Regular Paper

Making Risk Assessments More Comparable and Repeatable

David C. Hall*

MacAulay-Brown, Inc., 655 Discovery Drive, Suite 300, Huntsville, AL 35773

Received 30 December 2009; Revised 13 January 2010; Accepted 23 March 2010, after one or more revisions
Published online 3 June 2010 in Wiley Online Library (wileyonlinelibrary.com)
DOI 10.1002/sys.20169

ABSTRACT

Many of the objections to implementing Risk Management and acting upon risk results hinge on the subjectivity of the risk assessment system. This subjectivity makes it difficult to make risk assessments justifiable, repeatable, and comparable over an entire project, program, or organization. One cannot easily justify assigning a 30% likelihood to a risk occurring when others with more, the same, or less experience are ascribing a 60% likelihood of occurrence to a similar risk. How to get all (or most) risk assessments, regardless of type (software, hardware, integration, programmatic, external, etc.), justifiable, repeatable, and comparable has been one of the holy grails of Risk Management for years. The methodology outlined in this paper meets at least some of this requirement. The methodology requires incorporating the Likelihood of Occurrence into a set of specifically defined sublevels under each risk category rather than using it as a separate multiplication factor. Basically, the assumption behind this methodology is that the more mature the process, the more experience available, the more detailed the design, etc., the lower the likelihood of occurrence of a specific risk becomes. Making this assumption, incorporating the likelihood into each specific sublevel and requiring justification for each choice then allows the establishment of more representative scores for project risks and allows risk information to be presented in a justifiable, repeatable, and comparable fashion. © 2010 Wiley Periodicals, Inc. Syst Eng 14: 173–179, 2011

Key words: risk management; risk assessment; risk subjectivity; risk vocabulary; risk scoring system

1. INTRODUCTION

There are many definitions of risk that vary by specific application and situational context. This inconsistent and ambiguous use of the word "risk" is one of several current criticisms of the methods to manage risk [Hubbard, 2009]. Risk Management is a process that "everyone understands" and is used across many different professions (technical, financial, medical, environmental, etc.). This has led to a significant number of process and methodology interpretations linked to specific professions, work areas, and types of programs/projects/ organizations. Further confusing the issue is that Risk Management involves interacting with individuals at all levels in an organization in trying to establish and maintain a common language and approach.

There have been and are numerous different risk management standards (or guides) in use [see ISO, 2002, 2004, 2006, 2008, 2009a, 2009b; IEC, 2009; RAMP, no date; IRM, 2008; DAU Risk CoP, no date; AIRMIC, ALARM, and IRM, 2002; BSI, 2008; Office of Government Commerce, 2007; COSO, 2004], and others in development, but few of them were aimed at providing a risk management process that could be deployed by numerous organizations across different professions. And even those only tried to standardize the high level

* E-mail: Halld105048@yahoo.com

process steps and not the actual methodologies used in accomplishing the steps. Recognizing this, over the past couple of decades, more and more professions and organizations have adopted the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) Risk Management Standard and various process and vocabulary documents. Over the last few years, even ISO noted that the fragmentation of its risk management standards needed to be addressed. ISO 31000 [2009] was developed to combine all aspects of risk management and represent "best practice" in risk management processes and facilitate organizations investing in risk management infrastructure. The standard provides generic principles and guidelines for the implementation of risk management in general, and is supposed to be applicable to a broad range of stakeholders in all industry sectors and across all types of management systems. So the overall risk management *process* is becoming more and more standardized across all professions.

2. REVIEW

However, even though the overall Risk Management **process** has been fairly well standardized, the methodologies of how the steps are accomplished remain up to the individual to define. For example, how the actual assessments (or risk scoring system) are conducted remains extremely diverse and subjective. Each profession uses something different. As a case in point, one of the most popular in the financial profession is called Value-at-Risk (VaR). However, there are several different types of VaR—Long Term VaR, Marginal VaR, Factor VaR, and Shock VaR [Jorion, 2006]. But even with the different types, one of the noted authors stated that "the greatest benefit of using the VaR methodology lies in the *imposition of a structured methodology for critically thinking about risk*. Institutions that go through the process of computing their VaR are forced to confront their exposure to financial risks and to set up a proper risk management function. *Thus the process of getting to a VaR may be as important as_the number itself*" (italics mine) [Jorion and Taleb, 1997]. So, at least in the area of financial risk, use of a structured methodology has been recognized as significantly contributing to the usefulness of risk management and as important in ensuring that the process is used appropriately.

In the technical/engineering professions, most of the objections to implementing Risk Management and then actually acting upon risk results stem from the subjectivity of current risk assessment methodologies (or risk scoring systems). So in this paper we will focus on risk assessment methodologies. Currently all methodologies are accomplished with subjective input, and most do not required formalization of a justification. This makes it difficult to provide risk assessments that are justifiable, repeatable, and comparable within a single project, much less across multiple projects or throughout an organization. It is hard to justify a specific assessment when others of comparable, less, or greater experience are assessing the same or similar risks as significantly lower (or higher) than you. So the question to be answered is—how do we get risk assessments, regardless of type and coverage, justifiable, repeatable, and comparable?

It is interesting to note that different risk assessment methodologies (or risk scoring systems) have continued to emerge and proliferate as different types of users "define and refine" their risk assessment processes (e.g., Probabilistic Risk Assessment [NASA, 2002]). This proliferation of methods is considered healthy by some, but I think that the continuing divergence in many of the risk management process steps legitimizes those who refuse to believe such a subjective and ad hoc process can actually be useful. It would be very helpful if there had also been a convergence to one or a few risk assessment methodologies by virtue of the ongoing standardization efforts. However, since I started work in risk management in the 1970s, I have seen no indications that any group or organization has made or is making an attempt to standardize this specific part of the risk management process. Everyone believes that their needs are too different and are working out their own methodology without much motivation to try to standardize.

If one does a search on "Risk Management Subjectivity," one can find almost 200,000 citations discussing some aspect of this one problem within risk management. And in reviewing many of these citations, one finds quotations like the following:

> "As a result, risk assessment is largely guesswork. Guesswork means the savings can be just about anything the security manager chooses to report" [Bejtlich, 2007].
>
> "Subjectivity will not disappear, and we should not strive for that. We should try to build a good rationale for our risk assessments, but not aim for objectivity" [Bush, 2009].
>
> "Subjectivity is a necessary part of risk assessment. Even in quantitative risk assessments subjective judgment occurs" [Main, 2004].
>
> "The need for judgment introduces subjectivity and bias, and therefore uncertainty and the likelihood of inaccuracy. The results obtained by one risk analyst are unlikely to be obtained by others starting with the same information" [Redmill, 2001].
>
> "Subjectivity can never be removed completely from the risk assessment process" [Day, 2003].
>
> "Qualitative analysis is by definition approximate, but quantitative analysis is often assumed to be wholly objective. Yet there is considerable subjectivity in the analysis process" [Redmill, 2002].

So how do we attempt to standardize answering the following risk management question set in a justifiable, repeatable, and comparable way regardless of the type of risk identified: What can go wrong?[1] What is the likelihood that it will go wrong? What is the consequence if it goes wrong?

One methodology that should be useful in accomplishing this prodigious feat is called the Formal Analytical Scoring System (FASS). This methodology has been designed to force all assessors to use the same set of definitions for both likelihood and consequence when doing their assessments and requiring justification statements for each decision. To

[1] Note that "wrong" is also intended to mean unplanned and/or unexpected.

Table I. Examples of Universal Risks

Cost Development	Schedule Development
Requirements variability	Design and Engineering Maturity
Technology Maturity	Transportation Complexity
History/Experience	Component Maturity
Fabrication resources	Testing Required
Methodology and Process Maturity	Development Support Resources
Personnel	Hardware and Software Interfaces
Logistics Requirements	Facility/Site Resources
Data Requirements	Integration Environment and Resources

successfully accomplish this, it is necessary to (1) establish an appropriate set of risks that can be used to assess your project, program or organization, (2) incorporate a set of likelihood of occurrence definitions for each risk within that set of risks, and (3) establish a set of consequence definitions based on a combination of programmatic and technical impacts that are appropriate for the level of comparison you are dealing with.

3. RISK IDENTIFICATION—THE BASIS FOR RISK MANAGEMENT

Framing the risk management question set is a fundamental problem with all forms of risk assessments. Tversky and Kahneman [1981] note that, for example, there are two major areas that cause risks to be discounted or ignored. The first is: Because our brains get overloaded and we tend to take mental shortcuts, the risk of extreme events is usually discounted because the probability is too low to evaluate intuitively. This basically means that people tend to largely or totally ignore a serious or fatal risk. Likewise, an extremely catastrophic or mentally disturbing event may be ignored in an assessment despite the fact it has occurred before and has a nonzero likelihood. Second, an event that everyone agrees is inevitable may be ruled out of an assessment due to an unwillingness to admit that it is inevitable. Such human tendencies for wishful thinking and arrogance often affect even the most rigorous applications of risk management.

Lack of experience is another factor in getting an accurate risk set to consider. In many cases I have seen, valid risks were not being considered or assessed simply because no one involved in the project had any experience with the specific type of technology, customer, or process. The Black Swan example [Taleb, 2007] has been used to show how an outlier with an extreme impact can be completely missed because there is no knowledge of the possibility. For an average project there may not be any real Black Swans. But for projects with (for example) low Technology Readiness Levels or severe political/budget implications, a Black Swan could be a significant driver.

To try to ensure that all risks appropriate to your project, program, or organization are at least thought of, it is recommended that a set of universal risks that MUST be considered for every project or operation should be established and used to try and minimize leaving out risks due to wishful thinking, arrogance, inexperience, etc. Having a checklist created by subject matter experts and experienced risk managers detailing the risks to be contemplated provides a much greater potential that all risks have at least been considered. Many people and organizations have recognized that risk checklists are useful, and there are numerous ones specific to professions or technologies available (a search on Risk Checklists provides over 9 million citations). Doing this goes a long way toward reducing the subjectivity of establishing a risk list. This is another area that could use some standardization. There has only been one attempt that I know of to develop a Universal Risk List [Risk Management Research Collaboration, 2002].

4. FORMAL ANALYTICAL SCORING SYSTEM

The FASS methodology is based on the following assumption: the more mature the process, the more expertise available, the more detailed the design or the more you have built, the lower the likelihood of the risk event occurring. And vice versa, of course. This assumption is not always valid, but for purposes of simplifying the methodology, we accept it as **normally**[2] valid and need to be careful to note any specific risks for which it is determined not valid.

The first step in using this methodology is to establish a set of risks appropriate to your specific project, program, or organization. This set of risks should be comprehensive enough to allow all appropriate risks to be addressed. As noted above, there are numerous checklists available to aid in developing such a list specific to your project. Examples of "universal risks" are shown in Table I. Note that each of these generic titles have to be further defined for your specific project.

Once you have the risk set for your project, it should be winnowed out by eliminating those risks that could only minimally affect the project. Once you have completed and gotten acceptance of your specific set of risks, you then further define them by developing a set of Likelihood Level Statements for each of them. These Likelihood Level Statements are based on your specific project and incorporate the maturity of the process, the level of the design, the build level of the hardware, etc., for each risk. Examples are shown in Table II. In this example, I have chosen to use a five-level set of statements, but you can chose any number for your set. This

[2] Note that this assumption is not valid for natural disasters such as earthquakes, etc. But it is valid for most human actions.

Table II. Examples of Likelihood Level Statements for a Technology Development Project

Technology Maturity	Specific description of technology being considered
Level 5	Pre-Concept – Scientific research is required and no supporting technology base is available.
Level 4	Concept. Documented design meeting functional requirements is complete
Level 3	Engineering Model/Breadboard. Functional hardware model has passed performance/functional tests for component maturation
Level 2	Prototype. Fit, form, and function have been demonstrated by a technically analogous hardware component. Prototype passed qualification & acceptance tests.
Level 1	Operational. A technically identical (but not necessarily physically identical) hardware item is currently operational and deployed in an environment similar to XXX
Personnel	Specific description of personnel type/numbers
Level 5	No approved plan to staff the development activities exists
Level 4	An approved plan exits to staff the development activities, but sufficient personnel are not available
Level 3	Sufficient personnel exist, but have less than one year average experience.
Level 2	Sufficient personnel are available with average experience exceeding one year and are functioning as a team
Level 1	Sufficient personnel are available and have created similar hardware and have experience on XXX items
Transportation Complexity	Specific description of product/material to be transported
Level 5	Product is too large/heavy to be transported by rail/aircraft.
Level 4	Product is too large/heavy to be transported by highway.
Level 3	Product has to be handled via exception for transportation by rail, air or highway.
Level 2	Product has to be handled via discrete planning for transportation by rail, air or highway
Level 1	Product can be transported by rail, air or highway

five-level set was chosen because I intend to use a 5 × 5 matrix for the final risk assessment matrix.

If you want to be more specific in your risk assessments, you can determine a weighting factor for each of these risks and Likelihood Level statements. This can be accomplished by use of one of several tools, such as simple multivoting or the Analytic Hierarchy Process. Note that such weighting factors could be assigned from a project, program, or enterprise level depending on whether or not you want to compare risk levels within your organization's projects or programs. Table III shows how the same risks and Likelihood Level statements might look after having weighting factors developed and applied.

Use of weighting factors would enable you to use a true Likelihood times Consequence formula to come up with a risk number. It does take more time to establish, but provides a comparable and justifiable quantitative figure.

5. CONSEQUENCE DEFINITIONS

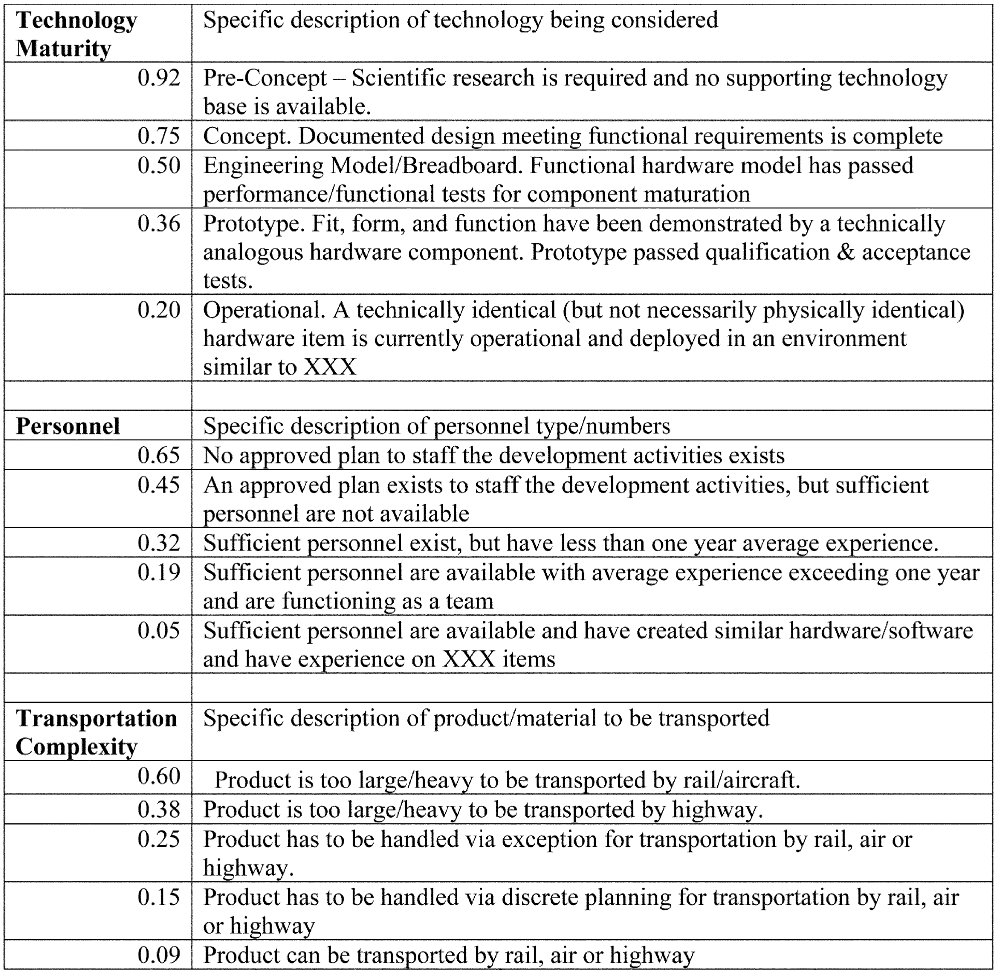
Now you need to establish a set of consequence definitions that matches the level of comparison you are trying to achieve. The single consequence statement that connects performance, cost, and schedule is required to allow comparisons throughout a project, program, or organization. An example of this type of consequence set is shown in Table IV.

To use this in a quantitative formula, you can establish specific weighing factors (using the same methodology used to establish weighing factors for the Likelihood Levels) for each consequence level. For example, 5 could equate to .95, 4 to .75, 3 to .55, 2 to .35, and 1 to .15.

6. ASSESSING RISKS

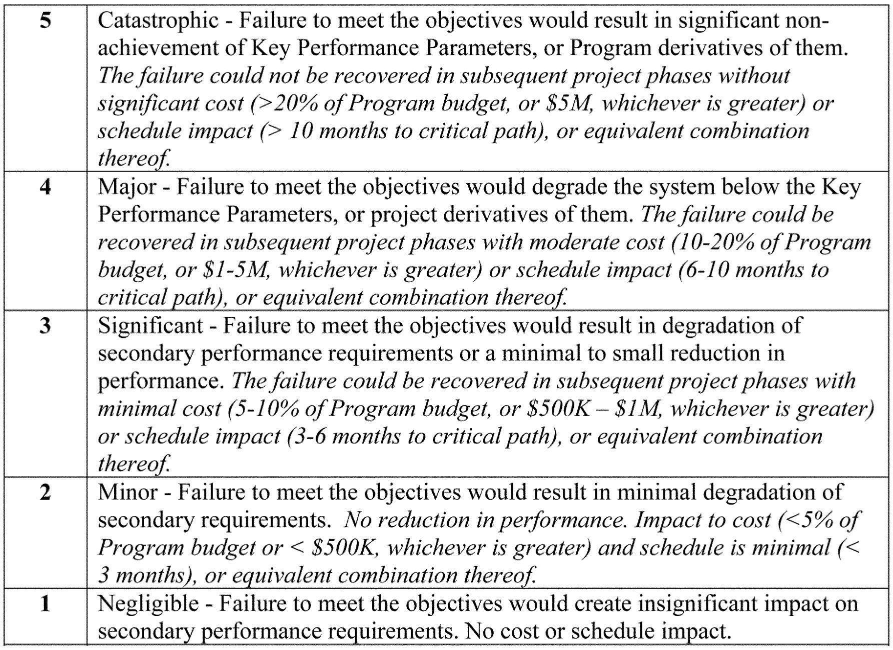
With this set of Risks, Likelihood Levels, and Consequences established and approved, each risk assessor now must assign specific likelihood levels and consequences to their risk. As the assessor decides which risk and which risk levels are appropriate to their situation, they must also justify why they chose a specific Likelihood Level and Consequence. Using predefined and accepted factors for risk assessments and requiring justifications enables (forces) assessors to be more objective and thoughtful. Simply assigning a likelihood of occurrence and a consequence for a risk like "We will not be able to accomplish this specific activity within planned cost

Table III. Examples of Weighted Likelihood Level Statements



Technology Maturity	Specific description of technology being considered
0.92	Pre-Concept – Scientific research is required and no supporting technology base is available.
0.75	Concept. Documented design meeting functional requirements is complete
0.50	Engineering Model/Breadboard. Functional hardware model has passed performance/functional tests for component maturation
0.36	Prototype. Fit, form, and function have been demonstrated by a technically analogous hardware component. Prototype passed qualification & acceptance tests.
0.20	Operational. A technically identical (but not necessarily physically identical) hardware item is currently operational and deployed in an environment similar to XXX
Personnel	Specific description of personnel type/numbers
0.65	No approved plan to staff the development activities exists
0.45	An approved plan exists to staff the development activities, but sufficient personnel are not available
0.32	Sufficient personnel exist, but have less than one year average experience.
0.19	Sufficient personnel are available with average experience exceeding one year and are functioning as a team
0.05	Sufficient personnel are available and have created similar hardware/software and have experience on XXX items
Transportation Complexity	Specific description of product/material to be transported
0.60	Product is too large/heavy to be transported by rail/aircraft.
0.38	Product is too large/heavy to be transported by highway.
0.25	Product has to be handled via exception for transportation by rail, air or highway.
0.15	Product has to be handled via discrete planning for transportation by rail, air or highway
0.09	Product can be transported by rail, air or highway

Table IV. Example Consequence Set



5	Catastrophic - Failure to meet the objectives would result in significant non-achievement of Key Performance Parameters, or Program derivatives of them. *The failure could not be recovered in subsequent project phases without significant cost (>20% of Program budget, or $5M, whichever is greater) or schedule impact (> 10 months to critical path), or equivalent combination thereof.*
4	Major - Failure to meet the objectives would degrade the system below the Key Performance Parameters, or project derivatives of them. *The failure could be recovered in subsequent project phases with moderate cost (10-20% of Program budget, or $1-5M, whichever is greater) or schedule impact (6-10 months to critical path), or equivalent combination thereof.*
3	Significant - Failure to meet the objectives would result in degradation of secondary performance requirements or a minimal to small reduction in performance. *The failure could be recovered in subsequent project phases with minimal cost (5-10% of Program budget, or $500K – $1M, whichever is greater) or schedule impact (3-6 months to critical path), or equivalent combination thereof.*
2	Minor - Failure to meet the objectives would result in minimal degradation of secondary requirements. *No reduction in performance. Impact to cost (<5% of Program budget or < $500K, whichever is greater) and schedule is minimal (< 3 months), or equivalent combination thereof.*
1	Negligible - Failure to meet the objectives would create insignificant impact on secondary performance requirements. No cost or schedule impact.

Table V. Example Risk Assessment

Risk	Level	Justification
Technology Maturity	2	Fit, form, and function have been demonstrated by the XXX unit. The XXX Unit has passed qualification test on June 2, 2009 and acceptance test on August 6, 2009.
Consequence	4	Major – If this unit failed to meet its performance objectives, it would degrade the system below the Key Performance Parameter of availability. *The failure would require moderate cost increase of $1M to redesign and replace the Unit within a subsequent Project Phase.*

and schedule" is very subjective and subject to numerous biases. In this methodology, all assessors are required to pick one of the predefined statements that best fits their risk and justify why it best fits. Table V provides the required formal assessment.

Once each assessor has completed assigning a specific set of Likelihood Levels and Consequences to their risk, you can now use either a 5 × 5 matrix (Fig. 1) or a quantitative formula to establish a Risk Level. This particular risk is noted in Figure 1 by the star in the appropriate box. This risk level is unique to each risk and can then be compared to all other risk levels within your methodology. But since it is based on a predefined set of Likelihoods and Consequences, it is easily repeatable and comparable. In comparing risks within a project or across multiple projects, each of these decisions, along with the associated justification statements, can be reviewed by management to determine if they agree with the assessment.

7. CONCLUSIONS

This methodology is one way of increasing the probability that all risks inherent in a project, program, or organization have been considered and reducing the subjectivity of the subsequent risk assessments. It does not provide absolute results, but does provide a basis for making risk assessments justifiable, repeatable, and comparable. This methodology can be used for all types of risk assessments; hardware, software, integration, management, external, etc., and can allow a valid comparison of risk level regardless of type and genesis of risks. Since we are mostly dealing with human actions, perceptions, feelings and concerns, I doubt that we will ever be able to get absolute results.

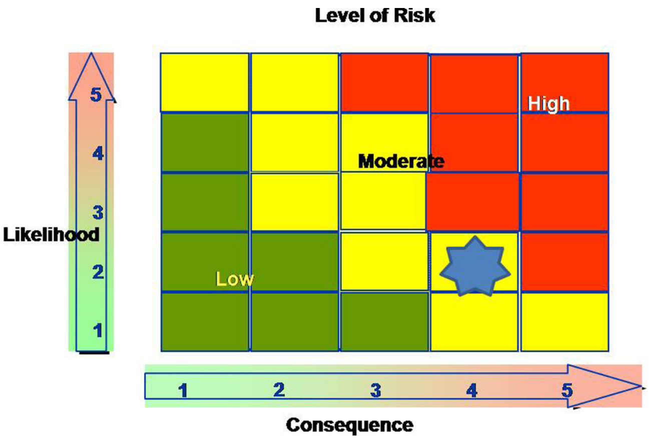


Figure 1. Example 5 × 5 matrix. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

However, for risk assessments to mean something—to be value-added—the numbers or levels developed do not need to be absolutely related to anything. They only need to be related to each other in an ordinal sense. As a comparison, it is meaningful to say that one mass weighs twice as much as another because it makes no difference whether you measure in grams, pounds, or tons.[3] It is just as meaningful to say one risk is twice as risky as another as long as each assessment is repeatable, justifiable, *and comparable*. It makes no difference whether you measure in cost, schedule, performance, public relations, or any other units *as long as you use the same measures and definitions to analyze all your risks.*

REFERENCES

Association of Insurance and Risk Managers (AIRMIC), ALARM, and Institute of Risk Management (IRM), Risk Management Standard, London, 2002.

R. Bejtlich, No ROI? No problem, http://taosecurity.blogspot.com/2007/07/no-roi-no-problem.html, July 14, 2007.

British Standards Institution (BSI), BS 31100:2008 Risk Management Code of Practice, London, 2008.

L. Bush, Quality risk management demystified at CMC strategy forum, BioPharm Int, August 12, 2009.

Committee of Sponsoring Organizations of the Treadway Commission (COSO), Enterprise Risk Management (ERM)—integrated framework, Washington, DC, 2004.

DAU Risk CoP; https://acc.dau.mil/CommunityBrowser.aspx?id=17607&lang=en-US, April 2010.

K. Day; Can risk assessments be fact-based? Security Management 47 (September 2003).

International Electrotechnical Commission (IEC), IEC 31010: Risk

[3] Of course, it does make a difference if you are required to pick the mass up. So *what* you are trying to accomplish does matter.

Management—Risk Assessment Techniques, Geneva, Switzerland, 2009.

International Organization for Standardization (ISO), ISO/IEC Guide 73:2002, Risk Management: Vocabulary, Geneva, Switzerland, 2002.

International Organization for Standardization (ISO), ISO/IEC/IEEE 16085:2004, Geneva, Switzerland, 2004.

International Organization for Standardization (ISO), ISO/IEC 16085: Risk Management 2006, Geneva, Switzerland, 2006.

International Organization for Standardization (ISO), ISO/IEC 27005:2008 Information Technology—Security Techniques—Information Security Risk Management, Geneva, Switzerland, 2008.

International Organization for Standardization (ISO), AS/NZS/ISO 31000:2009 Risk Management Systems Standard, Geneva, Switzerland, 2009a.

International Organization for Standardization (ISO), ISO 31000: 2009: Risk Management - Principles and Guidelines, Geneva, Switzerland, 2009b.

D. Hubbard, The failure of risk management: Why it's broken and how to fix it, Wiley, Hoboken, NJ, 2009.

The Institute of Risk Management (IRM), Risk Management Standard_030820, London, 2008.

P. Jorion, Value at risk: The new benchmark for managing financial risk, 3rd edition, McGraw-Hill, New York, 2006.

P. Jorion and N. Taleb, The Jorion/Taleb debate, Derivatives Strategy 2(4) (April 1997).

B.W. Main, Risk assessment: Basics and benchmarks, Design Safety Engineering, Ann Arbor, MI, 2004.

NASA, Probabilistic risk assessment procedures guide for NASA managers and practitioners, NASA Office of Safety and Mission Assurance, Houston, TX, August 2002

Office of Government Commerce, UK, Management of risk: Guidance for practitioner; London, 2007.

RAMP: Risk Analysis and Management for Projects, Institute of Actuaries and Institution of Civil Engineers, London.

F. Redmill, Exploring subjectivity in hazard analysis, Eng Management J 12(3) (June 2002), 139–144.

F. Redmill, Subjectivity in risk analysis, Redmill Consultancy, Felix.Redmill@ncl.ac.uk, July 2001.

Risk Management Research Collaboration, Universal Risk Project Report, February 2002.

N. Taleb, The Black Swan: The impact of the highly improbable, Random House, New York, 2007.

A. Tversky and D. Kahneman, The framing of decisions and the psychology of choice, Science New Series 211(4481) (January 30, 1981), 453–458.

David C. Hall received his BMe from Auburn University in 1975. He worked for 30 years as a USAF active duty and Reserve officer, retiring as a Colonel in 2006. He has worked in industry for over 25 years, mainly as a Systems Engineer and Risk Manager. His Risk Management experience includes being a Risk Management Train-the-Trainer for several federal government agencies, providing numerous risk management training courses and tutorials to over 3000 people, developing process improvements and accomplishing risk management functions for both commercial and government programs. He has received the INCOSE Expert Systems Engineering Professional (ESEP) Certification and also is an ISC^2 Certified Information Systems Security Professional (CISSP).