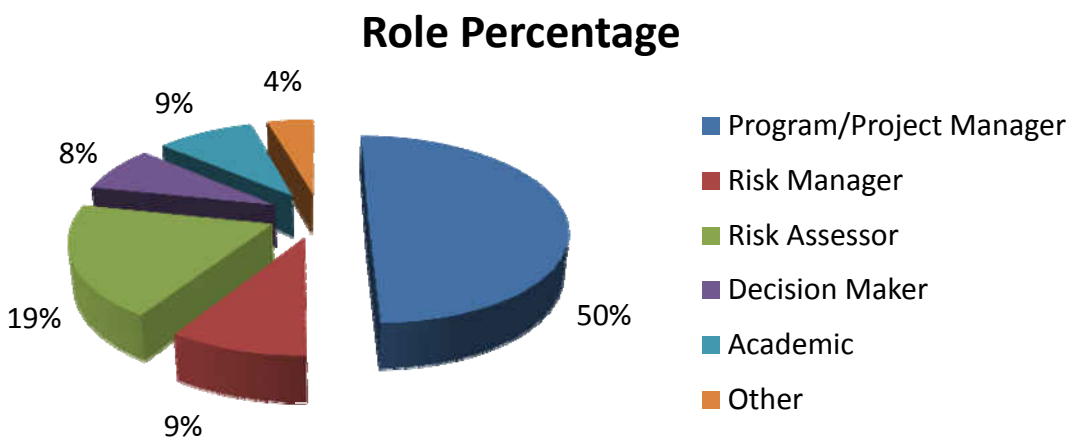


Figure 17 Roles by Percentage

The participation by role indicates a good variety of roles with half of the responses from the program/project Manager role which represents the ultimate decision maker. The survey respondents represent a broad variety of roles associated with decision making and risk management.

Participant Experience Level

Two demographics were collected as part of the survey. The demographics were the role of the survey participant and duration of that role. The duration of the role(s) constitutes the experience level of the participant. Table 11 shows the conversion of the survey narrative to the numerical rating assigned by the survey tool.

Table 11 Conversion of Survey Narrative to Numerical Label

Survey Narrative	Numerical Label
Less than 1 year	1
Equal to or greater than 1 year and less than 5 years	2
Equal to or greater than 5 years and less than 10 years	3
Equal to or greater than 10 years	4

Figure 18 shows the average experience level, broken out by the role of the survey participant.

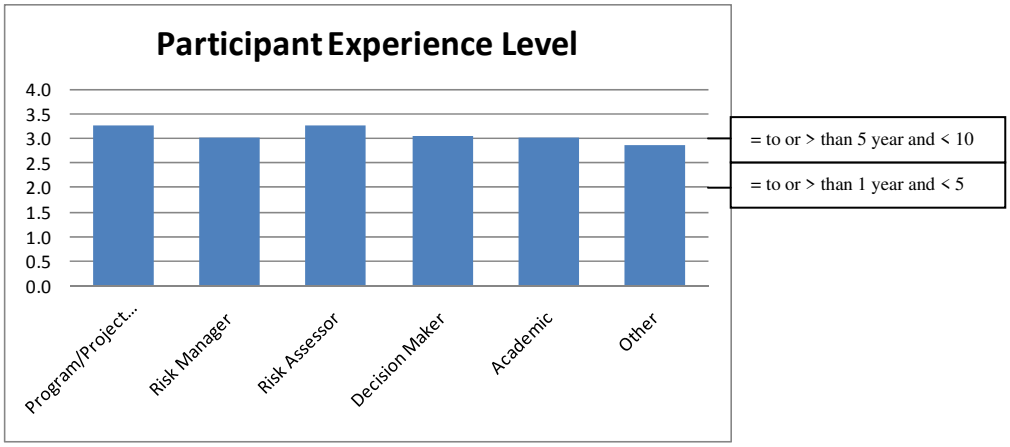


Figure 18 Participant Experience Level

Table 12 shows the interpretation of the numerical rating as assigned by the survey tool to the survey narrative text for the various participant levels.

Table 12Numerical Rating to Narrative Text

Role	Experience	Minimum	Maximum
Program/Project Manager	3.25	Equal to or greater than 5 years	Less than 10 years
Risk Manager	3.03	Equal to or greater than 5 years	Less than 10 years
Risk Assessor	3.26	Equal to or greater than 5 years	Less than 10 years
Decision Maker	3.03	Equal to or greater than 5 years	Less than 10 years
Academic	3.03	Equal to or greater than 5 years	Less than 10 years
Other	2.88	Equal to or greater than 1 year	Less than 5 years
Overall	3.17	Equal to or greater than 5 years	Less than 10 years

Based on the survey data, the overall average experience level of the survey participants is equal to or greater than 5 years, but less than 10 years. The survey does not have the fidelity to make any judgments as to where the average experience level falls within that range (low, mid or high end). The survey respondents represent experience levels from fairly new to experienced risk assessors.

Figure 19 shows for each role, the number of respondents by experience level.

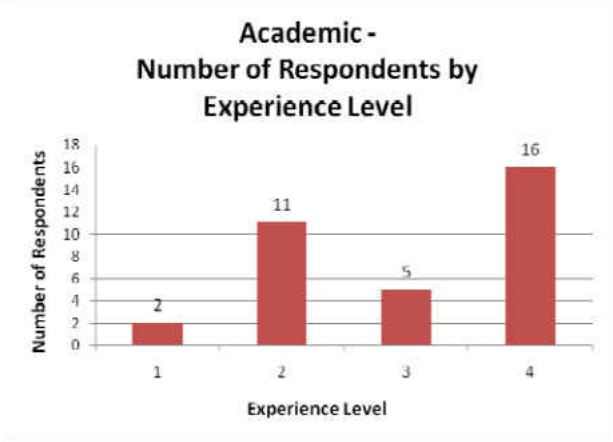


Figure 19 Numbers of Respondents by Experience Level

Confidence Level and Margin of Error

Based on the response to the survey, the confidence level can be calculated using any of the many on-line confidence level calculators. For the purposes of this research the Raosoft Sample Size Calculator (Raosoft 2008) provided these results:

- Confidence Level = 95%
- Margin of Error = +/-5.14%

Data Analysis

Minitab (Release 14.20) (Minitab Corp. 2004) and EXCEL (Version 2007, 12.0.6300.5000) (Microsoft Corp. 2007) were used as tools for the data analysis phase.

Statistical Analysis

The survey data were transferred to Minitab for statistical analysis.

Survey Reliability

The reliability of the survey instrument was measured using Cronbach's alpha. Calculation was performed using an excel based reliability calculator (Siegle 2008). The calculated value of Cronbach's alpha is 0.77. The generally accepted value of reliability is 0.70 or higher when measured using Cronbach's alpha.

Descriptive Statistics

Excel was used to calculate a high level view of the results of the survey. The data shown in Table 13 shows the number of negative, neutral and positive responses to each question. Negative responses are calculated by taking the total count of responses indicating "Strongly Disagree" or "Disagree" divided by the total count of responses to all possible responses. Neutral responses are calculated as the total count of "Undecided" divided by the total count of responses to all possible responses. Positive responses are calculated by taking the total count of responses indicating "Strongly Agree" or "Agree" divided by the total count of responses to all possible responses.

Table 13 Summary of Responses as Negative, Neutral or Positive

	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Negative	12%	10%	12%	9%	16%	7%	5%
Neutral	7%	10%	10%	14%	24%	10%	9%
Positive	80%	79%	78%	77%	60%	83%	86%

The survey was designed to investigate the usefulness of the proposed confidence levels. This involved measuring two aspects which were the importance and the usableness of the confidence levels, as proposed. Table 14 shows the negative, neutral and positive responses, broken down by importance and usableness. Questions 1, 3, 4 and 6 were designed to measure importance and Questions 2, 5 and 7 were designed to measure usableness.

Table 14 Summary of Responses as Negative, Neutral or Positive by Importance and Usableness

Importance	Q1	Q3	Q4	Q6
Negative	12%	12%	9%	7%
Neutral	7%	10%	14%	10%
Positive	80%	78%	77%	83%

Usableness	Q2	Q5	Q7
Negative	10%	16%	5%
Neutral	10%	24%	9%
Positive	79%	60%	86%

Figures 20 – 26 provide the descriptive statistics for the individual questions.

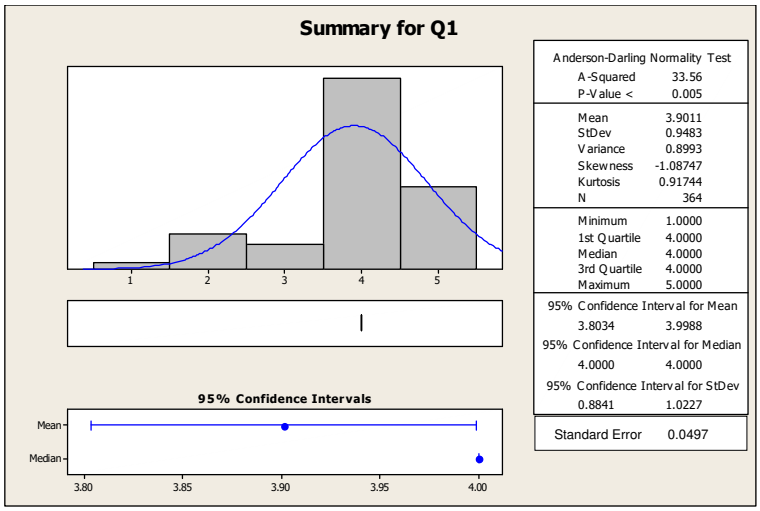


Figure 20 Descriptive Statistics for Q1

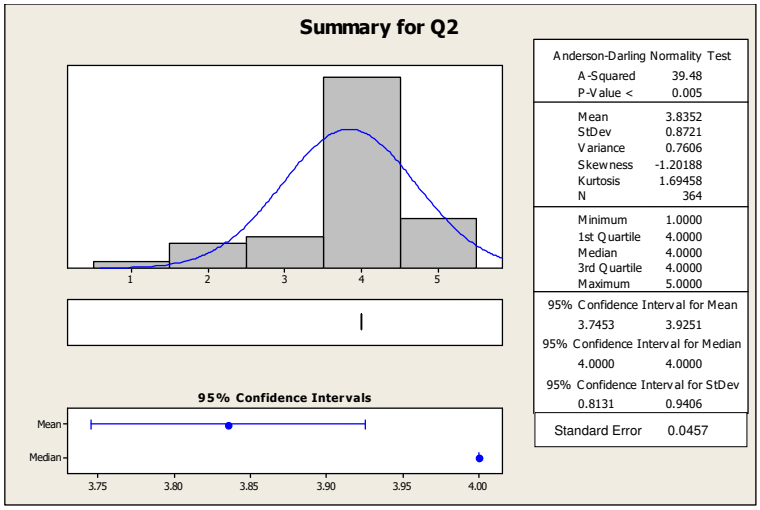


Figure 21 Descriptive Statistics for Q2

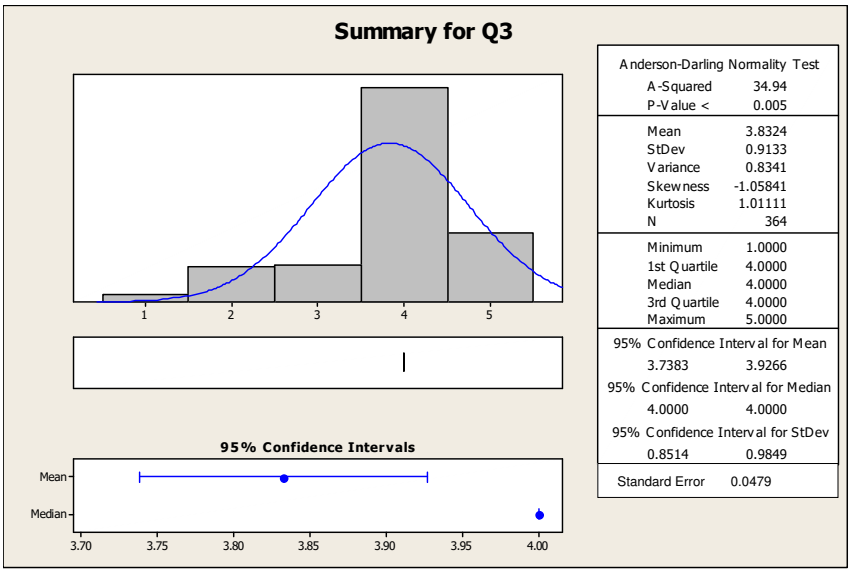


Figure 22 Descriptive Statistics for Q3

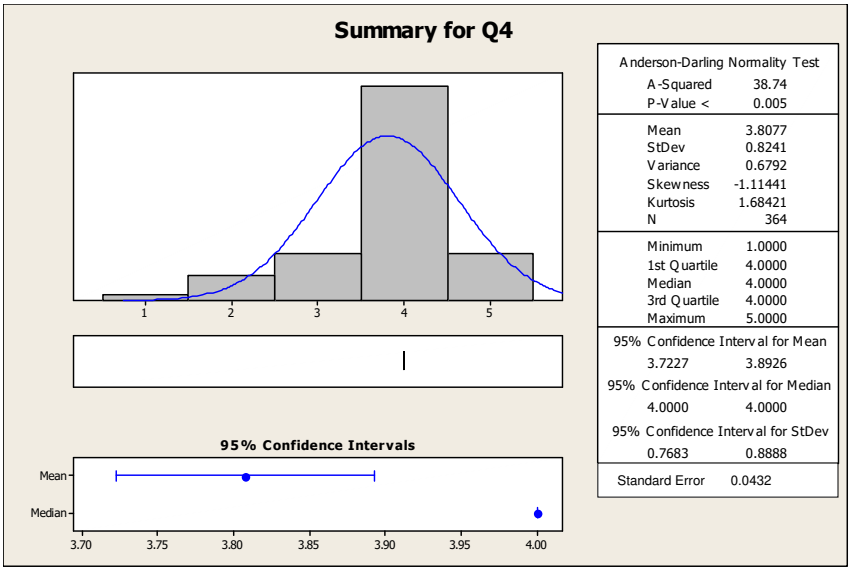


Figure 23 Descriptive Statistics for Q4

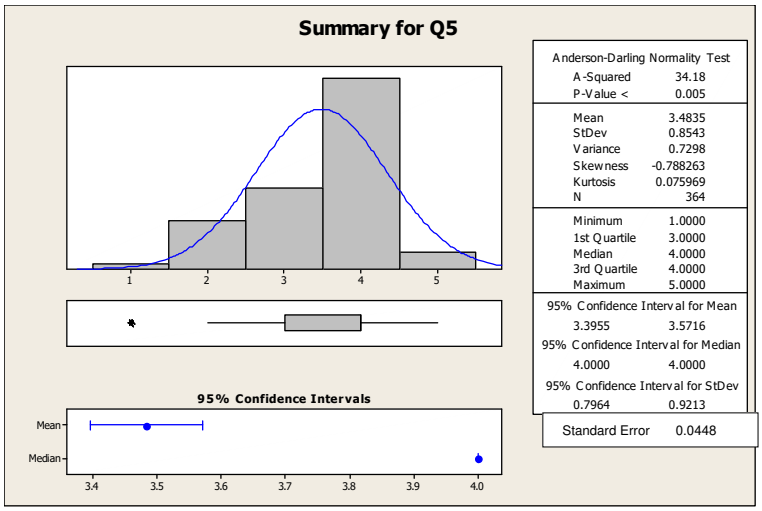


Figure 24 Descriptive Statistics for Q5

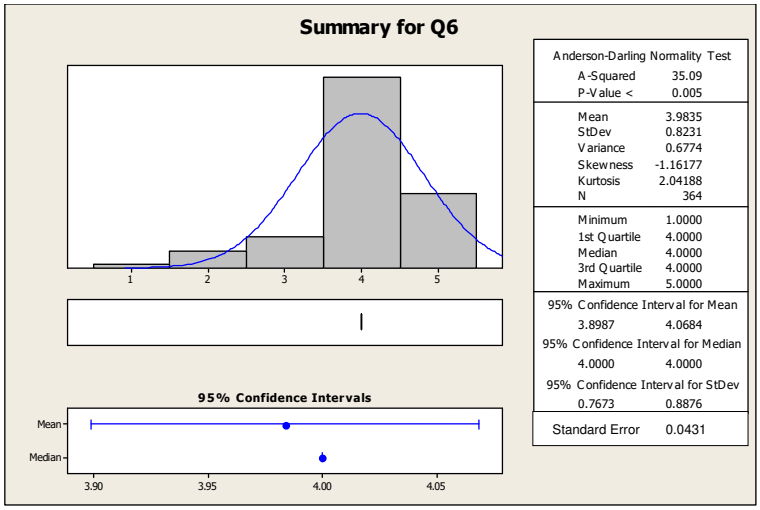


Figure 25 Descriptive Statistics for Q6

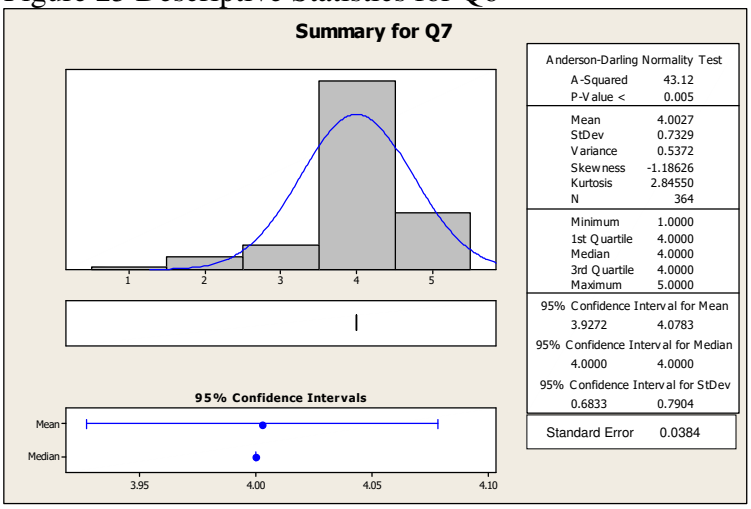


Figure 26 Descriptive Statistics for Q7

The overall statistics, as shown in Table 13, are the most telling as to the results of the survey. Based on breaking the data into negative, neutral and positive responses, Questions 1, 3, 4 and 6 which are related to importance had a range centering around 80% positive scoring. Questions, 2, 5 and 7 which are related to usableness scored lower on average with a 75% positive scoring and with more of a range. Question 5 which scored lowest of any question on the survey at 60% indicates that the actual definitions for the confidence levels are in need of

improvement. The overall evaluation of the importance and usableness factors is that while improvement in the definitions of the actual confidence levels will be necessary to implement them, respondents showed an overall response that confidence levels, as presented are useful to internal project communications. Other descriptive statistics were collected that included data such as mean, median, standard deviation etc., as shown in Figures 20 - 26.

The survey responses at a confidence level of 95% demonstrated a range of errors from 3.84 to 4.97%. Based on the results of the survey, 77 – 83% of those surveyed provided a positive response which indicated agreement that knowing the confidence the assessors have in their assessment is important and would improve a management decision. The survey showed that 60 – 86% of the respondents provided a positive response that the confidence levels and their definitions as presented in the survey were usable. The question with the lowest positive response (60%) was related to the way in which the individual levels were defined.

Testing for Equivalent Means of the Distributions

There are many ways that the survey data can be sorted and displayed. Of interest is how each role (program/project manager, risk manager, risk assessor, decision maker, academic and other) scored each question. Table 15 summarizes each question and the mean score as provided by participants in each role.

Table 15 Role versus Scoring

	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Program/Project Manager	3.84	3.91	3.81	3.86	3.53	3.99	4.02
Risk Manager	3.56	3.94	3.74	3.79	3.59	3.82	3.79
Risk Assessor	3.94	3.60	3.79	3.59	3.23	3.94	4.01
Decision Maker	4.10	3.93	3.93	4.00	3.38	4.10	4.24
Academic	4.12	3.79	4.03	3.68	3.65	4.06	3.88
Other	4.31	3.75	3.81	4.06	3.69	4.06	4.00

From Table 15 there is a definite range of positive responses regardless of the role which corresponds to the overall analysis shown in Table 13.

Additionally there is interest in how the experience level of the participant may affect how they scored each question. Table 16 summarizes each question and the mean score as provided by participants based on their experience level.

Table 16 Experience versus Scoring

	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Less than 1 year	3.85	3.70	4.00	3.95	3.75	4.00	4.00
Equal to or greater than 1 year and less than 5 years	3.81	3.94	3.70	3.88	3.52	4.08	3.95
Equal to or greater than 5 years and less than 10 years	3.88	3.90	3.86	3.86	3.50	4.00	4.10
Equal to or greater than 10 years	3.96	3.78	3.87	3.74	3.43	3.94	3.99

From Table 16 there is a definite range of positive responses regardless of the experience which corresponds to the overall analysis shown in Table 13. It is important to understand whether the role of the survey participant or their experience level affects the results of the survey. Likert survey results in general do not lend themselves to parametric testing due to the inability to assume a normal distribution. For the purposes of this research, a series of non-parametric Kruskal-Wallis tests were performed. The assumptions necessary to perform the Kruskal-Wallis are:

1. "The k samples are random and independent
2. There are five or more measurements in each sample
3. The observations can be ranked.

[Note: No assumptions have to be made about the shape of the population probability distributions.]"
(Scheaffer and McClave 1990 p. 574)

Minitab was used to calculate Kruskal-Wallis statistics for each question based on the role of the survey participant.

Kruskal-Wallis Test: Q1 versus Role

```
Kruskal-Wallis Test on Q1

Role                       N  Median  Ave Rank      Z
Academic                  34   4.000     206.2   1.38
Decision Maker            29   4.000     199.9   0.93
Other                     16   4.000     222.0   1.54
Program/Project Manager  181   4.000     175.0  -1.35
Risk Assessor             70   4.000     191.4   0.79
Risk Manager              34   4.000     147.0  -2.07
Overall                  364              182.5

H = 10.07  DF = 5  P = 0.073
H = 12.46  DF = 5  P = 0.029  (adjusted for ties)
```

Kruskal-Wallis Test: Q2 versus Role

```
Kruskal-Wallis Test on Q2

Role                       N  Median  Ave Rank      Z
Academic                  34   4.000     175.6  -0.40
Decision Maker            29   4.000     197.2   0.79
Other                     16   4.000     178.4  -0.16
Program/Project Manager  181   4.000     189.6   1.28
Risk Assessor             70   4.000     160.4  -1.96
Risk Manager              34   4.000     186.7   0.24
Overall                  364              182.5

H = 4.71  DF = 5  P = 0.452
H = 6.33  DF = 5  P = 0.276  (adjusted for ties)
```

Kruskal-Wallis Test: Q3 versus Role

```
Kruskal-Wallis Test on Q3

Role                       N  Median  Ave Rank      Z
Academic                  34   4.000     201.7   1.12
Decision Maker            29   4.000     191.9   0.50
Other                     16   4.000     185.1   0.10
Program/Project Manager  181   4.000     181.0  -0.28
Risk Assessor             70   4.000     180.2  -0.20
Risk Manager              34   4.000     167.1  -0.90
Overall                  364              182.5

H = 2.17  DF = 5  P = 0.825
H = 2.77  DF = 5  P = 0.735  (adjusted for ties)
```

Kruskal-Wallis Test: Q4 versus Role

```
363 cases were used
1 cases contained missing values


Kruskal-Wallis Test on Q4

Role                       N  Median  Ave Rank      Z
Academic                  34   4.000     164.6  -1.01
Decision Maker            29   4.000     208.5   1.42
Other                     16   4.000     213.1   1.21
Program/Project Manager  180   4.000     188.2   1.11
Risk Assessor             70   4.000     159.8  -1.97
Risk Manager              34   4.000     175.1  -0.40
Overall                  363             182.0

H = 8.08  DF = 5  P = 0.152
H = 10.87  DF = 5  P = 0.054  (adjusted for ties)
```

Kruskal-Wallis Test: Q5 versus Role

```
363 cases were used
1 cases contained missing values


Kruskal-Wallis Test on Q5

Role                       N  Median  Ave Rank      Z
Academic                  34   4.000     198.6   0.97
Decision Maker            29   4.000     168.1  -0.74
Other                     16   4.000     202.1   0.78
Program/Project Manager  180   4.000     186.2   0.75
Risk Assessor             70   4.000     160.1  -1.94
Risk Manager              34   4.000     190.9   0.52
Overall                  363             182.0

H = 5.52  DF = 5  P = 0.356
H = 6.80  DF = 5  P = 0.236  (adjusted for ties)
```

Kruskal-Wallis Test: Q6 versus Role

```
Kruskal-Wallis Test on Q6

Role                       N  Median  Ave Rank      Z
Academic                  34   4.000     190.3   0.45
Decision Maker            29   4.000     192.1   0.51
Other                     16   4.000     190.8   0.32
Program/Project Manager  181   4.000     181.6  -0.17
Risk Assessor             70   4.000     182.9   0.04
Risk Manager              34   4.000     166.7  -0.92
Overall                  364             182.5

H = 1.31  DF = 5  P = 0.934
H = 1.70  DF = 5  P = 0.889  (adjusted for ties)
```

Kruskal-Wallis Test: Q7 versus Role

```
362 cases were used
2 cases contained missing values


Kruskal-Wallis Test on Q7

Role                       N  Median  Ave Rank      Z
Academic                  34   4.000     163.8  -1.04
Decision Maker            29   4.000     212.9   1.69
Other                     16   4.000     180.9  -0.02
Program/Project Manager  180   4.000     182.1   0.11
Risk Assessor             69   4.000     188.8   0.65
Risk Manager              34   4.000     154.5  -1.58
Overall                  362             181.5

H = 6.20  DF = 5  P = 0.287
H = 8.82  DF = 5  P = 0.116  (adjusted for ties)
```

Minitab was used to calculate Kruskal-Wallis statistics for each question based on the experience of the survey participant.

Kruskal-Wallis Test: Q1 versus Experience Level

```
Kruskal-Wallis Test on Q1

Experience Level                                          N  Median  Ave Rank
Equal to or greater than 1 year, but less than 5 years   84   4.000     165.6
Equal to or greater than 5 years, but less than 10 years 74   4.000     182.7
Greater than 10 years                                   186   4.000     190.9
Less than 1 year                                         20   4.000     174.2
Overall                                                 364             182.5

Experience Level                                             Z
Equal to or greater than 1 year, but less than 5 years   -1.68
Equal to or greater than 5 years, but less than 10 years  0.02
Greater than 10 years                                     1.57
Less than 1 year                                         -0.36
Overall

H = 3.49  DF = 3  P = 0.322
H = 4.31  DF = 3  P = 0.229  (adjusted for ties)
```

Kruskal-Wallis Test: Q2 versus Experience Level

```
Kruskal-Wallis Test on Q2

Experience Level                                          N  Median  Ave Rank
Equal to or greater than 1 year, but less than 5 years   84   4.000     190.9
Equal to or greater than 5 years, but less than 10 years 74   4.000     188.9
Greater than 10 years                                   186   4.000     177.9
Less than 1 year                                         20   4.000     166.1
Overall                                                 364             182.5

Experience Level                                          Z
Equal to or greater than 1 year, but less than 5 years  0.84
Equal to or greater than 5 years, but less than 10 years 0.58
Greater than 10 years                                  -0.85
Less than 1 year                                       -0.72
Overall

H = 1.65  DF = 3  P = 0.647
H = 2.22  DF = 3  P = 0.528  (adjusted for ties)
```

Kruskal-Wallis Test: Q3 versus Experience Level

```
Kruskal-Wallis Test on Q3

Experience Level                                          N  Median  Ave Rank
Equal to or greater than 1 year, but less than 5 years   84   4.000     169.3
Equal to or greater than 5 years, but less than 10 years 74   4.000     183.6
Greater than 10 years                                   186   4.000     186.8
Less than 1 year                                         20   4.000     194.1
Overall                                                 364             182.5

Experience Level                                          Z
Equal to or greater than 1 year, but less than 5 years -1.31
Equal to or greater than 5 years, but less than 10 years 0.10
Greater than 10 years                                   0.79
Less than 1 year                                        0.51
Overall

H = 1.88  DF = 3  P = 0.598
H = 2.39  DF = 3  P = 0.495  (adjusted for ties)
```

Kruskal-Wallis Test: Q4 versus Experience Level

```
363 cases were used
1 cases contained missing values


Kruskal-Wallis Test on Q4

Experience Level                                          N  Median  Ave Rank
Equal to or greater than 1 year, but less than 5 years   84   4.000     191.3
Equal to or greater than 5 years, but less than 10 years 74   4.000     186.0
Greater than 10 years                                   185   4.000     175.2
Less than 1 year                                         20   4.000     191.0
Overall                                                 363             182.0

Experience Level                                          Z
Equal to or greater than 1 year, but less than 5 years  0.92
Equal to or greater than 5 years, but less than 10 years 0.37
Greater than 10 years                                  -1.26
Less than 1 year                                        0.39
Overall

H = 1.69  DF = 3  P = 0.640
H = 2.27  DF = 3  P = 0.518  (adjusted for ties)
```

Kruskal-Wallis Test: Q5 versus Experience Level

```
363 cases were used
1 cases contained missing values


Kruskal-Wallis Test on Q5

Experience Level                                          N  Median  Ave Rank
Equal to or greater than 1 year, but less than 5 years   84   4.000     186.0
Equal to or greater than 5 years, but less than 10 years 73   4.000     181.6
Greater than 10 years                                   186   4.000     177.7
Less than 1 year                                         20   4.000     206.8
Overall                                                 363             182.0

Experience Level                                          Z
Equal to or greater than 1 year, but less than 5 years  0.40
Equal to or greater than 5 years, but less than 10 years -0.04
Greater than 10 years                                  -0.80
Less than 1 year                                        1.09
Overall

H = 1.55  DF = 3  P = 0.670
H = 1.91  DF = 3  P = 0.591  (adjusted for ties)
```

Kruskal-Wallis Test: Q6 versus Experience Level

```
Kruskal-Wallis Test on Q6

Experience Level                                          N  Median  Ave Rank
Equal to or greater than 1 year, but less than 5 years   84   4.000     192.4
Equal to or greater than 5 years, but less than 10 years 74   4.000     180.6
Greater than 10 years                                   186   4.000     179.8
Less than 1 year                                         20   4.000     173.0
Overall                                                 364             182.5

Experience Level                                             Z
Equal to or greater than 1 year, but less than 5 years    0.98
Equal to or greater than 5 years, but less than 10 years -0.17
Greater than 10 years                                    -0.50
Less than 1 year                                         -0.42
Overall

H = 1.05  DF = 3  P = 0.788
H = 1.37  DF = 3  P = 0.713  (adjusted for ties)
```

Kruskal-Wallis Test: Q7 versus Experience Level

```
362 cases were used
2 cases contained missing values


Kruskal-Wallis Test on Q7

Experience Level                                          N  Median  Ave Rank
Equal to or greater than 1 year, but less than 5 years   84   4.000     173.9
Equal to or greater than 5 years, but less than 10 years 73   4.000     191.6
Greater than 10 years                                   185   4.000     182.0
Less than 1 year                                         20   4.000     172.1
Overall                                                 362             181.5

Experience Level                                             Z
Equal to or greater than 1 year, but less than 5 years   -0.76
Equal to or greater than 5 years, but less than 10 years  0.93
Greater than 10 years                                     0.09
Less than 1 year                                         -0.41
Overall

H = 1.29  DF = 3  P = 0.730
H = 1.84  DF = 3  P = 0.606  (adjusted for ties)
```

Does the role or experience of the survey participants affect the response to the survey questions? If the answer is no, then we would expect similar means for the distributions despite the role that the survey participant has. We state this in terms of a hypothesis as:

Hypothesis Test for Role

H_o: The probability distribution of the survey responses are identical regardless of the role of the survey participant.

H_a: At least two of the distribution of the survey responses differ

(Northwestern University Medical School 2008)

Hypothesis Test for Experience

H_o: The probability distribution of the survey responses are identical regardless of the experience of the survey participant.

H_a: At least two of the distribution of the survey responses differ

To determine if Ho should be rejected, the rejection region must be determined:

Rejection region: H>χ^2_α(k-1), which is the critical value (Scheaffer and McClave 1990 p. 574)

Table 17 Kruskal-Wallis Test Results by Role with Rejection Region

	H Test Statistic	P	H Test Statistic Adjusted for ties	P Adjusted for ties	Reject based on critical value? 11.0705
Importance					
Question 1	10.07	.073	12.46	.029	Yes
Question 3	2.17	.825	2.77	.735	No
Question 4	8.08	.152	10.87	.054	No
Question 6	1.31	.934	1.70	.889	No
Usableness					
Question 2	4.71	.452	6.33	.276	No
Question 5	5.52	.356	6.80	.236	No
Question 7	6.20	.287	8.82	.116	No

Critical Values from (Scheaffer and McClave 1990 p. 661)
DF = 5

The results are shown in Table 17. For six of the seven questions, the results of the Kruskal-Wallis test failed to reject the hypothesis that the distributions of the response means were identical. The exception is Question 1 which indicates that there are differences based on role and by inspection of the responses, it appears that decision makers, academics, and others more strongly support the need for confidence level information to reduce the difficulty in making risk based decisions in projects.

Table 18 Kruskal-Wallis Test Results by Experience with Rejection Region

	H Test Statistic	P	H Test Statistic Adjusted for ties	P Adjusted for ties	Reject base on critical value? 7.81473
Importance					
Question 1	3.49	.322	4.31	.229	No
Question 3	1.88	.598	2.39	.495	No
Question 4	1.69	.640	2.27	.518	No
Question 6	1.05	.788	1.37	.713	No
Usableness					
Question 2	1.65	.647	2.22	.528	No
Question 5	1.55	.670	1.91	.591	No
Question 7	1.29	.730	1.84	.606	No

Critical Values from (Scheaffer and McClave 1990 p. 661)
DF = 3

The results are shown in Table 18. For all of the seven questions a non-parametric Kruskal-Wallis test generally failed to reject the hypothesis that the means of the distributions of the survey response were identical.

Analysis of Ad-Hoc Comments

The survey allowed ad-hoc comments within a text block described as "Please add any comments and suggestions you might have". There were a total of 105 comments provided by the survey participants which constituted 28.8% of the respondents. During the analysis of the comments, some comments were marked as "Discard" because they were deemed as general and not providing any additional information useful for the analysis phase. During the initial review, 18 comments were discarded which left 87 comments or 23.9% of the respondents. The

comments from the survey participants were analyzed to improve the understanding of the overall survey responses. This analysis assumes that survey participants that had taken the time to write comments cared enough or felt strongly about the survey material and needed to be factored in the overall research conclusions. Analysis of the mean responses between the respondents that provided comments and those that did not provide comments is shown in Figure 27.

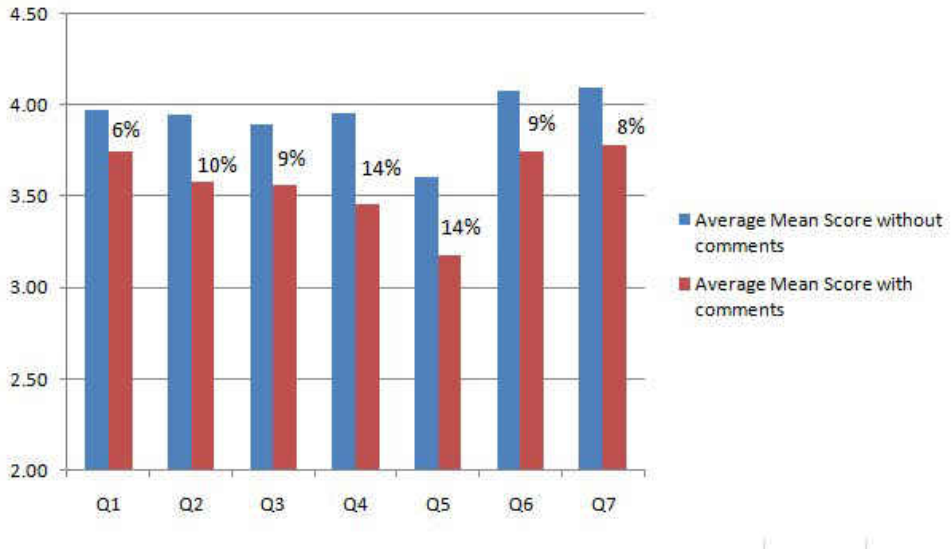


Figure 27 Comparison of Mean for Respondents Submitting or Not Submitting Comments

An observation that for every question, those that did submit comments on an average scored the questions consistently lower than those that did not submit comments.

The process for analyzing comments:

1. Categorize all comments into one of four categories:
    - General
    - Pertaining to Importance
    - Pertaining to Usableness
    - Pertaining to both Importance and Usableness

2. Mark as “Discard” any comments that were deemed as general and not providing any additional information useful for the analysis phase.

3. Comments were groups by similarity.

4. Examine the comments in light of the survey results.

5. Draw conclusions based on the ad-hoc survey comments.

Categorization of Comments

All 105 comments were categorized into these four categories:

- General
- Pertaining to Importance
- Pertaining to Usableness
- Pertaining to both Importance and Usableness

Table 19 shows the resulting numbers of comments assigned to each category.

Table 19 Number of Comments Assigned to each Category

Category	Number of Comments Assigned
General	48
Pertaining to Importance	27
Pertaining to Usableness	25
Pertaining to both Importance and Usableness	5

Discard of Comments

Each comment was reviewed and those deemed as general and not providing any additional information useful for the analysis phase were marked as “Discard”. Examples of

comments that were discarded from the analysis include “Questions 1 and 6 and 7 seem similar. Questions 3 and 4 seem similar,” “The questions in which I was 'undecided' were questions that I did not understand fully” and “Your method of looking at this seems to be from an academic model rather than that of a practitioner model.” Table 20 shows the results of the discard phase.

Table 20 Number of Comments Remaining after Discard Review

Category	Number of Comments Assigned	Discard	Number of Comments Remaining
General	48	16	32
Pertaining to Importance	27	0	27
Pertaining to Usableness	25	1	24
Pertaining to both Importance and Usableness	5	1	4
Totals	105	18	87

Comment Grouping

The remaining comments were grouped according to their similarity into one of eleven categories. These categories are:

- Confidence Level Definition
- Confidence in the assessor(s)
- Data Presentation
- Different Approach
- Informational
- Other Factors
- Practicality

- Relationship to Consequence and Likelihood
- Risk Matrices Issues
- Subjectivity
- Support

Figure 28 shows the frequency of comments by category.

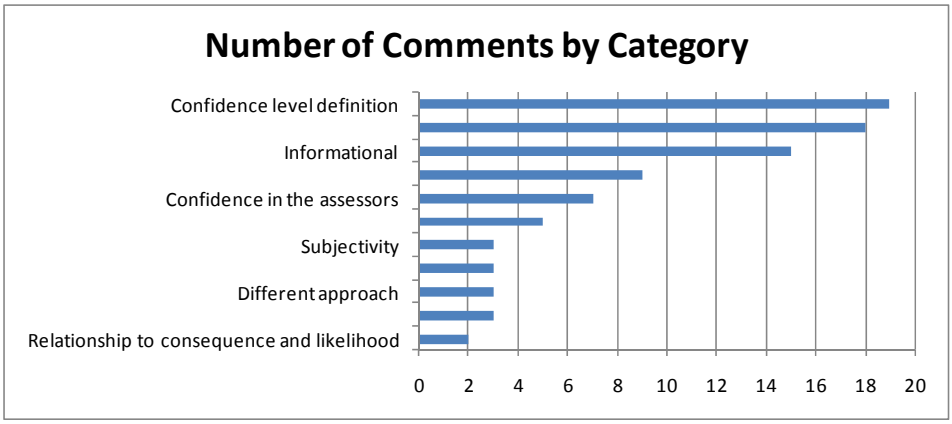


Figure 28 Number of Comments by Category

Comment Analysis

This section will present each comment grouping with analysis and examples of the comments received. The selection of comments presented was transcribed to eliminate identification of the respondent.

Confidence Level Definition

There were nineteen (19) comments that were grouped as related to the definition of confidence levels. The ad-hoc comments support the quantitative data in that they indicate that additional work is required to make the confidence level definitions clearer and more usable. In addition, there are several comments that indicate they do not agree

with the importance of confidence levels to the decision maker. Examples of survey comments related to the confidence level definition:

- "'Risk' cannot have confidence. The evaluator, assessor or manager can have confidence. I think 'Risk Assessor Confidence Level' is meaningful."

- "Might be easier to just have one confidence score for the risk. Manager will also have his internal confidence level, which may be the real deciding factor."

- "The challenge is always the grey area between the levels so it is very important to provide criteria that clearly helps the assessor make and defend the decisions."

- "I work in the risk assessment arena and found your description of confidence levels to be more academic than useful. It's better to simplify your terms so that the responder doesn't have to work so hard to understand your question."

- "Would prefer a 4 point scale. This would force the level into a generally favorable or unfavorable state and take out the 'neutral' (Medium) option which is otherwise likely to be overly relied upon."

- "I find this jargon confusing. You have a risk estimate, which is the probability of a specified consequence occurring. So now you're adding a consideration of confidence in the probability estimate and in the specification of the consequence? If your confidence in every key aspect of the assessment seems questionable, decision makers won't use it."

- "The term 'confidence levels' may be confusing to some, as this has a very specific and limited meaning as applied in traditional statistics. Additionally, what you are describing is presumably a mix of variability, uncertainty, and subjective judgment- which may be confusing. I would suggest consulting the text 'Uncertainty' by Morgan and Henrion' and reworking the term and the definition. "

- "A decisionmaker typically tends to use some aggregate of likelihood and consequence, rather than individual quantities, even if confidence levels are communicated separately. There is a lot of literature available on communication of uncertainty in risk assessment (see many papers published in Risk Assessment). The proposed definitions may not work in some cases like nuclear reactor safety studies, where confidence levels are often high even in situations where they are low using these definitions."

- "Your definitions are too confusing."

- "I would suggest that for complex systems, more granularity could be added to the confidence levels/definitions."

- "You could make the difference between low confidence levels based on 'relative past experiences' and medium confidence levels based on 'similar conditions being observed previously' clearer. The current definitions overlap somewhat and are ambiguous."

- "The 'risk confidence' definition is understandable, but not ideal - might still be confusing at first to people unfamiliar with risk terminology, I would recommend 'risk certainty' or even 'risk uncertainty' level as this is what you are really getting at. I think the information is very useful, but you would need to provide clearer examples, of how you would make decisions based on different risk confidence levels with the 3 tier system and risk matrix. Also, I would recommend including 1 more level, to give a 'low' and 'very low', because it seems to me that many judgments would fall into the low category, but some will be less uncertain then others."

- "I don't agree with the definition given here of a risk confidence level. This should have to do with the uncertainty in the risk estimate. The issues of the pedigree of the risk estimate, which seem to be the focus here, are important but not the main issue - the main issue is the uncertainty. Furthermore, qualitative definitions are vague and thus not useful."

Confidence in the assessor(s)

There were seven (7) comments that were grouped as related to the confidence in the assessor(s). These ad-hoc comments are interesting in the fact that they point out a different aspect of confidence, which is the importance of the decision maker's confidence in the risk assessor, versus the risk assessor's confidence in their assessment.

Examples of survey comments related to confidence in the risk assessor(s):

- "As a manager you have to know your risk manager and have confidence in that person. If they are new then the confidence level is, in reality, meaningless. If they have a history and do well, then it is very meaningful. Otherwise you have to be able to ascertain what the risk factors really are and decide for yourself."

- "From a management prospective confidence in the risk assessor is probably more significant than the risk assessor's confidence."

- "The risk assessor's confidence in the technical knowledge of the people actually performing the analysis and the previous history with these people is important when making decisions."

- "All of the above answers are highly dependent on the experience level of the risk assessor, and the quality of the inputs."

- "The utility of the information would depend in large part on the manager's trust in the person giving the assessment and confidence level. At least, that is true for me, and for the legislators with whom I work."

- "In a large project, risk assessment is solicited from a group of experts and hence the confidence level is always high."

Data Presentation

There were three (3) comments that were grouped as related to the presentation of data. These ad-hoc comments discuss the importance of having an appropriate justification for the assessment, that presenting a variety of possible outcomes is important and that perception of individuals may be different from person to person. Examples of survey comments related to the presentation of data:

- "Transparent, internally-consistent & reproducible justifications of why a specific confidence level was chosen (over another) is paramount. Without it, you are just introducing an unfounded 'quality factor' at best, or a 'fudge factor' at worst, in the risk management decisionmaking."

- "Wouldn't you normally expect to present various outcomes and their related confidence levels? It would seem to me that knowledge of all possible outcomes and the related confidence levels would be important to the decision maker."

- "I think these types of qualitative scoring are more helpful than nothing but there are some very good papers out there that show that people, including risk managers interpret them very differently -- High to one person might be 75%, to another 99%. (c.f. papers by Cox in Risk Analysis if I remember correctly)These probabilities might lead to very different decisions depending on the context. So I think you have to be very careful to understand how sensitive to small changes in probability decisions might be. I have a preference for getting assessors to put quantitative probabilities, even if they're ranges or distributions."

Different Approach

There were three (3) comments what were grouped together as related to different approaches that might be taken. These ad-hoc comments address three alternate themes for research. The first is that examining the real outcome versus the predicted may be a topic of research. The second example discusses using a range of values versus a single point. This is similar to the work by Malone and Moses as shown in Figure 8. (Moses and Malone, Jr. 2005 p. 15). The third example suggests another approach, looking at "shiftability". Examples of survey comments related to possible different approaches:

- "This survey misses the point. Risk assessors usually believe too strongly that they understand what is going on. The real question is whether or not real life will bear out their predictions. That is where research needs to be done, not whether risk assessors believe in their own work."

- "When I ask for risk data to be presented I often ask for a range of values and not a single point. For example I will get a likelihood range (1-3) not just 2. This gives me a idea of the confidence in the number."

- "If by 'confidence' you mean how firm the probability or other measure of uncertainty is, the value of assessing it depends on whether there is a decision to be made on gathering new information about the risk (as opposed to deciding with the information you have). It would be more useful to define confidence levels in terms of 'shiftability', which does not necessarily correspond to the TYPE of information the assessment is based on. See Brown, R.V. Assessment uncertainty technology for making and defending risky decisions. Journal of Behavioral Decision Making, 3, 213-228, 1990."

Informational

There were fifteen (15) comments that were grouped together because the provided additional information. These ad-hoc comments provided additional perspectives on risk assessment. Examples of survey comments that provided additional information:

- "In a project Risk assessment, the facts are discussed until a consensus L and C assessment is reached. Confidence levels might help if there were a practical way

to improve upon 'expert judgment' or unvalidated model analysis. Typically we use the best data available - and if the better data aren't there, then the decision are made on the data we have. I'm not sure a credible risk would be treated differently depending on how confident we were in the assessed L and C - as long as the consensus was that this is the best assessment possible in this situation. So the confidence level in the 'Best' assessment is more important than the confidence level in any one assessment 'box' in the matrix. In other words, if the assessment looks like a 2-D uncertainty distribution (in L and C), then the confidence that the centroid or expected value hovers over the right box is more important than how wide the tails of the distribution are."

- "Risk is inherently due to uncertainty. Uncertainty comes in two forms: aleatoric and epistemic. Classical probability concepts can be used in assessing aleatoric phenomena (e.g., performance variations for a task). However, they are not applicable to epistemic uncertainty where the 'unknown unknowns' lurk. The concept of 'confidence' seems restricted to the aleatoric phenmena similar in nature to confidence limits in probability theory. Reference David Hillson."

- "In order for confidence levels to have meaning to decision makers they must be translated into solutions. e.g., a risk management solution of attaining more information would results in a high priority solution when the alternatives to reduce risk are costly and uncertain. There are also confidence levels in the efficacy of risk management. This is really where the confidence and information gathering alternative come into play. If you are uncertain about original consequences, how can you be certain about the efficacy of an investment structured to manage those consequences? "

- "I've assumed that there is a single risk assessor providing assessments. In our quantitative risk analysis work we facilitate expert assessment inputs to provide a range of outcomes for project risk management. Single points numbers, assessments and other inputs are essentially useless in determining the confidence in an uncertainty assessment, in our experience. Quality decisions are best made when the potential impact of a negative event is fully understood."

- "The 'reasons' for the risk assessor's confidence are also important (even if not measurable)."

- "I have found that you cannot avoid some level of subjectivity when assigning risk."

- "While the addition of a confidence score looks potentially useful, this would have to be tested before I could tell for sure. I get a lot of value out of the risk management thought process, and the confidence score may not matter much for that. Also, I would not want to diminish the importance of assessments that might

have low confidence, so how to use confidence scores is a little less straight forward to me."

- "1. Where possible, the establishment of confidence levels should have Project Team consensus after suitable discussion and education in their meaning, effect etc. 2. The actual results should be reviewed against the original estimate of confidence levels. In my experience Engineering estimates of duration of projects are usually understated by a factor of three."

- "Sometimes you can determine the subjective confidence level of the risk assessor as they verbally present the risk to the decision-makers- through their body language or their confidence while presenting the risk."

- "I have worked in the Risk arena for more than 15 years. Experience is what gives decision makers reliance on risk models. I think that having experience in the real world, seeing the actual sites and assessing the potential for exposure is more important than any risk models based on such constructs."

- "In my work, I attempt to bound uncertainty and variability for the decision maker and project team through scenario analysis. For complex, long-tail risk problems, scenario analysis is a better tool to inform decision makers by making confidence issues explicit."

- "Next to last statement: The risk assessor's confidence in the score assessment SHOULD BE important, but seldom is. Similarly, for the last statement: it CAN BE useful, but might not actually be used."

- "Modeling and test often will provide stronger confidence in the risk scoring than past experience and analysis, but not always. The important thing is for the decision maker to understand the basis of the risk scoring. For example, programmatic risks are not typically amenable to testing."

- " You've added a whole another layer of complexity -- certainty about these two elements --that is way beyond the way anyone thinks. It's very important but light years away from what we do here in trenches. "

- "This is difficult stuff. Telling a risk manager how confident you are in the results is an important responsibility, but it may not be easy for the RM to understand the rationale behind it. Also, we don't want risk managers to not make decisions just because the confidence is low, we just want them to appropriately weight the information from the risk assessment as one of many factors."

Other Factors

There were five (5) comments that were grouped together that discussed other factors related to the survey topic. These ad-hoc comments point out that there are additional factors either internal or external that drive the decision maker. Examples of survey comments that identified other factors to consider:

- "Had to score somewhat incongruently for a key reason: While the risk confidence criteria would be quite useful in mgmt making an informed decision, that decision process often is driven by important external factors or limiters (such as budget pressures for NASA or a rapid market shift in private sector)."
- "The confidence level also depends on the maturity and stability of the project or entity indentifying the risk."
- "While confidence in the assessment can be factor it is often secondary to other factors not related to the risk being estimated. The relative weight of these other factors can often make confidence statements of limited value."
- "The survey adequately focuses on the role of the risk assessor, but answers to some of the questions (4,6,7 in particular) require an assumption about the abilities and motivation of the ultimate decision maker. The less objective the decision maker, the less meaningful the CI. This variation should be accounted for in the model."

Practicality

There were nine (9) comments that were grouped together that discussed the practicality of the research. These ad-hoc comments specifically address the survey participants concern that confidence levels add to the complexity of the risk process or will not improve the decision process. Examples of survey comments that discussed the practicality of the research proposal:

- "Providing confidence levels would not improve management decisions as some decisions are for the most part subjective."
- "In my opinion, the intentions behind attempting to define and quantify confidence levels for likelihood and consequences are good. However, the

confidence level definitions provided here are quite subjective and would be ineffective (in my opinion) without further questioning of the risk assessor by the decision maker. I believe the decision maker should always invoke ample questioning in order to make a truly informed decision. Good decision makers will probe to the appropriate level of understanding before making their decision, regardless of whether confidence levels are quantified by the risk assessor or not. Thus, I am not convinced that asking the risk assessor to quantify confidence levels is of much added value to the decision making process for good decision makers. By the same token, I question the value of this for poor decision makers since poor decision makers may not ask additional probing/follow-up questions needed to effectively make a decision."

- "A confidence is another subjective evaluation of something that is already subjective (risk scoring). The better approach, in my opinion, is to ensure rationale for LxC scoring (and changes) is always captured within the risk documentation. Each person reviewing the risk can then decide for themselves whether the scoring is appropriate, and whether they have confidence in the scoring applied. Additionally, scoring is typically reviewed at the time the risk is evaluated for validity and may be changed at that time. I see no value in adding a confidence level to qualitative risk assessments."

- "Having recently read Taleb's 'The Black Swan', and looking at the literature of 'unintended consequences' I believe there are relatively few risk assessments that satisfy the conditions where confidence intervals can be well specified. Especially in regimes where human behavior is involved."

- "Personal professional observations indicate a disparity between theoretical and pragmatic uses of the subject data. From a theoretical perspective, risk management and confidence level assessment lends significant values. From a daily practitioner perspective (in the majority of cases), risk and confidence level assessments are considered only on 'high' consequence decisions."

- "Probably adds value, but adds another layer of complexity to and already complex process (risk management)."

- "Often the consequence(s) cannot be adequately envisioned in advance in order to support such as analysis."

- "Comments: Q1 - Note, What use is the assessment if there no or little confidence in it? Low and medium confidence levels are therefore relatively meaningless. Q3&Q7 - Confidence level definitions at the detail described in the survey add an un-necessary level of complexity to the assessment. Knowing the assessors overall confidence level is meaningful."

Relationship to Consequence and Likelihood

There were two (2) comments that were grouped together that discussed the relationship of confidence levels to consequence and/or likelihood scores. These ad-hoc comments address questions on the relationship between consequence, likelihood and the confidence levels. Examples of survey comments that discussed the practicality relationship to consequence and likelihood:

- "Sounds like you are trying to measure how strongly of a score you have, such as how '5' is it. How does a Likelihood of 5 with a Confidence of Low (5/Low), relate to a Likelihood of 4 with a Confidence of High (4/High)? I can believe that a 4/High is more significant to me than 5/Low. If that is not accurate, then perhaps the measure of confidence levels needs to have more clearly defined striations in the measures."

- "It's not clear how these rank within consequence...for example, how does a consequence assessment of 3 with low confidence rank against a consequence assessment of 2 with high confidence....I suspect if this is not well defined and ironed out, many managers will interpret and use the scale differently...I think it's very grey and these lines I drew in the example cross, just need to be sure folks are on the same page and how much they cross."

Risk Matrices Issues

There were three (3) comments that were grouped together that specifically reference an article titled "What's Wrong with Risk Matrices?" by L. A. Cox Jr. (Cox, Jr. 2008). These ad-hoc comments were of interest because they focused on a relatively new article by Tony Cox that questions the overall use of the risk matrix as it is commonly used today. Examples of survey comments that discussed the recent literature on issues with risk matrices:

- Have you read 'What's Wrong with Risk Matrices?' Tony Cox, Jr. Risk Analysis, Vol 28, No 2, 2008?? Think top down and put your work into context - risk matrices have their problems so concentrating on one aspect of them may not significantly improve decision making?!

- "Interesting and important work. I assume you've read and understand Tony Cox's critique of frequency/consequence risk matrices (Risk Analysis 28(2):497-512, 2008). Good luck!"

- "Not sure about matrix-based qualitative approaches. See Cox, LA. 2008. 'What's wrong with Risk Matrices' Risk Analysis 28:497-512. If a matrix itself has limited accuracy, putting CLs on it likely won't improve the outcome. Maybe try to apply a quantitative model to show it can work?"

Subjectivity

There were three (3) comments that were grouped together that addressed subjectivity. These ad-hoc comments note the subjectivity involved in the risk assessment process. Each comes at the subject from different perspectives. One comment notes that confidence levels are a method to formalize some of the subjectivity and increase trust. A second comment is totally against the use of confidence levels as a subjective measure and the third points out that the effort to focus on a better risk score may not be worth the effort. Examples of survey comments that discussed subjectivity:

- "Risk judgment, and risk management decisions will always have a subjective component, depending on the assessor. The risk confidence levels are an attempt to formalize the efforts to minimize the subjectivity bias and in the end to raise the trust."

- "Qualitative confidence level has limited usefulness in risk management decisions."

- "The Risk Assessment Criteria are 1) subjective and 2) defined by a range of values. So they are already characterized by ambiguity. Adding a confidence factor does not significantly reduce the level of ambiguity. When I score a risk as a 3x3, I understand that it could accurately scored in any of the neighboring boxes on the 5x5 matrix. But calling it a 3x3 gives me enough information to make a decision and move forward. The expense of collecting more information on the likelihood or consequence for the purpose of improving the risk severity score is hardly worth the effort. Such research is best applied toward finding the best handling approach."

Support

There were eighteen (18) comments that were grouped together that provided words of support for the research. These ad-hoc comments were supportive of one or more aspects of the proposal for confidence levels. The number of these comments is supportive of the overall statistics showing agreement that the risk assessor's confidence level in the assessment is important to the decision maker. Examples of survey comments that provided support to the survey proposal:

- "Thanks for distributing this survey to the risk community. I believe any time we are asked to stop and think about the questions you raise, it improves the entire risk community. Thanks!"

- "Your means of expressing risk and confidence is very simple and certainly if the internal team is of a technical background. It's difficult to comment more as you have provided very little that describes the context of your work or about who may populate the team."

- "Although you have nicely defined risk assessor confidence levels, you have left me to guess that a likelihood of 2 corresponds to 'medium' and a consequence of 3 corresponds to 'high'. At any rate those were my assumptions."

- "This is an interesting way of communicating confidence/uncertainty in qualitative risk assessment (risk scoring). My concern is that introducing the concept of confidence will further encourage using the risk score as a quantitative result, when it generally is not."

- "In my experience, the estimation and communication of the uncertainty in the scoring of an event's likelihood and consequence generally is overlooked and it shouldn't be overlooked. "

- "The problem is that most managers do not understand confidence levels. Many do not even ask what the confidence level is when likelihood is present. The three levels proposed above are probably more helpful than the usual confidence levels that come out of statistical analysis."

- "I think this is a very interesting idea to add the risk confidence level as another dimension to the Risk Matrix."

- “Risk Assessment is the probability of an adverse event occurring. It should be a quantitative number with calculated confidence intervals. It seems that the above questions are related to semi quantitative Risk Assessments. For the less quantitative risk assessments, an estimation of confidence level could be useful if it was applied uniformly by risk assessors.”

- "Our organization requires us to make decisions based on facts (ISO 9000). The confidence level provides a very useful measure of quantifying the level of uncertainty on various 'Facts' provided in making decisions. For example, I have a bridge that provides documentation as to pile depths. This has been undisputed until more recent evidence has been provided questioning the original data provided. The pile depths provided is not questioned - the level of confidence that the pile depths are accurate has been reduced from 'High' to 'Low'.”

- “This aids both qualitative and quantitative risk assessment”

- “The confidence level gives reassurance to the assessor that the likelihood or consequence is solid. But without it the +/- on the score most likely would not move much. It is not mandatory in making a decision but results in higher confidence in turn of the decision you make.”

- "Even with modeling and simulations, there is still some degree of subjectivity involved. Agreed, this is a useful decision making tool, however, too often decisions are based on 'gut feeling' rather than data. To some extent, maybe it is human nature that people feel they can beat the odds, even when presented to them - just look at Las Vegas and the gambling industry."

- "Confidence Level Definitions - I don't see a large difference between 'Relative Past Experience' vs 'Similar Conditions being observed previously'. I'd think relative past experience would be more of a Medium category. Managers usually draw out the confidence level of the risk matrix by asking questions, and determining what the levels were based on; ie what went into the decision for a 3X2 vs a 3X3. I'd say the confidence level is necessary in making a decision, it's just how does the presenter explain it. If the confidence level is low, then more resources or better qualified people are needed to assist in the review of the risk matrix before it is presented. I'm not sure I'd spend a significant amount of money to improve the confidence level above a Medium level, as those resources could be used to fix the problem instead."

- "While I generally agree with the approach and definitions for confidence level they may be too specific. Another way to approach this is describe it in terms of lines of evidence Insight --> direct empirical, and coherence across lines of evidence. I suggest looking at Crawford-Brown (2006) for a framework which may be useful. Another way to describe this which may be more meaningful is whether the results would likely change with additional information. Low would

imply that additional information may greatly impact the assessment and high would mean the estimate is fairly robust and would not likely change. One needs to consider that the basis of information such as direct empirical versus theoretical insight does not always imply that they are good. It is possible that even an estimate based on direct empirical evidence may change with additional data collection. By contrast, estimates based on indirect empirical evidence and theory may not change with additional information if there is a strong c"

- "I personally believe that use of an uncomplicated risk confidence level score should be mandatory BUT (1) it has to be very simple - too many options doesn't work. Also how well the risk management decision will rely on confidence in the risk assessment depends on issues such as level of the manager making the decision (the higher up the less likely the manager is likely to pay attention) and how 'political' the risk management decision is. Good luck with your research."

- "In the end, having an idea of the risk evaluator's confidence is useful; however it is still high subjective. The decision maker must have confidence on the evaluator's assessment, and even then the decision makers experience will weigh heavily on how the confidence level is used. Please don't misinterpret - I think you are on the right track here, the trick is to determine how to move this thought process to a more objective measure in the future."

The ad-hoc comments provided a wide-range of ideas and opinion across the risk management process. Several points are noted as being mentioned through-out several of the comments. There are several comments that note from their view, the confidence levels as presented add complexity to the process and are not required because the decision maker will always assume that the risk assessor has high confidence in the assessment. Others note the usefulness of the confidence levels, but note that looking at the number of levels and wording could improve them. A common theme is the subjectiveness of the risk management process itself and the need to document the basis of the assessment for the decision maker. Related to this is the concept of providing a range of values instead of point values. Overall, the ad-hoc comments provide a range of ideas that would be very useful in continuing the research into improving the definition of confidence levels.

Data Analysis Summary

The overall response to the survey was sufficient to make conclusions as to the usefulness of confidence levels by providing data on both their importance and usableness. The 364 completed surveys resulted in a confidence level of 95% with a range of errors from 3.84 to 4.97%. There was an excellent range of roles indicated by the survey participants, but of particular note is that approximately 50% indicated their role as a program/project manager which would be the most focused role for using confidence levels (Reference Figure 17). The participant experience level was good with the average being in the equal to or great than 5 years and less than 10 years range. The survey did not have enough fidelity to narrow the range any lower than this in order to find out if the actual experience level was at the low or high end of that scale.

The overall statistics as shown in Table 13 are the most telling as to the results of the survey. Based on breaking the data into negative, neutral and positive responses, Questions 1, 3, 4 and 6 which are related to importance had a range centering around 80% positive scoring. Questions, 2, 5 and 7 which are related to usableness scored much lower and with more of a range. Specifically, Question 2 which relates to the understandability of the confidence level definition scored at 79% positive which would indicate that some improvement is necessary on the actual confidence level definition. Question 5 which scored lowest of any question on the survey at 60% indicates that the actual definitions for the confidence levels are in need of improvement. Question 7 which scored highest of any question on the survey noted that using risk confidence levels is useful when combined with the likelihood and consequence scoring.

The overall evaluation of the usableness factor is that improvement in the definitions of the actual confidence levels will be necessary to implement them.

A non-parametric Kruskal-Wallis was conducted for additional understanding of whether the makeup of the roles or experience levels of the survey participants may affect the survey results. For six of the seven questions, the results of the Kruskal-Wallis test failed to reject the hypothesis that the distributions of the survey mean responses were identical. The exception was Question 1 where there are differences based on role and by inspection of the responses, it appears that decision makers, academics, and others more strongly support the need for confidence level information to reduce the difficulty in making risk based decisions in projects.

The ad-hoc comments provided in the survey were analyzed and broken into eleven groups for analysis. The comments added additional insight into the survey results and provide useful information for future research efforts to refine the definition of confidence level and its definition.

Chapter 5 presents the conclusions from the research in this dissertation.

CHAPTER 5: CONCLUSION

This chapter summarizes the conclusions drawn by this dissertation.

The purpose of this research was to define risk assessment confidence levels for use in internal project risk management communications. In order to achieve this, the research has:

- Demonstrated that there exists a gap in the existing literature addressing the communication of confidence levels in risk assessments.
- Defined confidence level as "The degree of certainty that the likelihood or consequence score (assigned by the risk assessor) reflects reality."
- Defined confidence levels:

Low Confidence Level:	Risk assessment is based on subjective opinion or relative past experiences.
Medium Confidence Level:	Risk assessment is based on similar conditions being observed previously and/or analysis. Analysis is based on unverified models and/or data.
High Confidence Level:	Risk assessment is based on testing, high fidelity modeling or simulation, use of prototype or experiments. Analysis is based on verified models and/or data. Assessment is based on an historical basis.

- Provided evidence that confidence levels would be useful to a representative sample of risk assessment professionals from industry and academia.

A comprehensive literature review was conducted, providing the reader with an overview of the existing literature as it relates to the subject of risk, risk communication, decision analysis,

confidence level, uncertainty, risk management and risk analysis. Table 4 classifies the literature cited in this dissertation according to the subjects they covered. The observation resulting from the literature review was that gaps exist in the existing literature regarding the use of confidence levels in project risk communication, definitions for confidence levels and data to show the usefulness of confidence levels.

After the literature review was completed, a preliminary definition of confidence level and definitions for high, medium and low levels of confidence were developed. A survey instrument was developed based on work by Larcker and Lessig (Larcker and Lessig 1980) to test the usefulness of the definitions, as determined by the attributes of importance and usableness. A focus group consisting of participants with experience in aerospace, the shipping industry and academia was used to evaluate the survey instrument. Modifications were made to the survey instrument based on the focus group observations and a draft survey was sent to a pilot group of participants consisting of representatives from industry and academia. Response rate from the pilot survey was only 8% due to bounce back and SPAM/firewall filtering which resulted in an increase in the number of surveys to be sent out for the primary survey.

The desired confidence level for the survey was 95% with a margin of error of 5%, which require a sample size of 384 for a large population. Based on the pilot survey, it was determined that a minimum size of 4, 800 surveys would be required. Actually 5,302 surveys were sent out. As noted in Chapter 4, for the purposes of this survey, response rate was calculated based on surveys that were viewed but not completed and those that were completed. Based on the receipt of 364 surveys instead of the 384 desired, the confidence level at 95% has a range of errors from 3.84 to 4.97% which was considered sufficient for this analysis. There were two (2) demographics collected as part of the survey which was the role of the participant and experience

level. The data collected is summarized in Table 10 and Table 12. The survey participants indicated that they had a role as a program/program manager in 50% of the responses. An examination of the roles of the survey respondents indicated that no single group was dominant. Based on the data summarized in Table 10 there was sufficient participation in roles such as program/project manager, risk manager, risk assessor and decision maker to validate the results. The respondents represented a broad variety of roles associated with decision making and risk management with experience levels from fairly new to experienced risk assessors. The experience level on average was equal to or greater than 5 years, but less than 10 year which was sufficient experience to validate the results. Once it was determined that the survey responses were sufficient in number, roles and experience, the initial analysis was to examine the seven questions for negative, neutral and positive responses. Note that Question 3 was negatively phrased it was necessary to reverse score this for the purposes of this analysis. This calculation consists of counting the number of responses (negative – 1, 2; neutral – 3; positive – 4, 5) and then dividing by the total to find a percentage. Table 13 shows the summary of all seven (7) questions. Since the analysis is looking at two attributes (importance and usableness), the results as shown in Table 13 were broken down in Table 14 by these attributes.

The survey results show that for the attribute of importance, there was an overall result that the participants thought positively that knowing the risk assessor's confidence in their assessment was important in making a decision as demonstrated in Questions 1 and 6 (80%, 83% respectively). Also in a positive response, while not as strong as the other two questions on the attribute of importance, there was a positive response that confidence levels would improve a management decision and were meaningful to the decision maker as shown in Questions 3 and 4 (78%, 77% respectively). From an overall look across all four (4) questions related to the

importance attribute, there was a positive response that confidence levels are important and meaningful to the decision maker.

The survey results showed for the attribute of usableness, there were mixed results. Questions 2 and 7 had an overall positive response (79%, 86% respectively), while Question 5 was narrowly a positive response (60%). Questions 2 and 7 address whether the definition of "confidence level" was understandable and whether confidence levels were useful in supplementing likelihood and consequence. Question 5 directly addresses whether the confidence levels as presented adequately described the differences in the assessment confidence. Based on this result, the confidence level definition itself is usable, but the confidence level definitions are not in their current form. In addition to these results, there were other descriptive statistics collected, but not used for the analysis.

The survey's reliability was calculated using Cronbach's alpha. The calculation resulted in a reliability of 0.77. The generally accepted value for reliability is 0.70 or higher. Other descriptive statistics were collected that included data such as mean, median, standard deviation etc., as shown in Figures 20 - 26. An area of interest in analyzing the data was to examine if the role of the survey participant or their experience level affects the results of the survey. For the purposes of this research, a series of non-parametric Kruskal-Wallis tests were performed. As shown in Table 17 for role, for six of the seven questions, the results of the Kruskal-Wallis test failed to reject the hypothesis that the probability distributions of the survey responses were identical. The exception was Question 1 which based on role and by inspection of the responses, it appears that decision makers, academics, and others more strongly support the need for confidence level information to reduce the difficulty in making risk based decisions in projects

As shown in Table 18 for experience, for all seven questions, the results of the Kruskal-Wallis test failed to reject the hypothesis that the probability distributions of the survey responses were identical.

The last data set were the ad-hoc comments that were entered by the survey participants into the free text block on the survey. Comments were provided by 28.8% of the survey participants. As described in Chapter 4, the entire comment set was examined and some comments were discarded as they were deemed as general and not providing any additional information useful in the analysis phase leaving 86 comments from 23.6% of the survey participants.

The usefulness of the comments is adding insight into the survey results and also in providing information for future research efforts to refine the definition of confidence level and its definition.

Based on the results of the survey, it is clear that the majority of those surveyed found that knowing the confidence the assessor has in their assessment is important and would improve a management decision and would agree with the definition what a confidence level is, but would not agree with the definitions of the individual confidence levels as presented. Overall, there is agreement that confidence levels are useful if the attributes of importance and usableness are taken together.

Based on the research there is a basis for a new tool for a decision maker to implement for communication. The tool needs to be refined, but in the end, it will complement existing tools such as the risk matrix, likelihood and consequence scoring matrices. The first need is to refine the definition of what a "risk confidence level" is. The idea of renaming it as "risk assessment confidence level" or "risk assessor's confidence level" has merit. Within the

comments, there was a recommendation to rename it as “risk certainty” or even “risk uncertainty” which could be examined further. The research provides evidence that the definition as presented “The degree of certainty that the likelihood or consequence score (assigned by the assessor) reflects reality” was understandable to a majority of the respondents, but given data, the definition should be revisited to see if it can be re-written to make it more understandable. What needs further work is the definitions of the confidence levels themselves. The research data shows that the definitions as stated are not perceived as adequate. A case can be made that this should not be an academic exercise, but taken on at the implementation level because these definitions, both the number of levels and the actual descriptions may be tailored to a specific industry or even a program or project. While three levels of definition would seem to be enough, from the comments there may be rationale for four levels. Just as the likelihood and consequence scoring is normally tailored to a project or program, the definition of the confidence levels should be also tailored. This research presented one set of definitions for the confidence levels themselves, but also assumed that a confidence level would be assigned to the likelihood score and another to the consequence score. This adds to the complexity of the communication of the score. Further research should be done to look at whether an aggregated confidence score could be defined that would take both the likelihood assessment confidence and consequence assessment confidence into account. This research has demonstrated a basis for further development of confidence levels as a risk management tool for use in internal project communications.

Contributions Of Research

This research has demonstrated that there exists a gap in the existing literature addressing the communication of confidence levels in risk assessments. Once this gap was identified, this research defined confidence level as "The degree of certainty that the likelihood or consequence score (assigned by the risk assessor) reflects reality." A set confidence level definitions were defined as:

Low Confidence Level:	Risk assessment is based on subjective opinion or relative past experiences.
Medium Confidence Level:	Risk assessment is based on similar conditions being observed previously and/or analysis. Analysis is based on unverified models and/or data.
High Confidence Level:	Risk assessment is based on testing, high fidelity modeling or simulation, use of prototype or experiments. Analysis is based on verified models and/or data. Assessment is based on an historical basis.

Using these definitions for confidence level, this research used a survey methodology to provide evidence that confidence levels are useful to a statistically significant sample of industry representatives and academia.

Areas of Future Research

One of the results from any research is the identification of areas that should be further researched. One of the benefits of collecting the ad-hoc comments in the survey was that it enabled survey participants to provide ideas for further research.

The survey results indicate that the actual definitions of the confidence level (provided here as "low," "medium" and "high") need additional development and research. Further research can look at whether the levels of definition should be four levels, as an example. Specifically one comment was that they use of four-levels instead of three takes out the neutrality of a "medium" option which assessors might be prone to rely on most of the time.

Comments were provided on the definitions that would be of interest in trying to modify the definitions of the levels. One comment was that there seemed to be some ambiguity between the definition of low which used "relative past experience" and medium which used "similar conditions being observed previously". Another related comment is that the definitions as presented are a mix of variability, uncertainty and subjectivity.

Whether different levels of granularity in the definitions would improve them is also worthy of additional research. While these specific definitions were designed to be used generally across various industries, it would be of interest to look at whether different definitions would be useful across different industries. Specific examples provided were use in the nuclear industry and complex projects.

A topic of comment that was very interesting was that risk assessors always believe strongly in their assessment. A possible research topic is to follow the history of some assessments from what was proposed by the assessor to what actually resulted from those

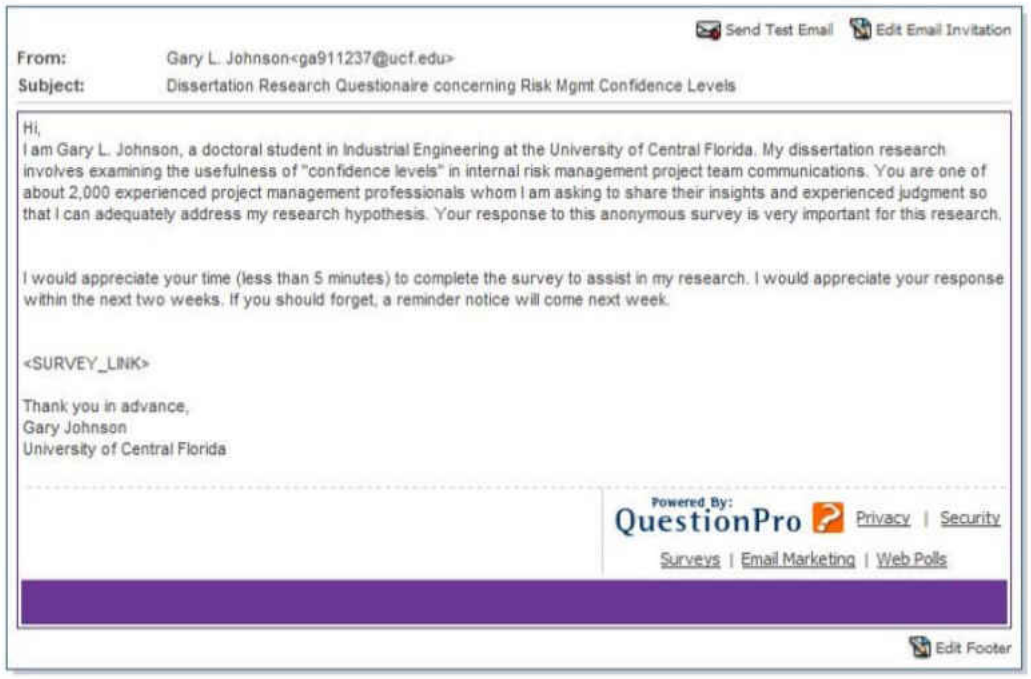
predictions in order to gain some insights into the reality of assessment prediction-reality. An aspect of this which might be interesting to document is the effect of shiftability on these risk assessments as proposed by Brown (1990).

The proposal by Moses and Malone (2005) to present a range of values for likelihood and consequence is of interest and would be interesting for additional research across industries to examine is usefulness.

Several comments were directed towards a recent article by Cox titled "What's Wrong with Risk Matrices?" (Cox, Jr. 2008). The research by Cox would seem to open many areas for future research as to the usability of risk matrices.

APPENDIX A: DRAFT SURVEY INSTRUMENT

Introductory E-mail:



Introductory Text:

Dear Survey Participant:

I am Gary L. Johnson, a doctoral student in Industrial Engineering at the University of Central Florida. I am also a project manager for NASA. My dissertation research involves examining the usefulness of "confidence levels" in internal risk management project team communications. I am asking about 2,000 experienced project management professionals to share their insights and experienced judgment so that I can adequately address my research hypothesis. Your response to this anonymous survey is very important for this research. It will take approximately 5 minutes to complete this survey.

Your participation in this study is completely voluntary. There are no foreseeable risks associated with this project. However, if you feel uncomfortable answering any questions, you can choose to not participate or not answer any specific question at any point.

The results of this study may be published, however your survey responses will be strictly confidential and data from this research will be reported only in the aggregate. Your information will be confidential and will remain confidential. If you have questions at any time about the

survey or the procedures, you may contact Gary L. Johnson at 281-244-7749 or email Gary L. Johnson at ga911237@ucf.edu.

Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (IRB). Questions or concerns about research participation and rights may be directed to UCF Institutional Review Board Office at the University of Central Florida Office of Research and Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246. The phone numbers are 407-823-2901 or 407-882-2276.

Thank you very much for your time and support. Please start with the survey now by clicking on the Continue button below.

Survey Background:

The purpose of this survey is to provide data on the usefulness of "confidence levels" n risk communications.

CONFIDENCE LEVEL – BACKGROUND:

The use of confidence levels for risk communications is a new concept that is being explored in this research.

Risk confidence levels are defined as **"A measure of the degree of surety or certainty that the likelihood or consequence score assigned by the risk assessor reflects reality".**

The use of risk confidence level for the purposes of this survey is LIMITED TO INTERNAL PROJECT COMMUNICATIONS only. Internal communications are defined as those between Program or Project team members (i.e. communication to the public or shareholders would be considered external communication).

The risk confidence level is used in conjunction with the likelihood and consequence assessment and is intended to provide the Program or project team members a sense of the risk assessor's confidence in the likelihood and consequence assessment.

Confidence level definitions are intended to be tailored for each project based on how that project defines likelihood and consequence (3x3 matrix, 5x5 matrix, 1-5, Low-Med-High, etc.) For the purposes of this survey the confidence levels are defined for a 3x3 (Likelihood x Consequence) risk matrix with magnitudes of Low-Medium-High.

Low, medium, or high, risk confidence levels for either likelihood or consequence may be defined as:

Confidence Level Definitions:
Low Confidence Level: Risk assessment is based on subjective opinion or relative past experiences.

Medium Confidence Level: Risk assessment is based on similar conditions being observed previously and/or analysis. Analysis is based on unverified models and/or data.

High Confidence Level: Risk assessment is based on testing, high fidelity modeling or simulation, use of prototype or experiments. Analysis is based on verified models and/or data. Assessment is based on historical basis.

Example and Survey:

EXAMPLE

The risk assessor presents the following risk assessment to the project team:

Likelihood of 2 with a confidence level of "high"
Consequence of 3 with a confidence level of "low"

In the back-up data you would find that the high confidence level for the likelihood of 2 is based on the use of proven high fidelity simulations using verified models. The low confidence level for the consequence is based on the opinion of the risk assessor and technical experts because even thought the events which trigger the risk are well modeled, the resulting effects are not models known with certainty.

The project team will now use this information to decide if additional resources should be allocated to raise the confidence in the consequence assessment to medium or high.

	Strongly Agree	Disagree	Undecided	Agree	Strongly Agree
It would be extremely difficult to complete a risk management decision if you do not know the risk assessor's confidence in the assessment.	O	O	O	O	O
The risk assessor's confidence in the assessment is important in making a risk management decision.	O	O	O	O	O
The risk confidence level, when supplementing the likelihood and consequence ranking, is useful in making a decision.	O	O	O	O	O

Please add any comments and suggestions you might have:

Risk confidence levels are defined as: "A measure of the degree of surety or certainty that the likelihood or consequence score assigned by the risk assessor reflects reality".

Confidence Level Definitions:

Low Confidence Level: Risk assessment is based on subjective opinion or relative past experiences.

Medium Confidence Level: Risk assessment is based on similar conditions being observed previously and/or analysis. Analysis is based on unverified models and/or data.

High Confidence Level: Risk assessment is based on testing, high fidelity modeling or simulation, use of prototype or experiments. Analysis is based on verified models and/or data. Assessment is based on an historical basis.

	Strongly Agree	Disagree	Undecided	Agree	Strongly Agree
The definition of a Risk Confidence Level as presented is understandable.	O	O	O	O	O
Most risk management decisions would be better if the risk confidence level was provided along with the likelihood and consequence.	O	O	O	O	O
The specific confidence level definitions as presented are meaningful to a decision maker.	O	O	O	O	O
The specific confidence level definitions as presented are understandable.	O	O	O	O	O
The specific confidence level definitions as presented adequately describe differences in the assessment confidence.	O	O	O	O	O

Please add any comments and suggestions you might have:

Demographics:

In which role do you have the most experience as related to risk management?

- O Program/Project Manager
- O Risk Manager
- O Risk Assessor
- O Decision Maker
- O Academic
- O I have no role
- O Other

What was the duration in your most experienced role?

- O Less than 1 year
- O >= 1 year and < 5 years
- O <= 5 years and < 10 years
- O >= 10 years

Please contact ga911237@pegasus.cc.ucf.edu if you have any questions regarding this survey.

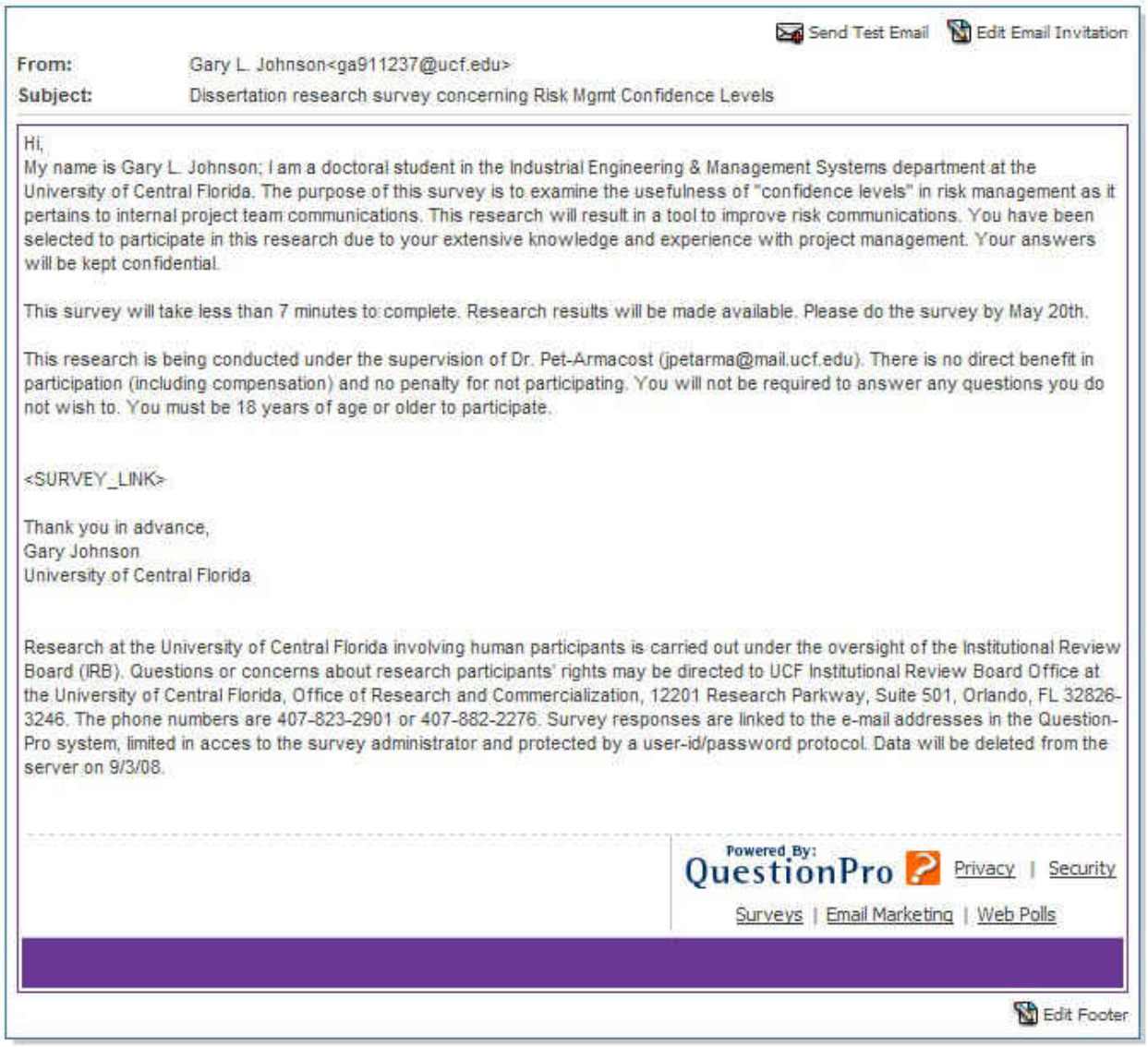
APPENDIX B: FOCUS GROUP WORKSHEET

Introductory e-mail	
1. Is the e-mail clear as to what is being asked of the recipient?	
2. Would recipients be more likely to take or not take the survey based on the e-mail?	
Introductory text – Starts with "Dear Survey Participant"	
3. Is the introductory text clear as to what is being asked of the recipient?	
4. Would recipients be more likely to continue or not continue the survey based on the introductory text?	
Survey Background	
5. Is the Survey Background clear?	
6. Is the Survey Background sufficient?	
Confidence Level – Background	
7. Is the Confidence Level – Background clear?	
8. Is the Confidence Level - Background sufficient?	
Confidence Level Definitions	
9. Are the Confidence Level Definitions clear?	
10. Are the Confidence Level Definitions sufficient?	
Example	
11. Is the Example clear?	
12. Is the Example sufficient?	
Three questions	
13. Is there redundancy in the questions?	
14. How do you feel about the questions?	
15. Are there any negative feelings about the questions?	

16. Does the order of the questions make a difference?	
Reminder of confidence level definition(s)	
17. Is the reminder of confidence level definitions clear?	
18. Is the reminder of confidence level definitions sufficient?	
Five questions	
19. Is there redundancy in the questions?	
20. How do you feel about the questions?	
21. Are there any negative feelings about the questions?	
22. Does the order of the questions make a difference?	
Experience in role related to risk management question	
23. Is the question about experience clear?	
24. Is the question about experience sufficient?	
Duration of that role	
25. Is the question about duration of that role clear?	
26. Is the question about duration of that role sufficient?	
Thank you statement	
27. Is the thank-you at the end sufficient?	

APPENDIX C: SURVEY INSTRUMENT

Introductory E-Mail:



Send Test Email Edit Email Invitation

From: Gary L. Johnson<ga911237@ucf.edu>

Subject: Dissertation research survey concerning Risk Mgmt Confidence Levels

Hi,
My name is Gary L. Johnson; I am a doctoral student in the Industrial Engineering & Management Systems department at the University of Central Florida. The purpose of this survey is to examine the usefulness of "confidence levels" in risk management as it pertains to internal project team communications. This research will result in a tool to improve risk communications. You have been selected to participate in this research due to your extensive knowledge and experience with project management. Your answers will be kept confidential.

This survey will take less than 7 minutes to complete. Research results will be made available. Please do the survey by May 20th.

This research is being conducted under the supervision of Dr. Pet-Armacost (jpetarma@mail.ucf.edu). There is no direct benefit in participation (including compensation) and no penalty for not participating. You will not be required to answer any questions you do not wish to. You must be 18 years of age or older to participate.

<SURVEY_LINK>

Thank you in advance,
Gary Johnson
University of Central Florida

Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (IRB). Questions or concerns about research participants' rights may be directed to UCF Institutional Review Board Office at the University of Central Florida, Office of Research and Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246. The phone numbers are 407-823-2901 or 407-882-2276. Survey responses are linked to the e-mail addresses in the Question-Pro system, limited in acces to the survey administrator and protected by a user-id/password protocol. Data will be deleted from the server on 9/3/08.

Powered By: QuestionPro Privacy | Security

Surveys | Email Marketing | Web Polls

Edit Footer

Consent Statement:

Dear Survey Participant:

The purpose of this survey is to examine the usefulness of "confidence levels" in risk management as it pertains to internal project team communications. This research will result in a tool to improve risk communications.

This survey will take less than 7 minutes to complete. Research results will be made available to you at Research Results

This research is being conducted under the supervision of Dr. Pet-Armacost (jpetarma@mail.ucf.edu). There is no direct benefit in participation (including compensation) and no penalty for not participating. You will not be required to answer any questions you do not wish to. You must be 18 years of age or older to participate.

The results of this study may be published, however your survey responses will be strictly confidential and data from this research will be reported only in the aggregate. If you have questions at any time about the survey or the procedures, you may contact Gary L. Johnson at ga911237@ucf.edu. Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (IRB). Questions or concerns about research participants' rights may be directed to UCF Institutional Review Board Office at the University of Central Florida, Office of Research and Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246. The phone numbers are 407-823-2901 or 407-882-2276. Survey responses are linked to the e-mail addresses in the Question-Pro system, limited in acces to the survey administrator and protected by a user-id/password protocol. Data will be deleted from the server on 9/3/08.

Thank you very much for your time and support. Please start with the survey now by clicking on the **Continue** button below.

Continue

Please contact ga911237@pegasus.cc.ucf.edu if you have any questions regarding this survey.

Survey Background Section:

CONFIDENCE LEVEL - BACKGROUND

Risk confidence levels are defined as: ***"The degree of certainty that the likelihood or consequence score (assigned by the risk assessor) reflects reality".***

Internal communications are defined as those between program or project team members (i.e. communication to the public or shareholders would be considered external communication). The use of risk confidence level for the purposes of this survey is LIMITED TO INTERNAL COMMUNICATIONS only.

The risk confidence level is used in conjunction with the likelihood and consequence assessment and is intended to provide the program or project team members a sense of the risk assessor's confidence in the likelihood and consequence assessment.

Confidence level definitions are intended to be tailored for each project based on how that project defines likelihood and consequence (3x3 matrix, 5x5 matrix, 1-5, Low-Med-High, etc.) For the purpose of this survey the confidence levels are defined for a 3x3 (Likelihood x Consequence) risk matrix using magnitudes of Low-Medium-High.

Risk Assessor's Confidence Level Definitions:

Low: Risk assessment is based on subjective opinion or relative past experiences.

Medium: Risk assessment is based on similar conditions being observed previously and/or analysis. Analysis is based on unverified models and/or data.

High: Risk assessment is based on testing, high fidelity modeling or simulation, use of prototype or experiments. Analysis is based on verified models and/or data. Assessment is based on an historical basis.

Example:

EXAMPLE

The risk assessor presents the following risk assessment to the project team:

Likelihood score of 2 with a confidence level of "high"

Consequence score of 3 with a confidence level of "low"

In the back-up data you would find that the high confidence level for the likelihood of 2 is based on the use of proven high fidelity simulations using verified models. The low confidence level for the consequence is based on the opinion of the risk assessor and technical experts because even thought the events which trigger the risk are well modeled, the resulting effects are not modeled or known with certainty.

The project team will now use this information to decide if additional resources should be allocated to raise the confidence in the consequence score to medium or high.

Survey Questions:

SURVEY QUESTIONS

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
It is difficult to complete a risk management decision if you do not know the risk assessor's confidence in the score assessment.	○	○	○	○	○
The definition of "risk confidence level" is understandable.	○	○	○	○	○
Providing confidence levels would not improve management decisions.	○	○	○	○	○
The confidence level definitions are meaningful to a decision maker.	○	○	○	○	○
The confidence level definitions adequately describe differences in the score assessment confidence.	○	○	○	○	○
The risk assessor's confidence in the score assessment is important in making a risk management decision.	○	○	○	○	○
The risk confidence level, when supplementing the likelihood and consequence scoring, is useful in making a decision.	○	○	○	○	○

Please add any comments and suggestions you might have:

Demographics:

In which role(s) do you have the most experience as related to risk management?

- [ ] Program/Project Manager
- [ ] Risk Manager
- [ ] Risk Assessor
- [ ] Decision Maker
- [ ] Academic
- [ ] Other

What was the total duration of your roles related to risk management?

- [ ] Less than 1 year
- [ ] >= 1 year and < 5 years
- [ ] >= 5 years and < 10 years
- [ ] >= 10 years

APPENDIX D: IRB APPROVAL LETTER

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901, 407-882-2012 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Notice of Exempt Review Status

From: **UCF Institutional Review Board**
FWA00000351, Exp. 5/07/10, IRB00001138

To: **Gary Johnson**

Date: **May 05, 2008**

IRB Number: **SBE-08-05636**

Study Title: **Research Survey concerning Risk Management Confidence Levels**

Dear Researcher:

Your research protocol was reviewed by the IRB Chair on 5/5/2008. Per federal regulations, 45 CFR 46.101, your study has been determined to be **minimal risk for human subjects and exempt** from 45 CFR 46 federal regulations and further IRB review or renewal unless you later wish to add the use of identifiers or change the protocol procedures in a way that might increase risk to participants. Before making any changes to your study, call the IRB office to discuss the changes. **A change which incorporates the use of identifiers may mean the study is no longer exempt, thus requiring the submission of a new application to change the classification to expedited if the risk is still minimal.** Please submit the Termination/Final Report form when the study has been completed. All forms may be completed and submitted online at https://iris.research.ucf.edu.

The category for which exempt status has been determined for this protocol is as follows:

2. Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey or interview procedures, or the observation of public behavior, so long as confidentiality is maintained.

(i) Information obtained is recorded in such a manner that the subject cannot be identified, directly or through identifiers linked to the subject, **and/or**

(ii) Subjects responses, if known outside the research would not reasonably place the subject at risk of criminal or civil liability or be damaging to the subject's financial standing or employability or reputation.

A **waiver of documentation of consent** has been approved for all subjects. Participants do not have to sign a consent form, but the IRB requires that you give participants a copy of the IRB-approved consent form, letter, information sheet, or statement of voluntary consent at the top of the survey.

All data, which may include signed consent form documents, must be retained in a locked file cabinet for a minimum of three years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained on a password-protected computer if electronic information is used. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

On behalf of Tracy Dietz, Ph.D., UCF IRB Chair, this letter is signed by:

Signature applied by Joanne Muratori on 05/05/2008 01:22:20 PM EDT

Joanne Muratori

IRB Coordinator

LIST OF REFERENCES

Reference List

Ahlberg, R. W. (1999). "Risk assessment and risk communication in chemical exposure." *International Archives of Occupational and Environmental Health*, 72(Suppl 3), M11-M13.

Aldag, R. J., and Kuzuhara, L. W. (2002). "Organizational Behavior & Management." South-Western, Cincinnati.

American Bureau of Shipping (2000). "Guidance Notes on Risk Assessment Applications For The Marine And Offshore Oil and Gas Industries." *American Bureau of Shipping Guidance Notes*.

Anderson, J., and Narasimhan, R. (1979). "Assessing Project Implementation Risk: A Methodological Approach." *Management Science*, 25(6), 512-521.

Andrews, C. J., Hassenzahl, D. M., and Johnson, B. B. (2004). "Accommodating Uncertainty in Comparative Risk." *Risk Analysis: An International Journal*, 24(5), 1323-1335.

Apostolakis, G. E. (2004). "How Useful is Quantitative Risk Assessment?" *Risk Analysis: An International Journal*, 24(3), 515-520.

Atman, C. J., Bostrom, A., Fischhoff, B., and Morgan, M. G. (1994). "Designing Risk Communications: Completing and Correcting Mental Models of Hazardous Processes, Part I." *Risk Analysis: An International Journal*, 14(5), 779-788.

Aven, T., and Pitblado, R. (1998). "On Risk Assessment in the Petroleum Activities on the Norwegian and UK Continental Shelves." *Reliability Engineering*, 61, 21-29.

Aven, T., and Vinnem, J. E. (2007). "Risk Management with Applications from the Offshore Petroleum Industry." Springer-Verlag, London.

Banon, H. (1994). "Assessing Fitness for Purpose of Offshore Platforms. II: Risk Management, Maintenance, and Repair." *Journal of Structural Engineering*, 120(12), 3613-3633.

Basu, R., and Lee, A. (2006). "A Flexible Approach to the Application of Risk-Based Methods to the Inspection of Hull Structures." *ABS Technical Papers 2006*, 21-29.

Ben-David, I., and Raz, T. (2001). "An integrated approach for risk response development in project planning." *The Journal of the Operational Research Society*, 52(1), 14-25.

Bernstein, P. L. (1996). "Against The Gods: The Remarkable Story of Risk." John Wiley & Sons, Inc., New York.

Bostrom, A., Atman, C. J., Fishhoff, B., and Morgan, M. G. (1994). "Evaluating Risk Communications: Completing and Correcting Mental Models of Hazardous Processes, Part II." *Risk Analysis: An International Journal*, 14(5), 789-798.

Bradt, D. J. (2004). "Risk Management Communication Barriers [for human space flight]." American Institute of Aeronautics and Astronautics, 1801 Alexander Bell Drive, Suite 500, Reston, VA, 20191-4344, USA, http://www.aiaa.org.

Brown, R. V. (1990). "Assessment Uncertainty Technology for Making and Defending Risky Decisions." *Journal of Behavioral Decision Making*, 3, 213-228.

Carbone, T. A., and Tippett, D. D. (2004). "Project Risk Management Using the Project Risk FMEA." *Engineering Management Journal*, 16(4), 28-35.

Componation, P. J., Utley, D. R., and Swain, J. J. (2001). "Using Risk Reduction to Measure Team Performance." *Engineering Management Journal*, 13(4), 27-34.

Conrow, E. H. (1995). "The Use of Ordinal Scales in Defense Systems Engineering." Proceedings of the 1995 Acquisition Research Symposium, 455-463.

Conrow, E. H., and Fredrickson, M. A. (1996). "Some Considerations for Implementing Risk Management in Defense Programs." *PM*, 6-11.

Covello, V. T., Winterfeldr, D., and Slovic, P. (1986). "Risk Communication: A Review of the Literature." *Risk Abstracts*, 3, 171-182.

Cox III, E. P. (1980). "The Optimal Number of Response Alternatives for a Scale: A Review." *Journal of Marketing Research*, 17(4), 407-422.

Cox, L. A., Jr. (2008). "What's Wrong with Risk Matrices?" *Risk Analysis: An International Journal*, 28(2), 497-512.

Cox, L. A. Jr., Babayev, D., and Huber, W. (2005). "Some Limitations of Qualitative Risk Rating Systems." *Risk Analysis: An International Journal*, 25(3), 651-662.

Cross, R. B., and Ballesio, J. E. (2003). "A Quantitative Risk Assessment Model for Oil Tankers." Society of Naval Architects and Marine Engineers, Society of Naval Architects and Marine Engineers Conference.

Defense Systems Management College (1989). "Risk Management Concepts and Guidance." *Rep. No. ADA214342*, Defense Systems Management College.

Defense Systems Management College (1990). "Systems Engineering Management Guide." *Rep. No. ADA223168*, Defense Systems Management College.

Dey, P. K. (2002). "Project Risk Management: A Combined Analytic Hierarchy Process and Decision Tree Approach." *Cost Engineering*, 44(3), 13-26.

Dillman, D.A., Tortora, R.D., Bowker, D. (1998). "Principles for Constructing Web Surveys.". Available at http://survey.sesrc.wsu.edu/dillman/papers.htm (Mar. 15, 2007).

Edwards, A., Elwyn, G., Covey, J., Matthews, E., and Pill, R. (2001). "Presenting Risk Information - A Review of the Effects of "Framing" and other Manipulations on Patient Outcomes." *Journal of Health Communication*, 6, 61-82.

Emblemsvag, J., and Kjolstad, L. E. (2002). "Strategic Risk Analysis--A Field Version." *Management Decision*, 40(9), 842-853.

EPA (2004). "An Examination of EPA Risk Assessment Principles and Practices." *Rep. No. Report EPA/1-B-04-001*, Environmental Protection Agency.

Finkel, A. M. (1994). "Stepping Out of Your Own Shadow: A Didactic Example of How Facing Uncertainty Can Improve Decision-Making." *Risk Analysis: An International Journal*, 14(5), 751-761.

Gifford, W. E., Bobbitt, H. R., and Slocum, J. W., Jr. (1979). "Message Characteristics and Perceptions of Uncertainty by Organizational Decision Makers." *The Academy of Management Journal*, 22(3), 458-481.

Greenberg, H. R., and Cramer, J. J. (1991). "Risk Assessment and Risk Management for the Chemical Process Industry." Van Nostrand Reinhold, New York.

Grose, V. L. (1987). "Managing Risk: Systematic Loss Prevention for Executives." Prentice Hall, Englewood Cliffs, New Jersey.

Haimes, Y. Y., Kaplan, S., and Lambert, J. H. (2002). "Risk Filtering, Ranking, and Management Framework Using Hierarchical Holographic Modeling." *Risk Analysis: An International Journal*, 22(2), 383-397.

Hassenzahl, D. M. (2006). "Implications of Excessive Precision for Risk Comparisons: Lessons from the Past Four Decades." *Risk Analysis: An International Journal*, 26(1), 265-276.

Hillson, D. (2003). "Using a risk breakdown structure in project management." *Journal of Facilities Management.*, 2(1), 85-97.

Ibrekk, H., and Morgan, M. G. (1987). "Graphical Communication of Uncertain Quantities to Nontechnical People." *Risk Analysis: An International Journal*, 7(4), 519-529.

Jaafari, A. (2001). "Management of Risks, Uncertainties and Opportunities on Projects: Time for a Fundamental Shift." *International Journal of Project Management*, 19(2), 89-101.

Jardine, C. G., and Hrudey, S. E. (1997). "Mixed Messages in Risk Communication." *Risk Analysis: An International Journal*, 17(4), 489-498.

Joint Technical Committee OB/7 . Australian Standard: Risk Management. AS/NZS 4360:1999. 1999. Strathfield, Australia, Standards Association of Australia.

Kahlor, L., Dunwoody, S., Griffin, R. J., Neuwirth, K., and Giese, J. (2003). "Studying Heuristic-Systematic Processing of *Risk Communication*." *Risk Analysis: An International Journal*, 23(2), 355-368.

Kangari, B., and Riggs, L. S. (1989). "Construction Risk Assessment by Linguistics." *IEEE Transactions on Engineering Management.*, 36(2), 126-131.

Kanner, M. D. (2005). "A Prospect Dynamic Model of Decision-Making." *Journal of Theoretical Politics*, 17(3), 311-338.

Kaplan, S., and Garrick, B. J. (1981). "On The Quantitative Definition of Risk." *Risk Analysis: An International Journal*, 1(1), 11-27.

Klinke, A., and Renn, O. (2001). "Precautionary principle and discursive strategies: classifying and managing risks." *Journal of Risk Research*, 4(2), 159-173.

Knuth, B. A., and Connelly, N. A. (1998). "Evaluating Risk Communication: Examining Target Audience Perceptions About Four Presentation Formats for Fish Consumption Health Advisory Information." *Risk Analysis: An International Journal*, 18(5), 649-659.

Larcker, D. F., and Lessig, V. P. (1980). "Perceived Usefulness Of Information: A Psychometric Examination." *Decision Sciences*, 11(1), 121-134.

Leiss, W. (1996). "Three Phases in the Evolution of Risk Communication Practice." *Annals of the American Academy of Political and Social Science*, 545, 85-94.

Lissitz, R. W., and Green, S. B. (1975). "Effect of the Number of Scale Points on Reliability: A Monte Carlo Approach." *Journal of Applied Psychology*, 60(1), 10-13.

Long, J., and Fischhoff, B. (2000). "Setting Risk Priorities: A Formal Model." *Risk Analysis: An International Journal*, 20(3), 339-352.

Lyons, T., and Skitmore, M. (2004). "Project risk management in the Queensland engineering construction industry: a survey." *International Journal of Project Management*, 22(1), 51-61.

Macgill, S. M., and Siu, Y. L. (2005). "A new paradigm for risk analysis." *Futures*, 37(10), 1105-1131.

Malone, R. W. Jr., and Moses, K. D. (2005). "Development of Risk Assessment Matrix for NASA Engineering and Safety Center (NESC)." Project Management Challenge Conference 2005.

McNeill, D., and Freiberger, P. (1993). "Fuzzy Logic." Simon & Schuster, New York, NY.

Microsoft Corp. Excel. [12.0.6300.5000 SP1 MSO 12.0.6213.1000]. 2007. Microsoft Corp.

Minitab Corp. Minitab. [Release 14]. 2004. Minitab, Inc.

Moschandreas, D., and Karuchit, S. (2005). "Risk Uncertainty Matters: An Engineer's View." *International journal of risk assessment and management*, 5(2/3/4), 167-192.

Moses, K. D., and Malone, R. W., Jr. (2005). "Development of Risk Assessment Matrix for NASA Engineering Safety Center." *Rep. No. 20050123548*.

Moynihan, G. P., Fonseca Daniel J., and Ray, P. S. (2002). "A Hybrid System Approach for Space Station Risk Analysis." *International Journal of Risk Assessment and Management*, 3(2-4), 292-309.

NASA NPR8000.4 . NPR 8000.4 Risk Management Procedures and Guidelines. 2004. National Aeronautics and Space Administration.

National Research Council (1989). "Improving Risk Communication." National Academy of Sciences, Washington, D.C.

National Research Council (2005). "The Owner's Role in Project Risk Management." The National Academies Press, Washington, D.C.

"National Statistical Service."(2008). http://www.nss.gov.au/nss/home.NSF/pages/Sample+Size+Calculator+Description?OpenDocument.

Northwestern University Medical School (2008). "PROPHET StatGuide: Glossary." *<http://www. basic. northwestern. edu/statguidefiles/sg_glos. html#location>*.

Oxford English Dictionary Online (1989). "The Oxford English Dictionary." *<http://www. oed. com/>*, 2nd Edition.

Parrott, R., Silk, K., Dorgan, K., Condit, C., and Harris, T. (2005). "Risk Comprehension and Judgments of Statistical Evidentiary Appeals." *Human Communication Research*, 31(3), 423-452.

Pfleeger, S. L. (2000). "Risk Business: What We Have Yet To Learn About Risk Management." *The Journal of Systems and Software*, 53, 265-273.

PMI (2004). "PMBOK A Guide to the Project Management Body of Knowledge." Project Management Institute.

Porter, S. R., and Whitcomb, M. E. (2003). "The Impact of Contact Type on Web Survey Response Rates." *Public Opinion Quarterly*, 67, 579-588.

"Raosoft Sample Size Calculator."(2008). Available at http://www.raosoft.com/samplesize.html.

Rose, P. R. (2001). "Risk Analysis and Management of Petroleum Exploration Ventures." The American Association of Petroleum Geologists, Tulsa.

Rowe, W. D. (1977). "An anatomy of risk." John Wiley and Sons, Inc., New York.

Scheaffer, R. L., and McClave, J. T. (1990). "Probability and Statistics for Engineers." PWS-KENT Publishing Company, Boston, Massachusetts.

Siegle, Del. "Neag School of Education UConn Reliability Calculator."(2008). Available at http://www.gifted.uconn.edu/Siegle/research/Instrument%20Reliability%20and%20Validity/reliabilitycalculator2.xls.

Smith, M., and Taffler, R. (1994). "Improving the Communication Function of Published Accounting Statements." *Accounting and Business Research*, 139-146.

Thamhain, H. J., and Wilemon, D. L. (1987). "Building High Performing Engineering Project Teams." *IEEE Transactions on Engineering Management.*, EM-34(3), 130-137.

Thompson, K. M. (2002). "Variability and Uncertainty Meet Risk Management and Risk Communication." *Risk Analysis: An International Journal;* 22(3), 647-654.

Thompson, K. M., and Bloom, D. L. (2000). "Communication of risk assessment information to risk managers." *Journal of Risk Research*, 3(4), 333-352.

University of Central Florida Institutional Review Board (2008). "University of Central Florida Institutional Review Board Principal Investigator's Manual.".

Verzbolovskis, M. (2004). "ABS Approach to Classification using Risk Analysis." *ABS Technical Papers 2004*, 47-58.

Viscusi, W. K., and Zeckhauser, R. J. (1996). "Hazard Communication: Warnings and Risk." *Annals of the American Academy of Political and Social Science*, 545, 106-115.

Whittemore, A. S. (1983). "Facts and Values in Risk Analysis for Environmental Toxicants." *Risk Analysis: An International Journal*, 3(1), 23-33.

Zio, E. (1997). "A Methodology for Interactive Decision-Making in Environmental Management Involving Multiple Stakeholders.", Massachusetts Institute of Technology.

Zmud, R. W. (1978). "An Empirical Investigation of the Dimensionality of the Concept of Information." *Decision Sciences*, 9(2), 187-195.