Risk Cube Data Biases

William T. Siefert
F-15 Electrical Standards
Core Avionics Team
The Boeing Company
william.t.siefert@boeing.com

Eric D. Smith
Systems Engineering Program
Research Institute for Manufacturing and Engineering Systems (RIMES)
University of Texas at El Paso
ESmith2@UTEP.edu

Abstract. Risk cubes, actually a 5x5 matrix, used in industry characterize particular risks in terms of the likelihood of occurrence, and the consequence of the actualized risk. Human cognitive bias research led by Daniel Kahneman and Amos Tversky exposed systematic translations of objective probability and value as judged by human subjects. An examination of industry-generated risk cube data reveals evidence of biases in the judgment of likelihood and consequence. In addition, other biases -- including a 'diagonal bias' -- are revealed in the risk cube data. The evidence presented could improve risk cube based risk analysis.

RISK CUBES

In the industry data analyzed for this paper, the two parameters of risk cube data are likelihood (L) and consequence (C). While the Term "Risk Cube" is used, the graphic is actually a 5x5 matrix. It is unclear why the term a cube took hold as it has only two dimensions. The use of the term "cube" is used as a matter of convenience for the reader. The cognitive bias literature uses the terms subjective probability and utility, respectively. In this paper, the possible distinctions between these terms will not be developed. (Note: For historical reasons, "value" and "value function" are used in the psychology literature in reference to what we here more clearly refer to as "utility" and "utility function.") Risk is usually computed as the product of Likelihood and Consequence: R = L x C. Usually, likelihood is scaled from 0-1, as is consequence.

Table 1: Equivalent terms: Likelihood and probability, consequence and value

Subjective Parameters	
Likelihood (L)	Consequence (C)
Subjective Probability, $\pi(p)$	Utility (negative), $U^{-}(v)$
Shown on:	
Ordinate, Y axis	Abscissa, X axis
Objective Parameters	
Objective Probability, p	Objective Value, v

Risk Cubes in Industry. The most common tool used to track and manage risks is the risk cube, an example of which is shown in Figure 1. The shape and number of zones are determined by individual companies, and vary in different risk cubes. Lines dividing the zones are iso-risk lines. Allowing the zones to be fuzzy eventually leads to continuous numbers. Risk points appear inside the cube. The cube provides a visual communication aid of the risks posed by particular events. The closer to the lower left corner of the cube the risk is placed, the lower the risk.

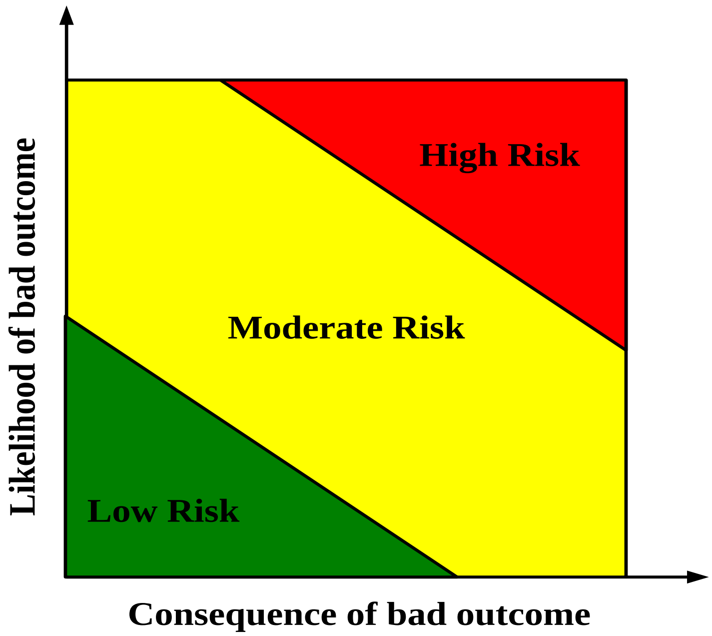

Figure 1: Risk cube with zones

This basic graphical representation has several weaknesses, including a lack of specific granularity and the lack of a repeatable way to locate the risk in the cube. A repeatable means of placing risk points is generally required for human beings. The lack of granularity in the risk cube shown in Figure 1, when used in practice, exposes the fact that most people cannot repeatedly judge real risks and accurately place risks points in the same location. Further, the continuous scale hinders communication among humans, since un-resolvable discussion may occur about the location of risk points. Industry needs an appropriate risk cube with sufficient but not excessive granularity. A risk cube used by industry needs to be appropriate for the task of timely risk point placement, and significant risk point replacement toward the lower left after successful risk mitigation, without encouraging micro-management of risk point placement.

The risk cube shown in Figure 2 represents an approach to achieve repeatable values likelihood and consequence. The five-range, 25-square risk cube with integer assignment of 1-5 for both likelihood and consequence provides sufficient granularity, reducing ambiguity in deciding which position is correct. This cube, along with a reasonable set of qualitative definitions of the ranges, provides engineers with a repeatable way to categorize risks.

Qualitative descriptions/definitions criteria exist for each of the squares in the cube for both likelihood and consequence. These criteria or definitions add stability and repeatability to risk point placement. Without repeatability, the process would lose much of its value. Table 2 lists a set of definitions for a qualitative risk mapping from real risks to the five ranges of likelihood and consequence for technology. Any such definitions are tailored to company requirements. Further, more objectivity in the definitions provides more confidence in risk point placement. In fact, perfect objectivity in the determination of likelihood and consequence, though expensive and largely impracticable, would eliminate most of the biases described below. Likelihood can be made objective with objective historical data of frequency of events, while consequence can be made objective with historical data. Mil-Std 882d and Mil-Std 1629a describe approaches to more objective risk analysis.

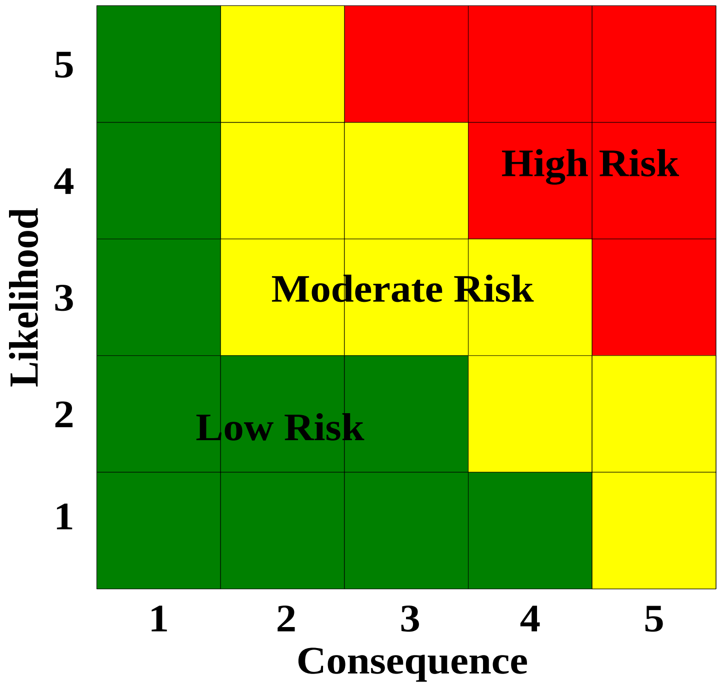

Figure 2: 5x5 risk cube

Table 2: Qualitative description for five ranges of likelihood and consequence

Qualitative scales for likelihood and consequence of technology risks				
Likelihood of failure – technology dependence				
Low	Minor	Moderate	Significant	High
No new technology – Systems are off the shelf	Minor modification of existing technology	Dependant on innovative use of existing technologies	Dependant on new technologies that are in development	Dependant on new technologies that are not yet funded
Consequence of failure – technical				
Low	Minor	Moderate	Significant	High
Little or no impact on program objectives	Minor reduction in technical performance with little or no impact on program objectives	Reduction in technical performance with limited impact on program objectives	Significant degradation in technical performance with a major impact on program objectives	Major degradation in technical performance that could jeopardize program success

COGNITIVE BIASES IN PROBABILITY AND VALUE

The most accepted descriptive theory of subjective expected value decision making by humans is Prospect Theory, developed by Daniel Kahneman and Amos Tversky (1979). Kahneman won the Nobel Prize in Economics in 2002 "for having integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty" (Nobel, 2002). (Since Nobel prizes are not bestowed posthumously, Tversky did not share this prize.) The contention of the Kahneman and Tversky school of “heuristics and biases” is that the presence of cognitive biases – even in extensive and vetted analyses – can never be ruled out. Innate human biases, and exterior circumstances, such as the framing or context of a question, can compromise estimates, judgments and decisions. It is important to note that subjects often maintain a strong sense that they are acting rationally while exhibiting biases.

Prospect Theory describes the subjective human decision-making process, specifically in the subjective assessment of probabilities and values, and their combination in gambles. Prospect Theory breaks subjective decision making into a preliminary ‘screening’ stage, and a secondary ‘evaluation’ stage. In the screening stage, probabilities and values are subjectively assessed, while the evaluation stage combines the subjective probabilities and values (utilities). Only the subjective assessment of probabilities and values is of interest in this paper.

Subjective Probability, π(p). Prospect Theory describes the subjective evaluation of probabilities, $\pi(p)$, according to the experimentally-obtained curve in Figure 3 (Tversky and Kahneman, 1992). This non-linear transformation, $\pi(p)$, shows how small probabilities are overestimated and large probabilities are underestimated.

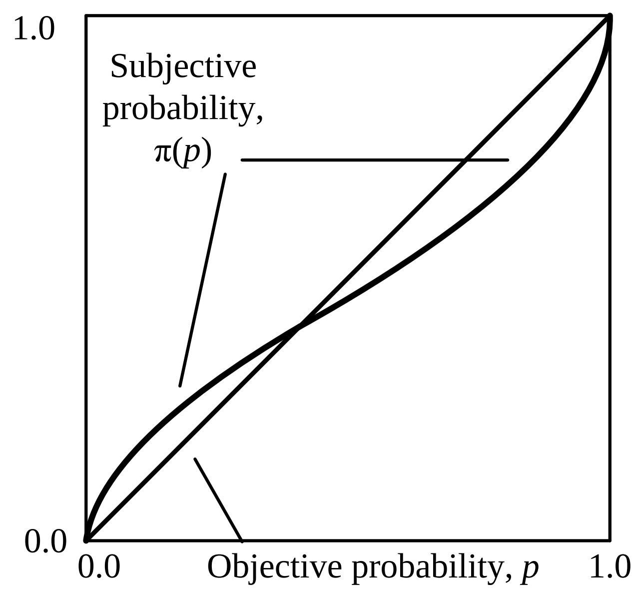

Figure 3: Subjective probability $\pi(p)$

The modeling equation for subjective probability is (Tversky and Kahneman, 1992)

$$\pi(p) = (p^{\delta}) / [p^{\delta} + (1- p)^{\delta}]^{(1/ \delta)}$$

where p = objective probability, with $0 < \delta \leq 1$. When δ =1, $\pi(p) = p$ = objective probability. A usual value for δ is $\delta = 0.69$ for losses (and $\lambda = 0.61$ for gains).

Subjective Utility. One of the effects of Prospect Theory's screening stage is that values are considered not in an absolute sense (from zero), but subjectively from the reference point established by the subject's perspective and wealth before a decision. This is an example of the psychological phenomenon called framing. The key graph that shows how objective values translate into subjective utilities is shown in Figure 4. Note the significant disparity in magnitude with which gains and losses are subjectively valued – approximately a 1-to-2 ratio.

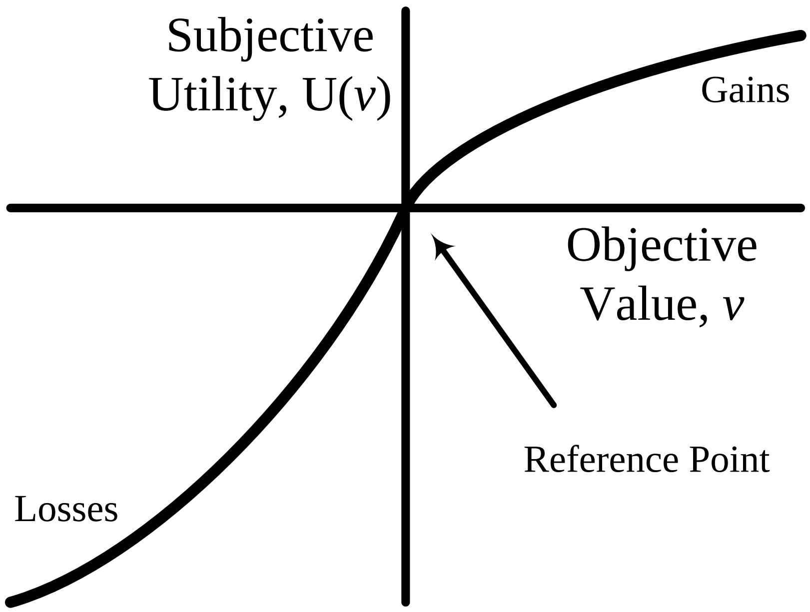

Figure 4: Subjective utility versus objective value according to Prospect Theory

For gains, $0 \le v \le \infty$, the utility function may be: $U^+(v) = \mathrm{Ln}(1 + v)$, where v = objective value. While for losses, $-\infty \le v \le 0$, the utility function may be: $U^-(v) = -(\mu)\mathrm{Ln}(1 - v)$, with $\mu \approx 2.3$.

Theoretical Implication of Prospect Theory for the Risk Cube. Figure 5 shows the theoretical implications of probability and value biases for the risk cube. The most obvious influence of non-symmetric subjective probability assessments will be to compress probability judgments inward toward the risk cube probability of three. Any non-symmetric effect may be difficult to observe.

For the assessment of values, note that only the loss section of the objective value to subjective value (utility) graph has been employed. The reasoning is that risks deal with possible losses, and not with opportunities for gain. The most likely true prediction of the loss section of the value curve is the following. Human subjects exaggerate the influences of losses when the losses will occur to their personal wealth. In industry, the value curve predicts that managers with an increased sense of corporate ownership will maintain a heightened awareness as to the monetary loss and corporate destabilization that a technical risk could cause. In fact, given the limits of any corporation, this managerial view is correct. Line engineers with less buy-in and a broader job market, will tend to see the objective value of the consequences of risks.

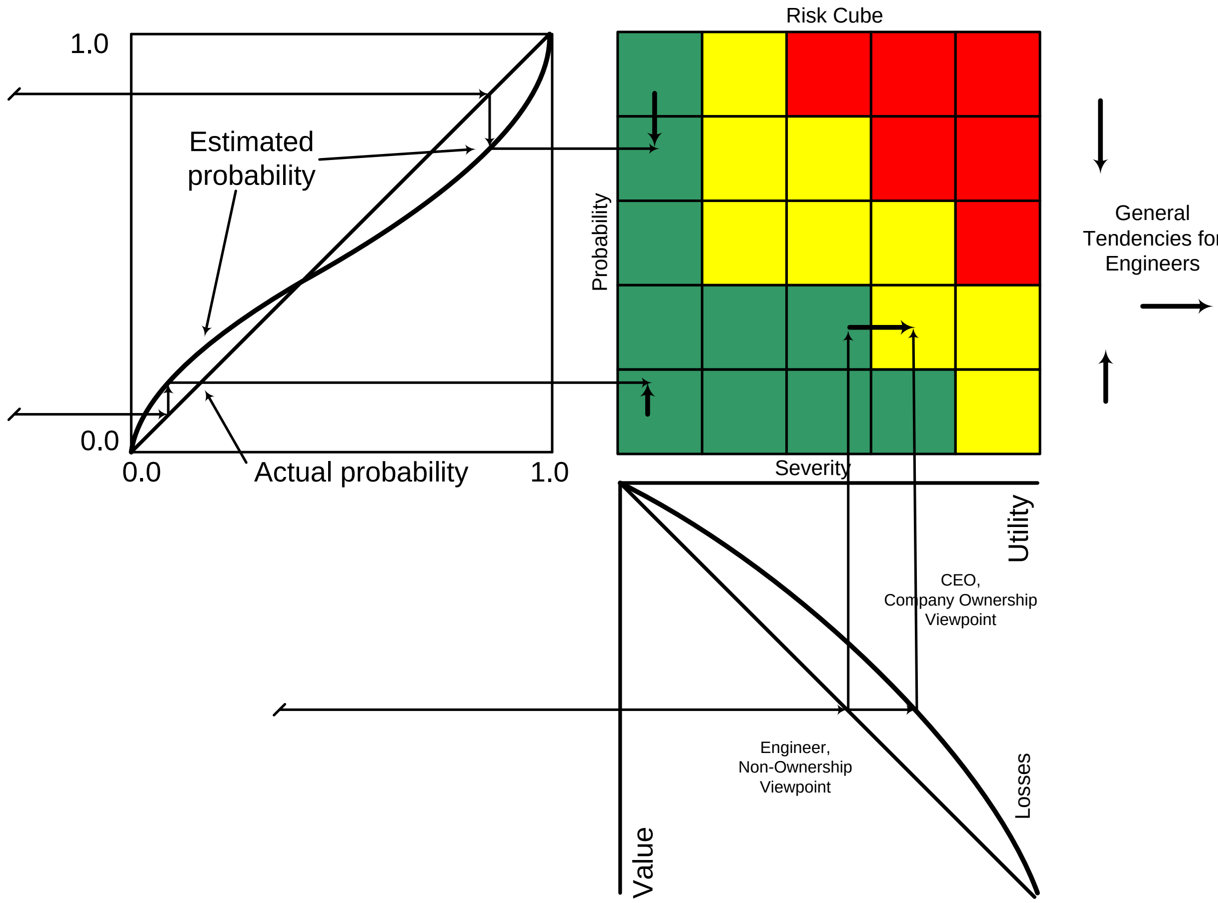

Figure 5: Theoretical implications of prospect theory for the risk cube

INDUSTRY CASE STUDY DATA, BIASES, AND ANALYSES

A set of high-quality industry data was analyzed to test for the presence of the theoretical implications. A sample of this data is given in Appendix A. The data were collected in the aerospace industry over a period of eight years for the purpose of risk management tracking. The engineers who entered the data did not know that the data would ever be used to test for the presence of cognitive biases. “Original data” refers to the first determination of likelihood and consequence numbers, while “current data” refers to the updated likelihood and consequence numbers, the update occurring months to years after the original datum point determination. Figure 6 shows the original data.

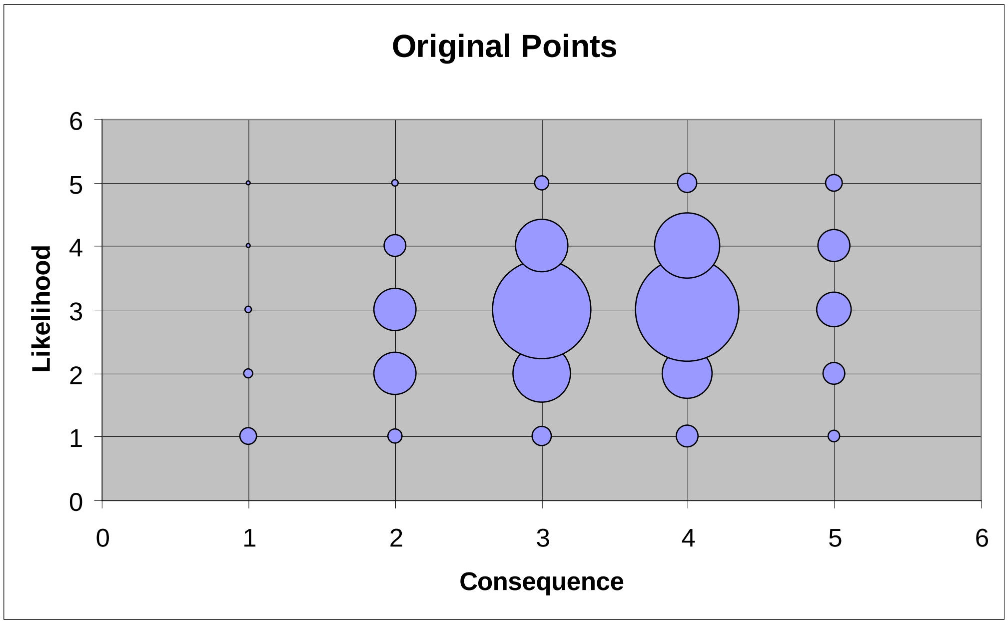

Figure 6: Bubble chart of original data. Bubbles sized by area for datum counts

ANALYSES AND OBSERVATIONS OF INITIAL DATA

First, we note that there are two impediments for the appearance of cognitive biases in the industry data: 1) The industry data are granular while the predictions of Prospect Theory are for continuous data, and 2) the qualitative descriptions for the five ranges of likelihood and consequence may have some non-linear influence in the placement of risk datum points.

Nevertheless, the evidence of cognitive biases emerges from the data as shown below.

Estimation in a Pre-Define Scale Bias. Schwarz (1990) conducted the following experiment where a response scale effected judgment. The following two questions (now provided with average answers) were each posed to a random 50% of a group of experimental subjects.

Please estimate the average number of hours you watch television per week:

____	____	__X__	____	____	____
1-4	5-8	9-12	13-16	17-20	More

Please estimate the average number of hours you watch television per week:

____	____	__X__	____	____	____
1-2	3-4	5-6	7-8	9-10	More than 10

With both random halves of the experimental population, the average answer lay in the central option of the five pre-determined ranges. Numerically, however, the average answers do not coincide or even overlap. Obviously, humans cannot be trusted with a calibration problem, even when it pertains to their own personal experience. In this case, subjects seemed eager to calibrate a number in a way that seemed most compatible with the given range of scale.

Statistical Evidence: Figure 7 shows the original likelihood data centered in their 1-5 scale around L =3.

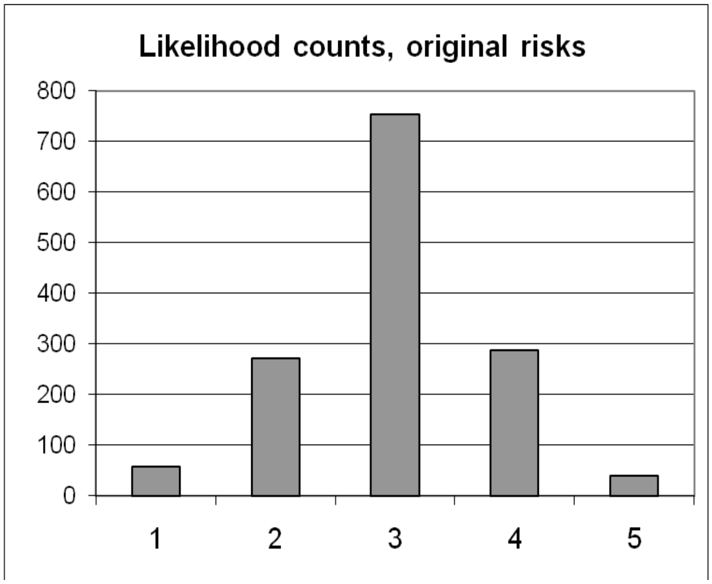

Figure 7: Likelihood data centered around L = 3

Figure 8 shows the original consequence data centered around C = 3.5, and balanced on C =3 and 4. An explanation for the shift toward higher consequence is provided in Section 3.5.

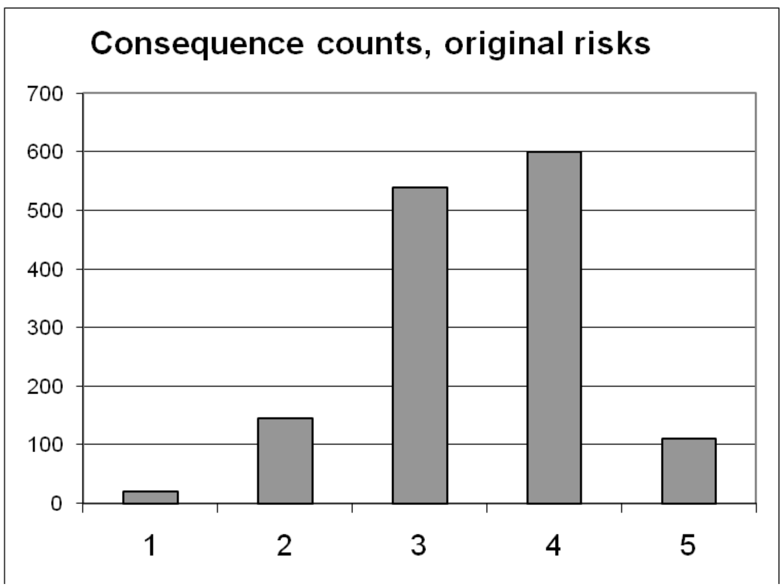

Figure 8: Consequence data balanced on C = 3 and 4

There are some infrastructure biases that help explain the centered data. A production aircraft will have few extreme risks. Also, a risk program on a production aircraft will not focus on the low magnitude risks.

Diagonal Bias. In anticipation of later moving the risk point toward the origin, risk points are withdrawn from the origin upward and rightward along the diagonal.

Statistical Evidence: Figure 9 shows a regression line plotted on the original data points.

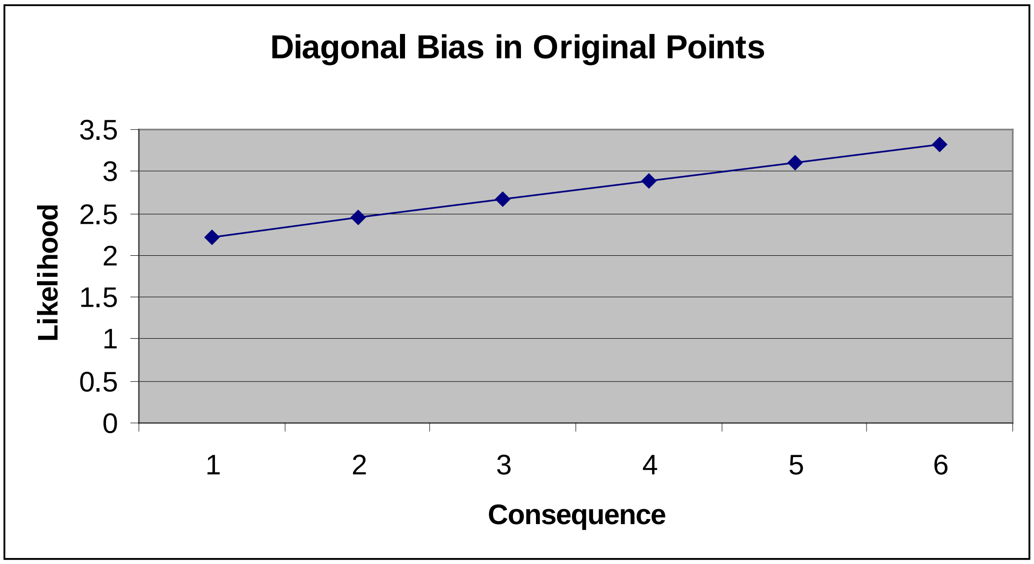

Figure 9: Regression line of original datum points showing a diagonal bias

Table 3: Regression parameters showing diagonal bias

Regression on 1412 Original Points		
Likelihood Intercept	Slope Coefficient	R
2.2	0.22	0.23

Probability Centering Bias. Likelihoods are pushed toward L = 3, symmetrically as a first-order approximation.

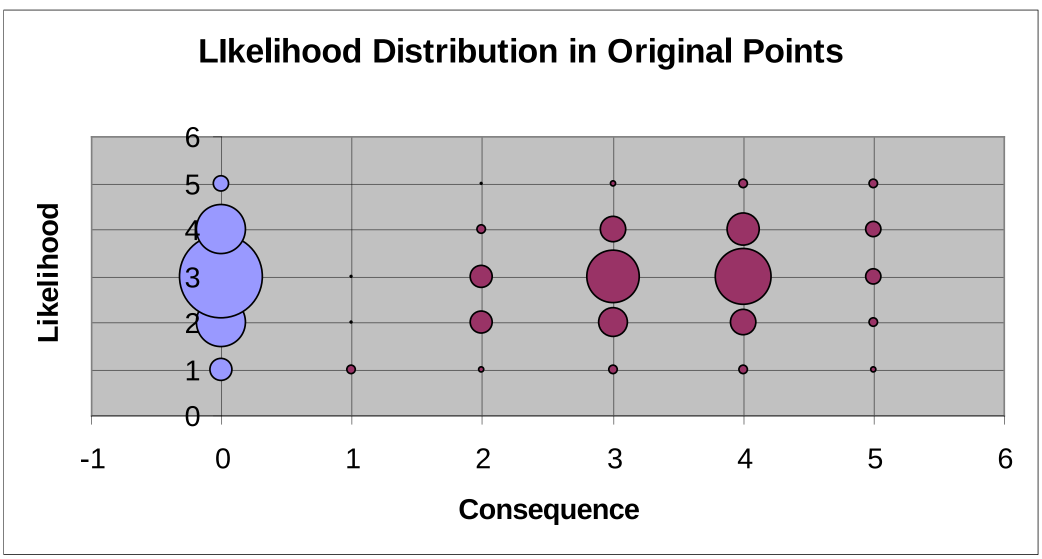

Figure 10: Likelihood marginal distribution of original datum points

Is seems that because 'people estimate probabilities poorly' (Cosmides and Tooby, 1996) they basically just distribute the original likelihood points symmetrically around likelihood=3.

Table 4: Likelihood original datum point counts compared with a normal distribution

Likelihood Original Datum point counts				
1	2	3	4	5
58	272	754	288	40
Normal distribution with mean = 3 and standard deviation = 0.78				
38	330	676	330	38

Statistical Evidence: The marginal distribution of likelihood of the original datum points show a high degree of symmetry: 58, 272, 754, 288, 40.

Asymmetrical Probability Bias. The subjective probability transformation, $\pi(p)$, predicts that likelihood data will be pushed toward L $\approx$ 3, with large probabilities translated down more than small probabilities are translated up, resulting in a reduced amount of large subjective probabilities, comparatively.

Statistical Evidence: The original datum counts for L =1 and L =5 are 58 and 40, respectively; 40<58, supporting the predicted asymmetry. The original datum counts for L = 2 and L = 4 are 272 and 288, respectively; 272<288, not supporting the predicted asymmetry. However, it can be supposed that real risks prompting the creation of an original datum point usually have L $\geq$ 3, and that such datum points have been in general pulled down from the L = 5 area by the significant translation of $\pi(p)$.

Consequence Bias. Judged consequence is pushed toward higher C values. Figure 11 shows marginal consequence.

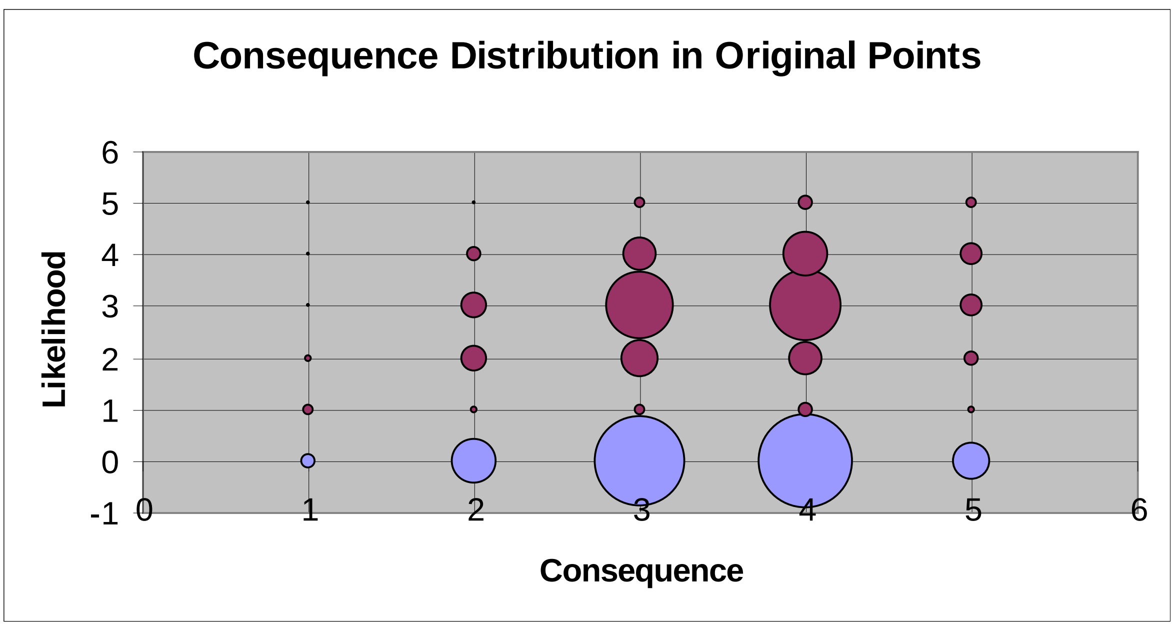

Figure 11: Consequence marginal distribution of original datum points

Consequence bias is theoretically predicted as occurring when an engineer seeks to communicate, through the placement of a risk datum point, that she/he is identifying the increased danger of risks as they can impact an entire corporation. On a personal basis, engineers usually experience greatly lesser losses, and so seem to perceive less risk.

Statistical Evidence. The original consequence counts in Table 5 show that most consequence counts are at C = 4.

Table 5: Consequence original datum point counts

Consequence Original Datum Points				
1	2	3	4	5
20	145	538	599	110

The C =1 counts are significantly less than C = 5 counts, and, the C =2 counts are significantly less than the C = 4 counts. Of course, some upward skewing of the original consequence data is expected, as the risk points address real concerns. However, the consequence count distribution is markedly different from the likelihood counts distribution; a Chi-Squared sum of 604.41 with 4 degrees of freedom gives ~0 probability that the distributions are equal.

An explanation for the spreading of the consequence distribution is that consequences in fact occur over a range of values. There is not one point value for a consequence, as consequences can occur over a range of magnitudes. For example, engine failure results in difference consequences, depending on other conditions.

ANALYSES AND OBSERVATIONS OF RISK MITIGATION

Figure 12 shows the new location of the risk datum points after mitigation efforts.

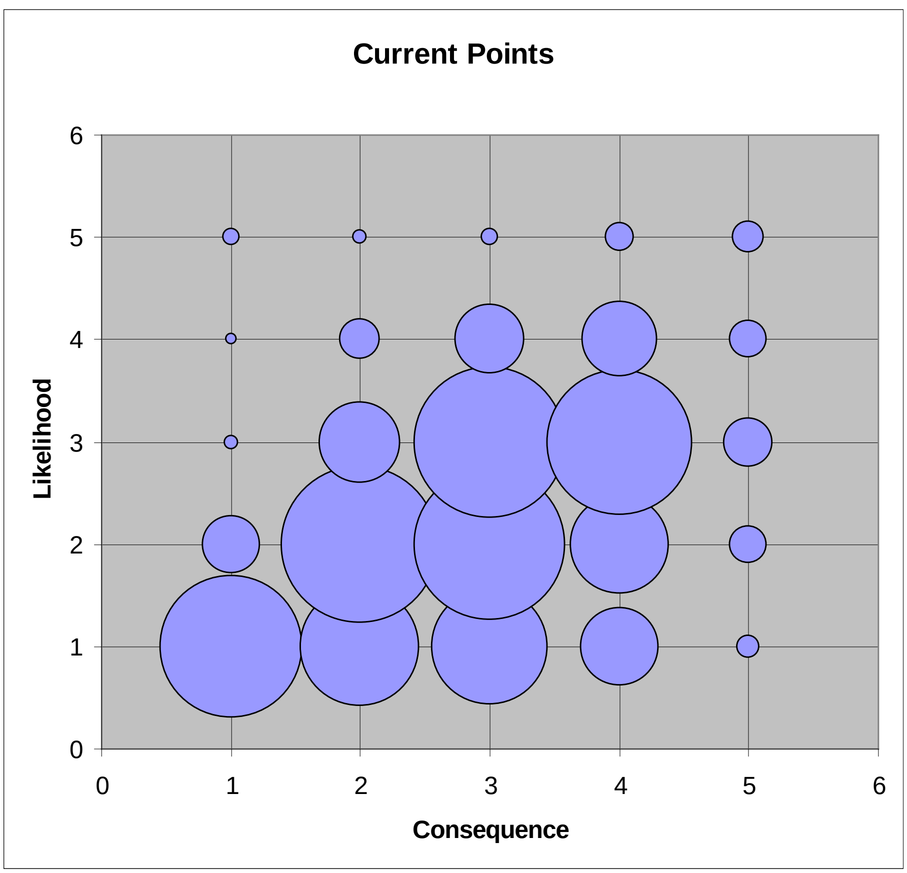

Figure 12: Mitigated, or current, risk datum points

Figure 13 shows the change of likelihood and consequence in the datum points.

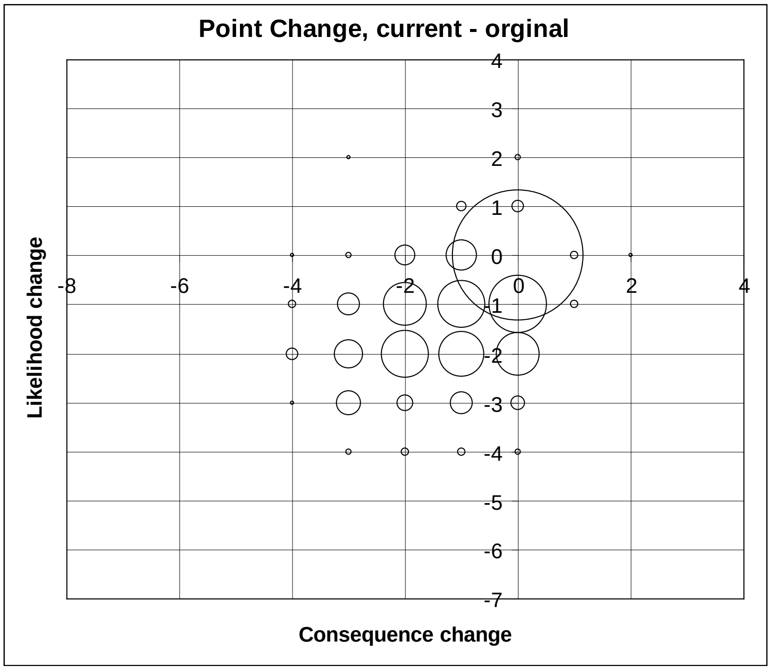

Figure 13: Change of likelihood and consequence in the datum points

Many Risk Datum Points Did Not Change Location. Explanation: Risk mitigation may have been difficult in many risks. Additionally, it is possible that many of the risk plans were not updated and were just abandoned by the risk originator.

No Risk Datum Points Were Moved To 0 Likelihood Or 0 Consequence. Explanation: No significant architectural changes that absolutely eliminated a particular risk were made. Many programs stop working on a risk when it reaches a certain level of risk verses consequences. Note: Because no risk was moved to 0 likelihood and 0 consequence, the word “abate” is not used in this paper.

Likelihood Is Mitigated More Than Consequence. Possible explanations: The usual risk situation in an aerospace program more readily allows a reduction in likelihood as opposed to consequence. Working engineers may be more able to reduce the likelihood of occurrence of a risk than the architecturally determined consequences.

Statistical Evidence: Consideration of all original and current risk point data yields Table 6.

Table 6: Likelihood mitigation versus consequence mitigation

Likelihood mitigation versus consequence mitigation				
	Likelihood		Consequence	
Total mitigation	1114		911	
Average mitigation	0.7890		0.6452	
99% confidence interval	0.7214	0.8566	0.5776	0.7128
Standard deviation	0.9850		0.9851	

The 99% confidence intervals show that the mitigation processes on likelihood and consequence are almost certainly different. Likelihood is mitigated more that consequence.

One school of thought is that you cannot ever reduce the consequence of a risk you can merely lower the likelihood of its occurrence. Not all risk practitioners subscribe to this theory, and there is no fast rule as to its application. The data seems to suggest that this group of engineers subscribes to this theory.

COST, SCHEDULE AND TECHNICAL RISK PERCEPTION AND MITIGATION

The original and current averages of cost, schedule and technical points are show in Figure 14.

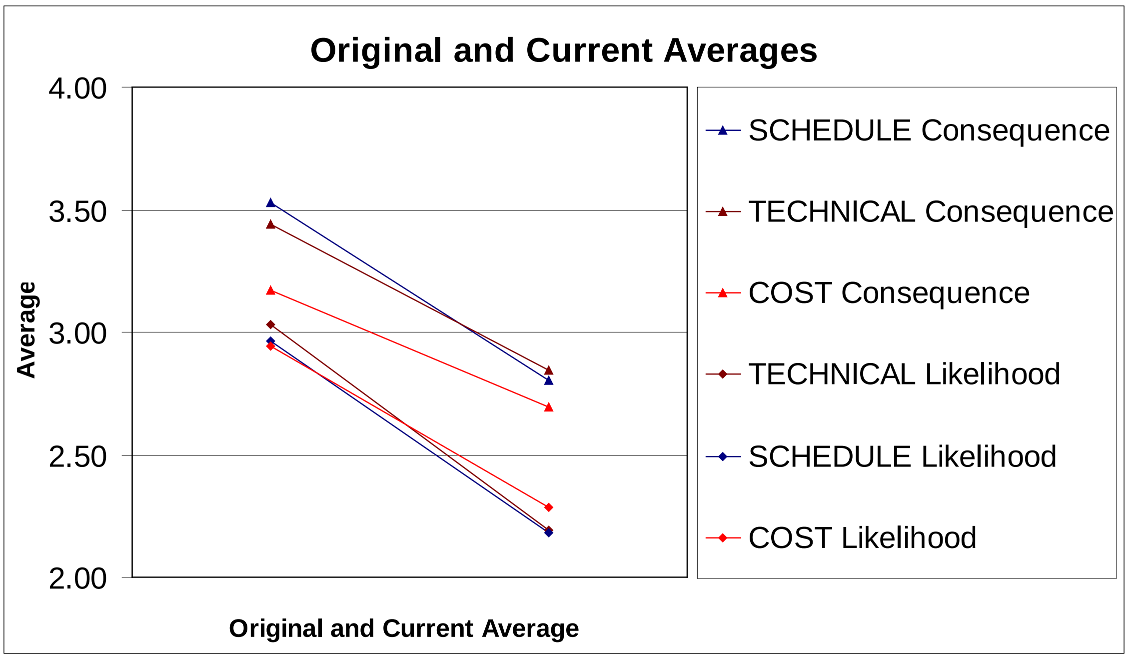

Figure 14: Average mitigation of cost, schedule and technical risks

Mitigation Of Likelihood. Figure 14 gives evidence that the line engineers originally perceives technical likelihood as greater than either schedule or cost likelihood. Subsequently, more effort is applied to mitigate technical likelihood than either schedule or cost likelihoods. In fact, cost likelihood is mitigated less than schedule likelihood.

Statistical Evidence: Figure 15 shows the order of likelihood mitigation emphasis from greatest to least: technical likelihood, schedule likelihood, and lastly cost likelihood.

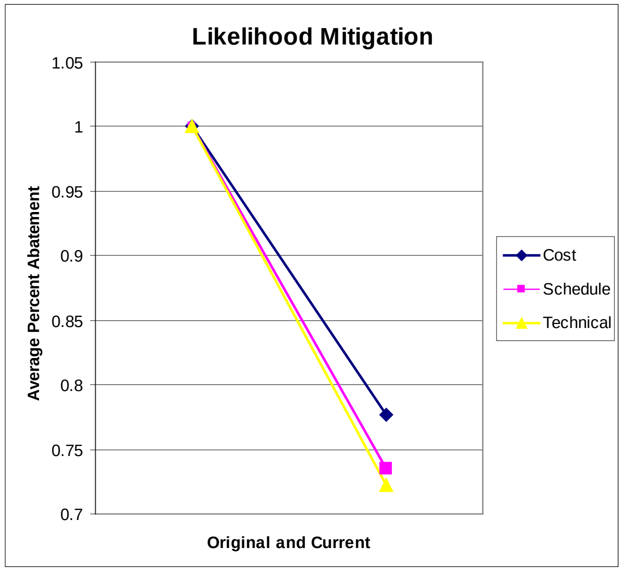

Figure 15: Likelihood mitigation compared

Table 7: Order of likelihood mitigation

Order of likelihood mitigation
1. Technical
2. Schedule
3. Cost

The data reveal differences in worker and manager foci that have been widely recognized since the beginning of management science. Namely, line engineers focus most on the technical risk likelihoods that they are most familiar with, they focus secondarily on schedule risk likelihood that management communicates to them, and least on cost risk likelihoods that are less frequently communicated by management. Lean models of value creation suggest that cognizance of cost risk likelihoods should be more effectively communicated.

Mitigation Of Consequence. Figure 14 shows that the line engineer originally perceives schedule consequence as greater than either technical or cost consequence.

Statistical Evidence: Figure 16 shows consequence mitigation emphasis, showing that the line engineer does more to mitigate schedule consequence than either technical or cost consequences. Cost consequence is mitigated less than technical consequence.

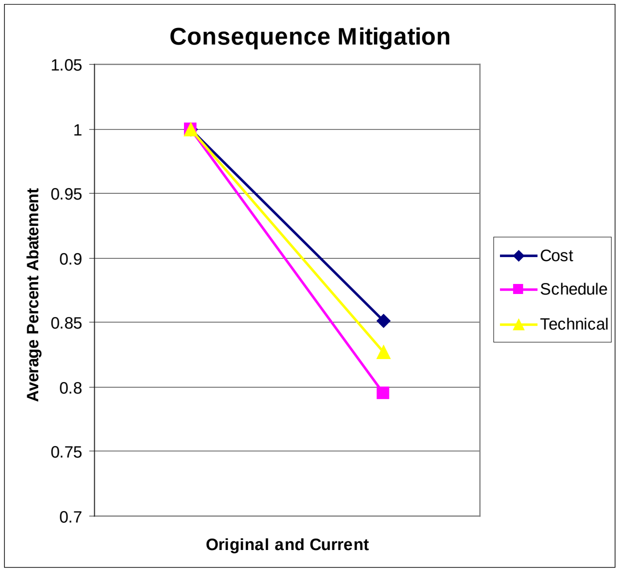

Figure 16: Consequence mitigation compared

The order of consequence mitigation emphasis is thus, from greatest to least: schedule consequence, technical consequence, and lastly cost consequence.

Table 8: Order of consequence mitigation

Order of consequence mitigation
1. Schedule
2. Technical
3. Cost

The data thus reveal that line engineers focus most on schedule consequences that effect their careers, they focus secondarily on technical consequences that effect their job performance reviews, and focus least on cost consequences that are more remote and associated with management. Lean models of value creation suggest that a higher cognizance of cost risk will be valuable at the engineering level.

CONCLUSION

To the authors' knowledge, this is the first time that the effects of cognitive biases have been documented within the risk cube. The data show clear evidence that probability and value translations, as likelihood and consequence judgments, are present in industry risk cube data. The straightforward steps in the development of this conclusion were as follows: 1) the translations were predicted by prospect theory, and, 2) historical data from 2 programs within 1 company confirmed the main predictions. Risk cubes have thus been shown to be, not objective number grids, but subjective, albeit useful, means to verify that risk items have received risk-mitigating attention. Confirmation of the presence of probability and value biases in risk data from other industries or companies can be the subject of future papers. The real world effects of using biased risk mitigation data should also be explored.

Biographies

William T. Siefert earned his B.S. in Business Administration in 1987 from the University of Redlands and his Masters in Systems Engineering from the University of Missouri-Rolla in 2007. His Masters Thesis was on “Cognitive Biases in Risk Management”. He is currently the Co-Chairman of the INCOSE Risk Management Working Group. He has been a practicing engineer in aerospace since 1973, working for several major corporations including The Boeing Company, where is currently employed.

Eric D. Smith is currently an Assistant Professor within the Industrial, Manufacturing and Systems Engineering (IMSE) Department at the University of Texas at El Paso (UTEP), and a Research Associate in the Research Institute for Manufacturing and Engineering Systems (RIMES). He earned a B.S. in Physics in 1994, an M.S. in Systems Engineering in 2003, and his Ph.D. in Systems and Industrial Engineering in 2006 from the University of Arizona in Tucson, AZ. His dissertation research lay at the interface of systems engineering, cognitive science, and multi-criteria decision-making.

References

Cosmides, L. and Tooby, J., *Are Humans Good Intuitive Statisticians after All? Rethinking Some Conclusions from the Literature on Judgment under Uncertainty*, Cognition 58 (1996), 1-73.

Kahneman, D. and Tversky, A., *Prospect Theory: An Analysis of Decision under Risk*, Econometrica 46 (2) (1979), 171-185.

Nobel, "The Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel 2002," Nobel Foundation, Stockholm, Sweden, 2002.

Schwarz, N., *Assessing Frequency Reports of Mundane Behaviors: Contributions of Cognitive Psychology to Questionnaire Construction*, Review of Personality and Social Psychology 11 (1990), 98-119.

Tversky, A. and Kahneman, D., *Advances in Prospect Theory: Cumulative Representation of Uncertainty*, Journal of Risk and Uncertainty 5 (1992), 297-323.