How to Measure Anything in Project Management

Hubbard Decision Research
2 South 410 Canterbury Ct
Glen Ellyn, Illinois 60137
www.hubbardresearch.com

Introduction

Topics for Today

- This is an introduction to basic concepts for measuring Project Management (PM).
- We will take a critical look at PM, how it is measured and then we will describe some solutions.
- We will treat this as an example of applying the ideas described in How to Measure Anything: Finding the Value of Intangibles in Business (HTMA).
- Some of the content are methods you can start using right away, others are aspirational and you should start on a path to master them.

I’m going to tend to focus on topics I don’t think are covered quite enough in PM.

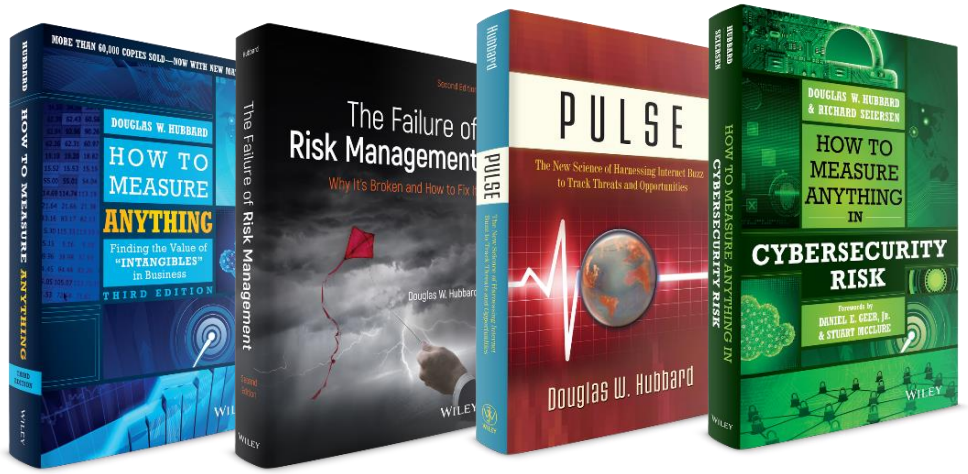
Hubbard Decision Research Background

In 200+ major analysis projects, HDR has been able to show that no matter how difficult the measurement and monetization problem appears to be, we find a way to evaluate it and communicate the results.

- The benefits and risks of ***dams on the Mekong River***
- Risks and benefits of ***Environmental policy*** for US farmers
- The benefits of ***Educational assistance*** in inner city schools
- The benefits of roads, schools and hospitals in Haiti and how to prioritize them for the ***United Nations***
- The relative value of ***R&D portfolios*** in aerospace, biotech, and pharma
- ***Logistics forecasts for the battlefield*** and the effectiveness of training for the US Military

(A Few) Information Technology Examples

- Risk of software development
- Value of better information access
- Cybersecurity risk
- Risk of obsolescence and optimal technology upgrades
- Performance metrics for the business value of applications



Introduction

Why Measure Projects?

Question: Why Measure Projects?
Answer: The risk of doing it wrong is high!

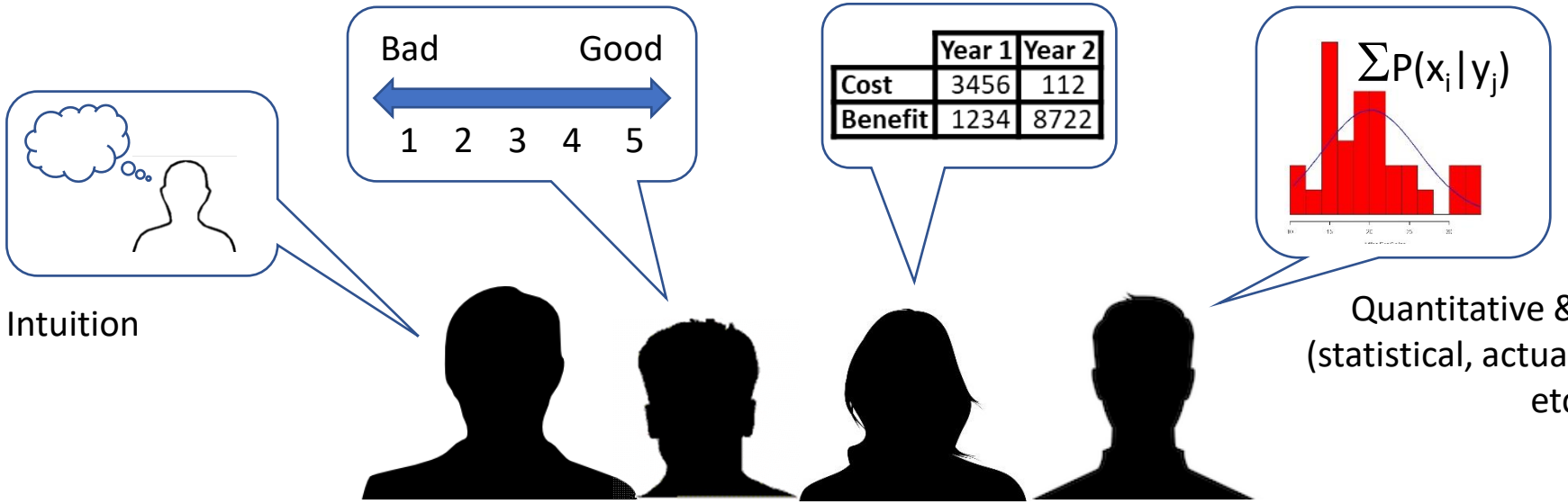
- They are large and time-consuming initiatives.
- Sometimes they have high visibility.
- Failure rates of projects from various sources: at least 10% maybe over 30% (failure includes cancellation but, in some cases, not meeting cost, schedule and deliverable goals).
- The cost of a failure can be greater than just wasted effort (loss in productivity, risk of customer loss, etc.).

Introduction

Types of Measurement Methods

Qualitative
(soft scores or “high/medium/low”)

Accounting-style
Cost estimate analysis
(point estimates, deterministic)

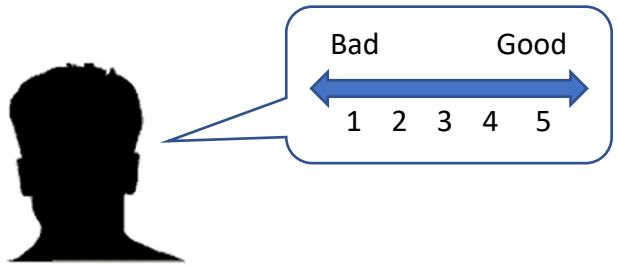


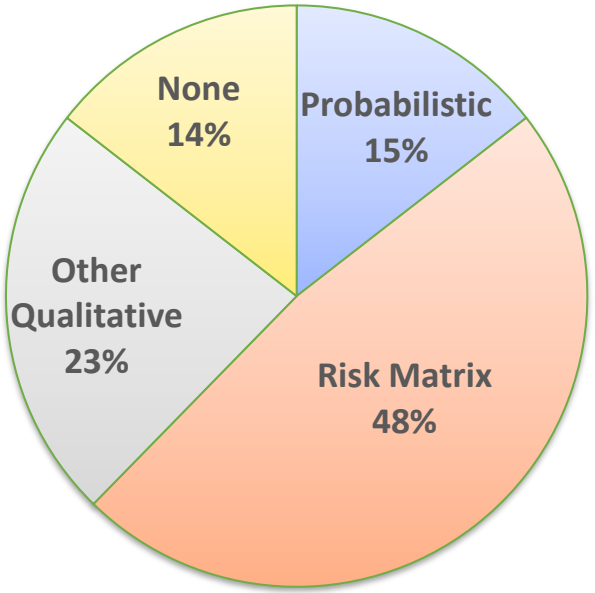
Expert Intuition

Quantitative & Probabilistic
(statistical, actuarial, simulations, etc.)

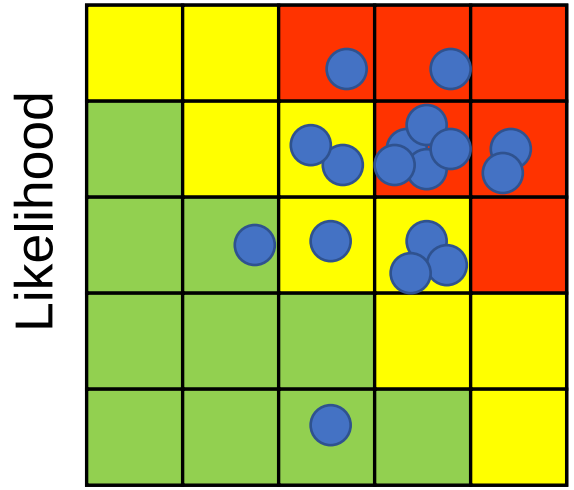
Do “Scores” and “Scales” Work?

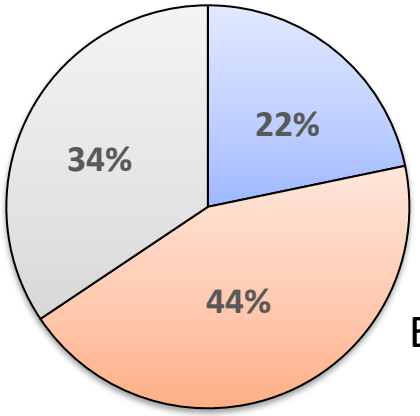
The Current Most Popular Method



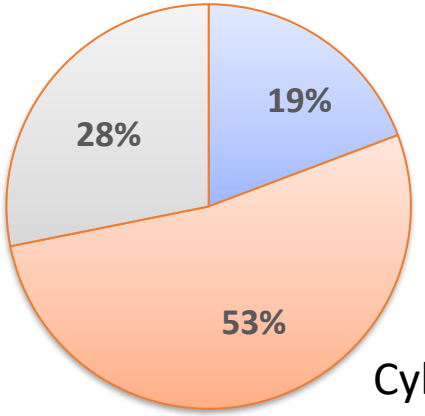


Project Management





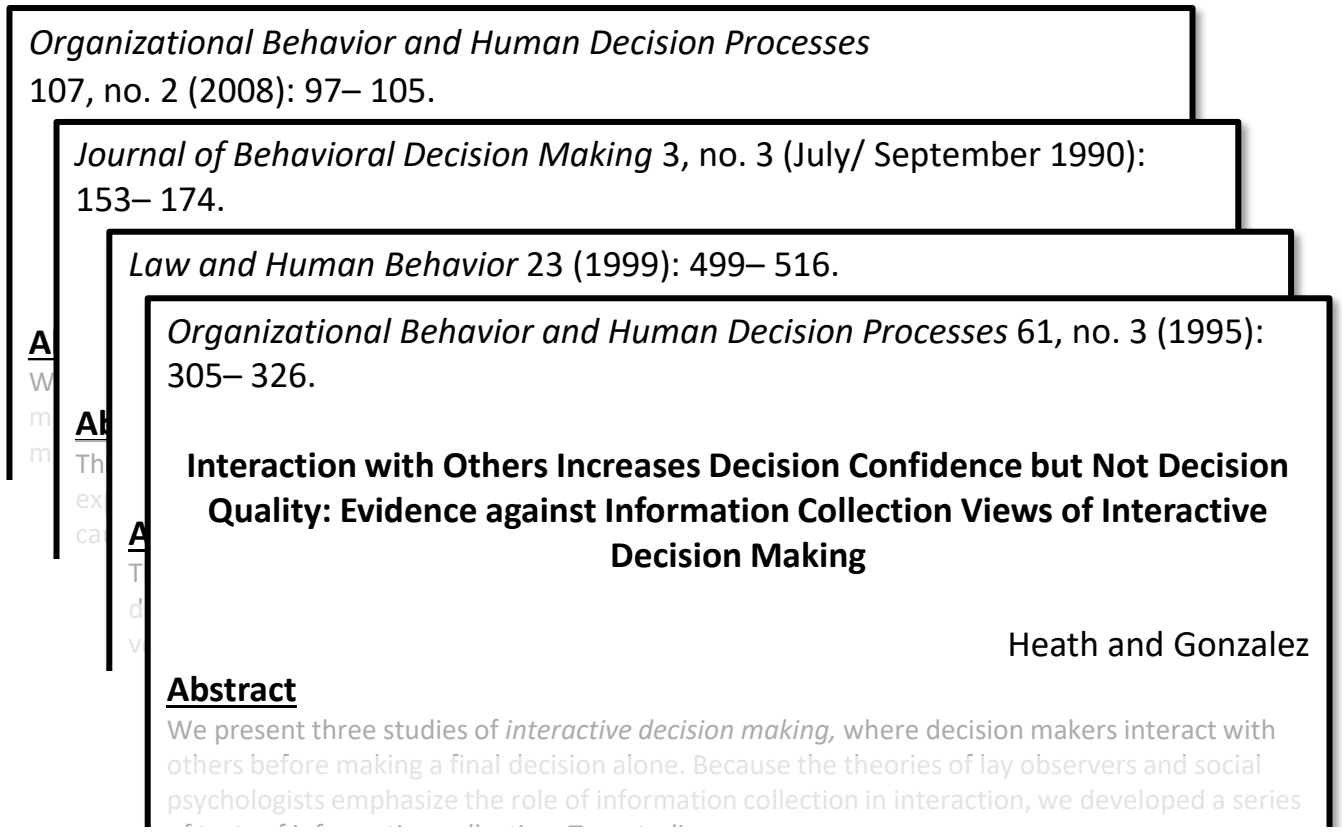
Enterprise Risk Management

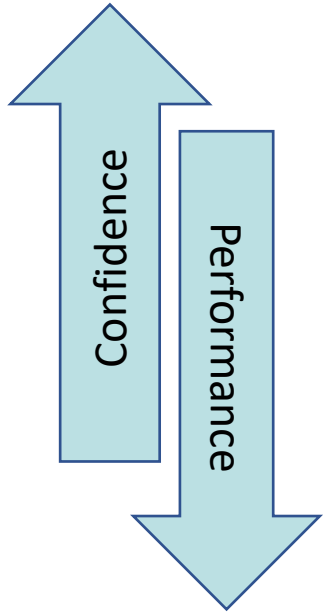


Cybersecurity

The Analysis Placebo

Confidence in decision making methods is detached from performance



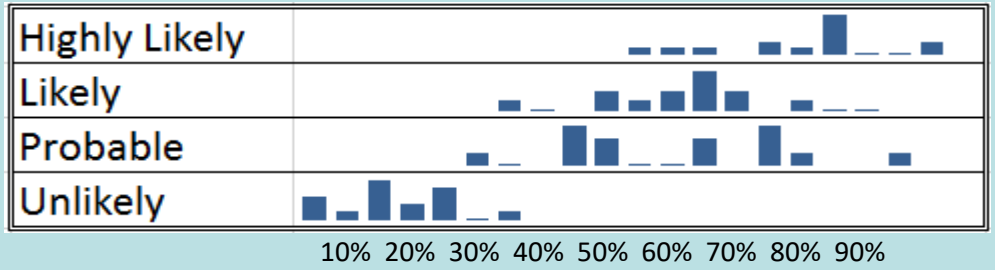


Do “Scores” and “Scales” Work?

Unintended consequences of simple scoring methods

Researchers uncovered several unintended consequences of simple ordinal scales and using words for probabilities.

- David Budescu and Dick Heuer (separately) researched the “illusion of communication” regarding interpretations of verbal labels for probabilities.

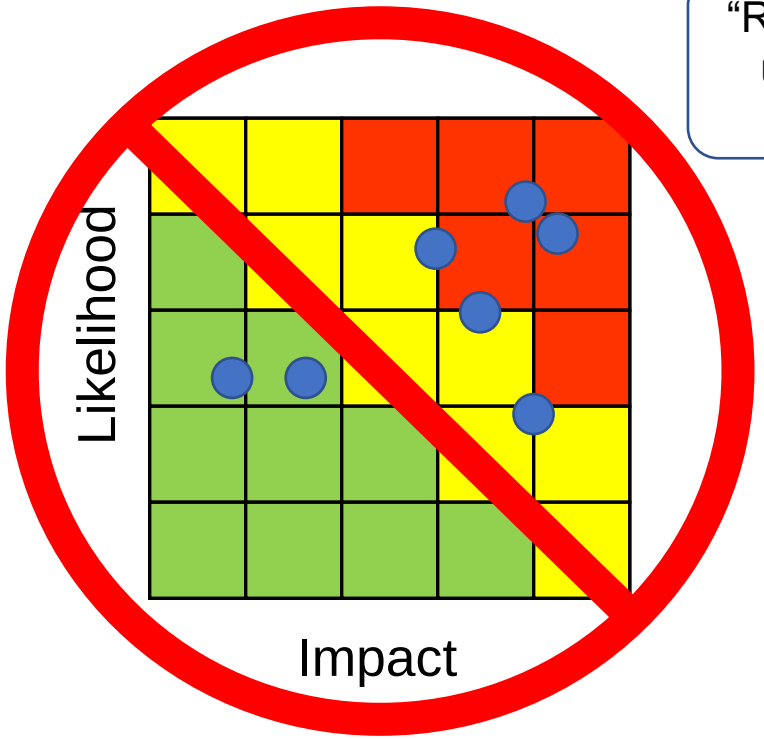


Craig R. Fox showed how arbitrary features of how scales are partitioned effects responses.

Example:

If “1” on a 5-point impact scale means “less than $1 million loss”, the share of that response is affected by the partition of *other* choices.

Do “Scores” and “Scales” Really Work?



“Risk Matrices should not be used for decisions of any consequence”

Society of Petroleum Engineers Economics &Management 6, no. 2 (April 2014): 56–66.

The Risk of Using Risk Matrices

P. Thomas, R. Bratvold, and J. E. Bickel

Abstract

The risk matrix (RM) is a widely espoused approach to assess and analyze risks in the oil & gas (O&G

in ris

the r

“[Risk Matrices] can be worse than useless”

Risk Analysis 28, no. 2 (2008).

What’s Wrong with Risk Matrices?

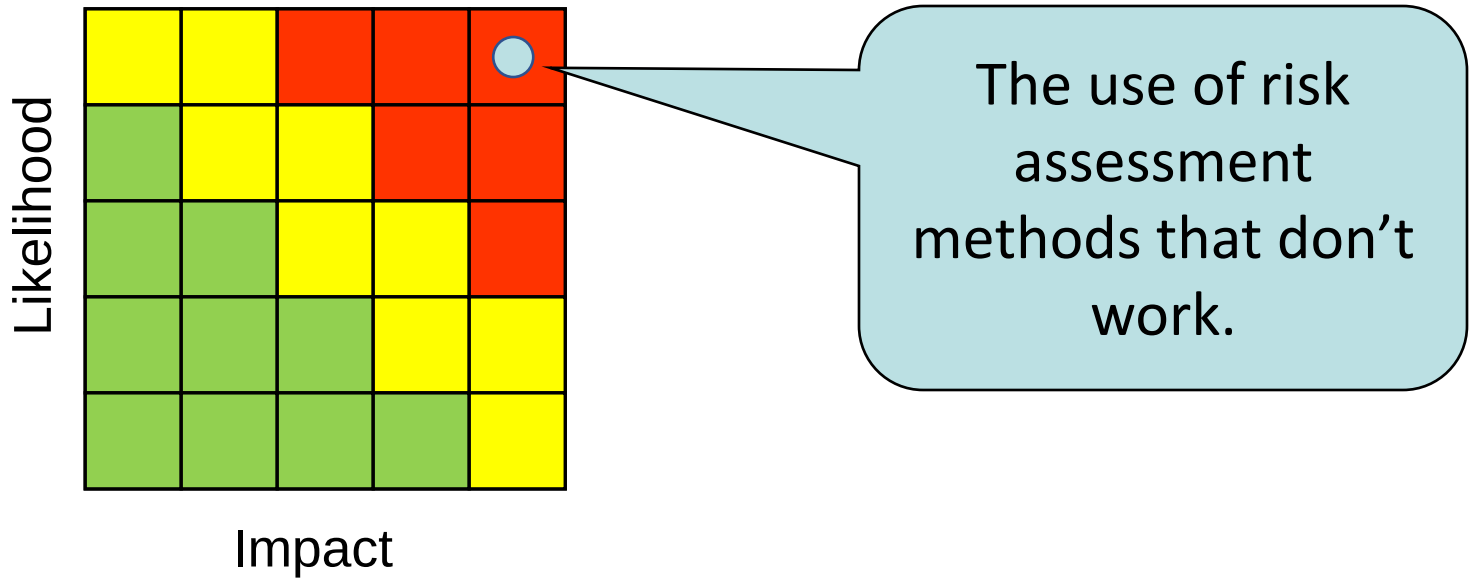
L. A. Cox, Jr.

Abstract

Risk matrices—tables mapping “frequency” and “severity” ratings to corresponding risk priority levels—are popular in applications as diverse as terrorism risk analysis, highway construction project management, office building risk analysis, climate change risk management, and enterprise risk management (ERM). National and international standards (e.g., Military Standard 882C and AS/NZS 4360:1999) have stimulated adoption of risk matrices by many organizations and risk consultants. However, little research rigorously validates their

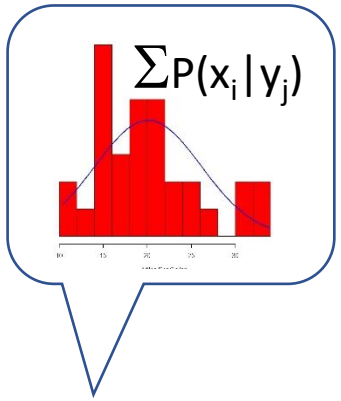
Do “Scores” and “Scales” Work?

The Only Risk Matrix You Need



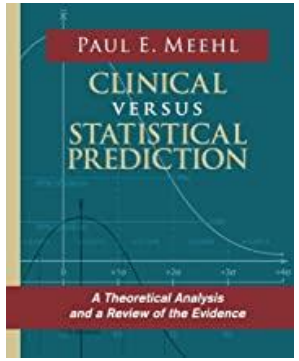
Experts vs. Algorithms

What the research says about statistical methods vs. Subject Matter Experts

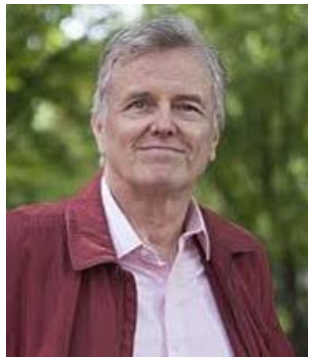


Paul Meehl assessed 150 studies comparing experts to statistical models in many fields (sports, prognosis of liver disease, etc.).

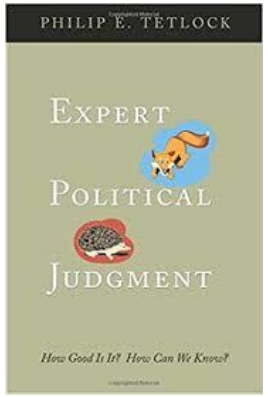
"There is no controversy in social science which shows such a large body of qualitatively diverse studies coming out so uniformly in the same direction as this one."



Philip Tetlock tracked a total of over 82,000 forecasts from 284 experts in a 20-year study covering politics, economics, war, technology trends and more.



"It is impossible to find any domain in which humans clearly outperformed crude extrapolation algorithms, less still sophisticated statistical ones."



So Why Don't We Use More Quantitative Methods?

Commonly stated reasons for not using quantitative methods

Have you heard (or said) any of these?

"We don't have sufficient data."

"Project cancellation is too complex to predict."

"Each situation is too unique and complex to apply scientific analysis of historical data."

"How do you know you have all the variables?"

The implied (and unjustified) conclusion from each of these is….

"Therefore, we are better off relying on our experience."

Measurement Misconceptions

CONCEPT of Measurement	The definition of measurement itself is widely misunderstood.
OBJECT of Measurement	The thing being measured is not well defined.
METHOD of Measurement	Many procedures of empirical observation are misunderstood.

Measurement Misconceptions

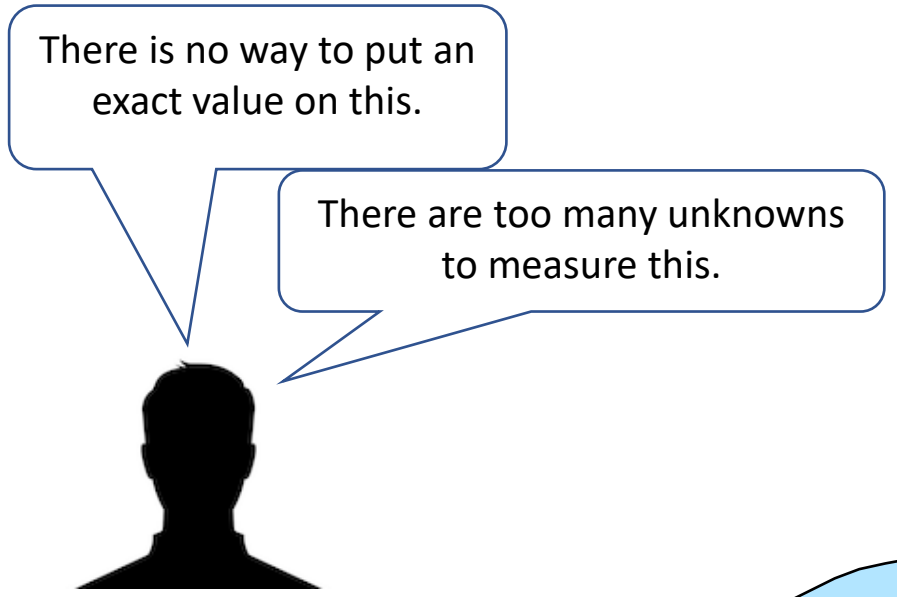
CONCEPT of Measurement	The definition of measurement itself is widely misunderstood.
OBJECT of Measurement	The thing being measured is not well defined.
METHOD of Measurement	Many procedures of empirical observation are misunderstood.

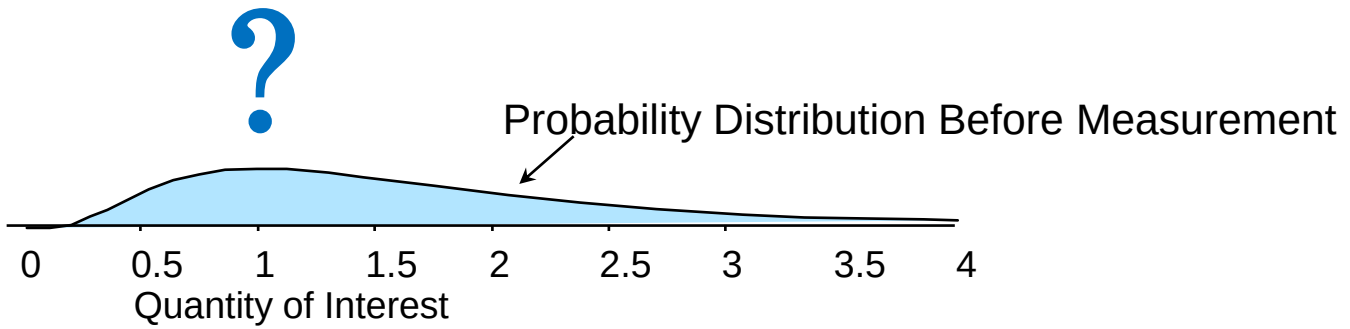
The Concept of Measurement

- What Measurement Really Means

It's not a point value.

Measurement: a quantitatively expressed reduction in uncertainty based on observation.



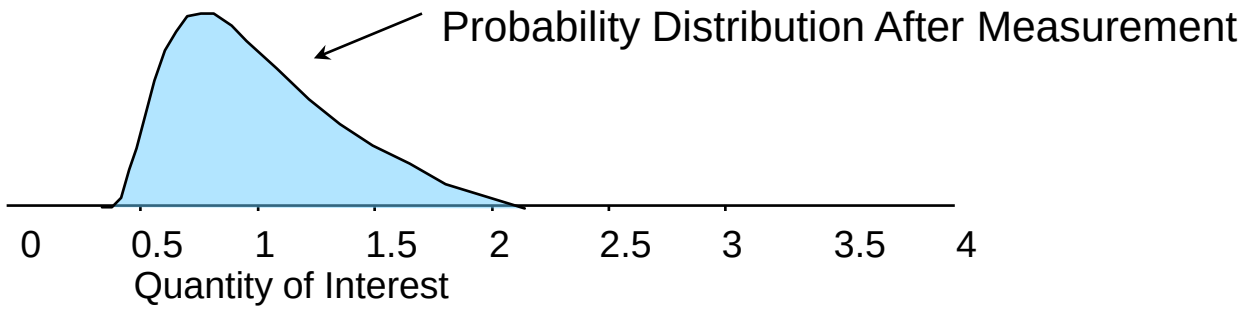


The Concept of Measurement

- What Measurement Really Means

It's not a point value.

Measurement: a quantitatively expressed reduction in uncertainty based on observation.



The Concept of Measurement

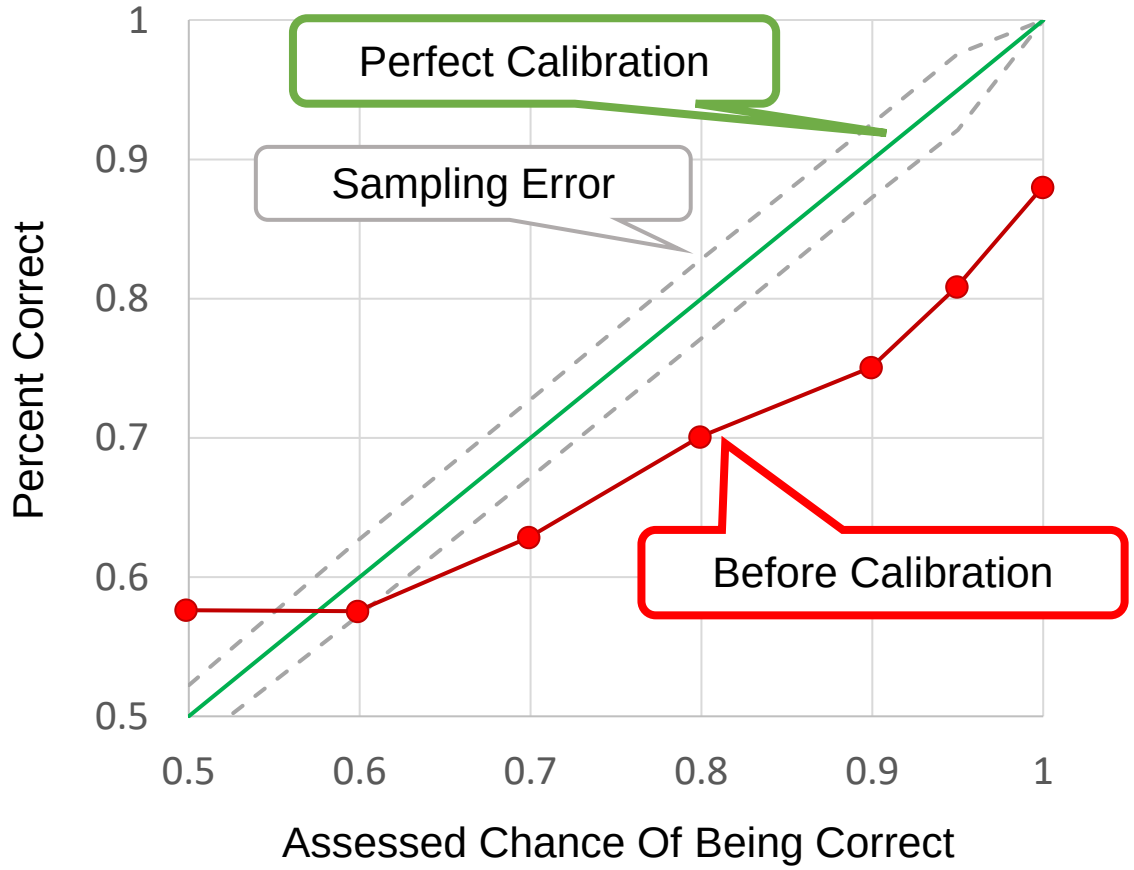
Calibrated Experts

"Overconfident professionals sincerely believe they have expertise, act as experts and look like experts. You will have to struggle to remind yourself that they may be in the grip of an illusion."

Daniel Kahneman, Psychologist, Economics Nobel

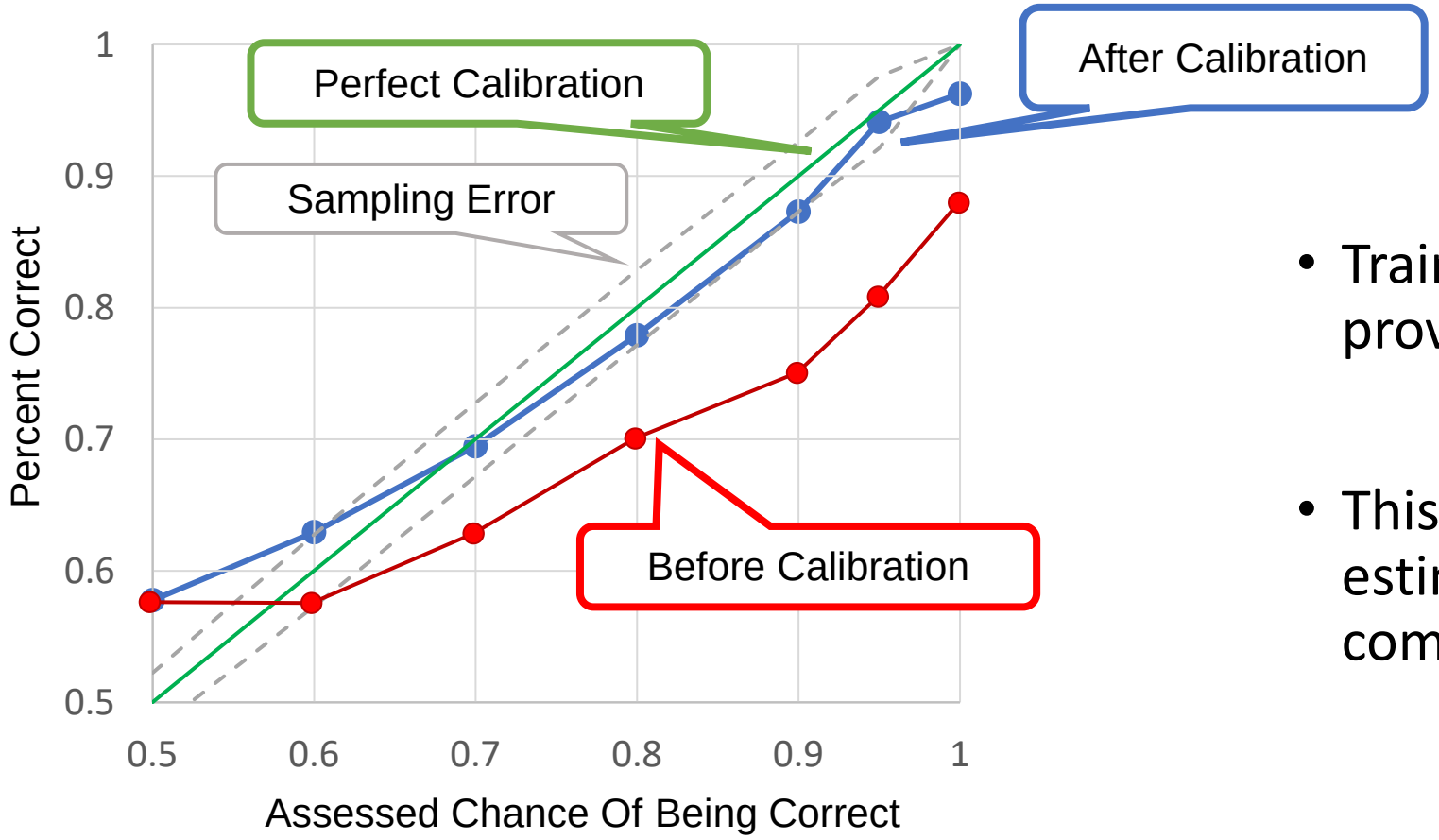
- Decades of studies show that most managers are statistically "overconfident" when assessing their own uncertainty.
- Studies also show that measuring *your own* uncertainty about a quantity is a general skill that can be taught with a ***measurable*** improvement.

Measuring Overconfidence



- HDR trained over 2,000 individuals in subjective estimation of probabilities.
- Almost everyone is overconfident on the first benchmark test.

Measuring Calibration Training

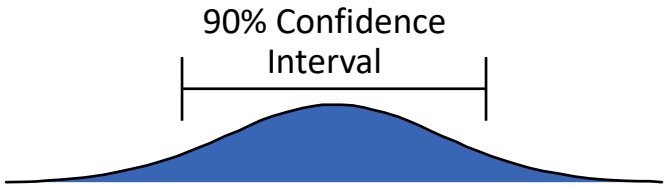


- Training improves the ability to provide calibrated estimates.
- This improves real-world estimates after training is complete.

The Concept of Measurement

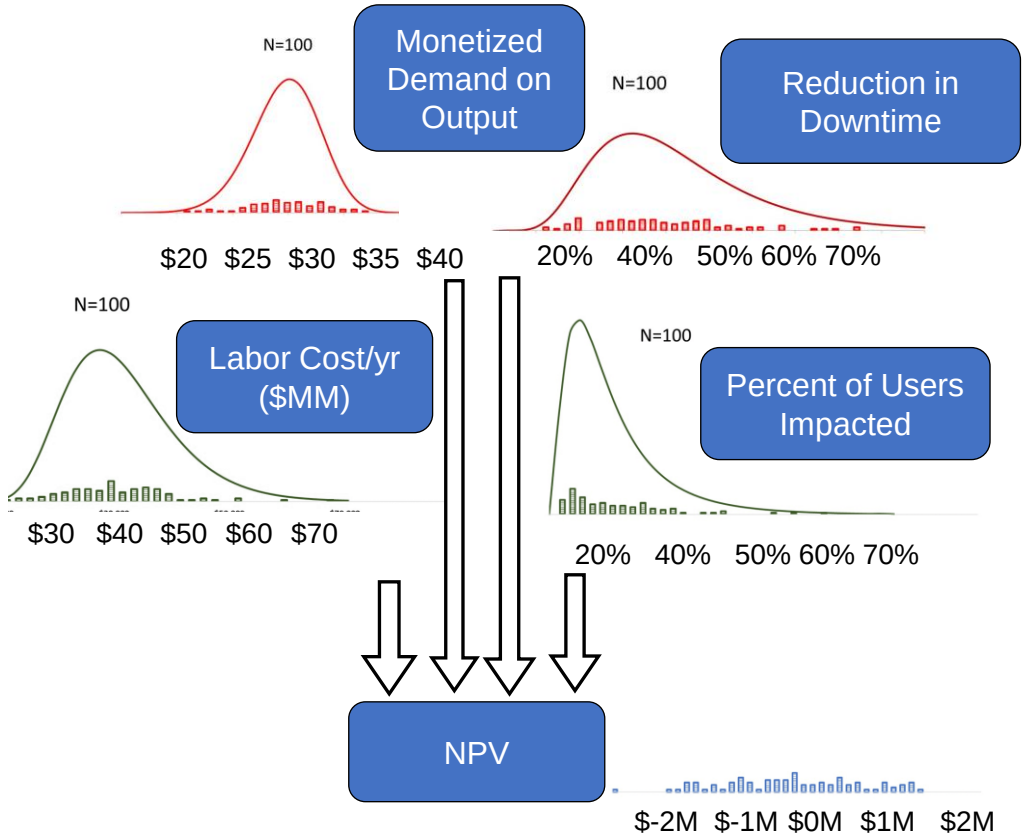
Overconfidence in Ranges

The same training methods apply to the assessment of uncertain ranges for quantities like the duration of project, the impact of a major data breach, etc.



Group	Subject	% Correct (target 90%)
Harvard MBAs	General Trivia	40%
Chemical Co. Employees	General Industry	50%
Chemical Co. Employees	Company-Specific	48%
Computer Co. Managers	General Business	17%
Computer Co. Managers	Company-Specific	36%
AIE Seminar (before training)	General Trivia & IT	35%-50%
AIE Seminar (after training)	General Trivia & IT	~90%

Doing The Math With Monte Carlo



Published in *International Journal of Forecasting*, 10 (1994), 495-906

Judgmental Decomposition: When Does It Work?

Donald G. MacGregor
Decision Research, Eugene, OR

Society of Petroleum Engineers (2000)

The Application of Probabilistic and Qualitative Methods to Asset Management Decision Making

G. S. Simpson, F. E. Lamb, J. H. Finch, and N. C. Dinnie

SSCAG/SCAF/EACE Joint International Conference (2008)

An Assessment of the Inherent Optimism in Early Conceptual Designs and Its Effect on Cost and Schedule Growth

D. Bearden, C. Freaner, R. Bitten, and D. Emmons

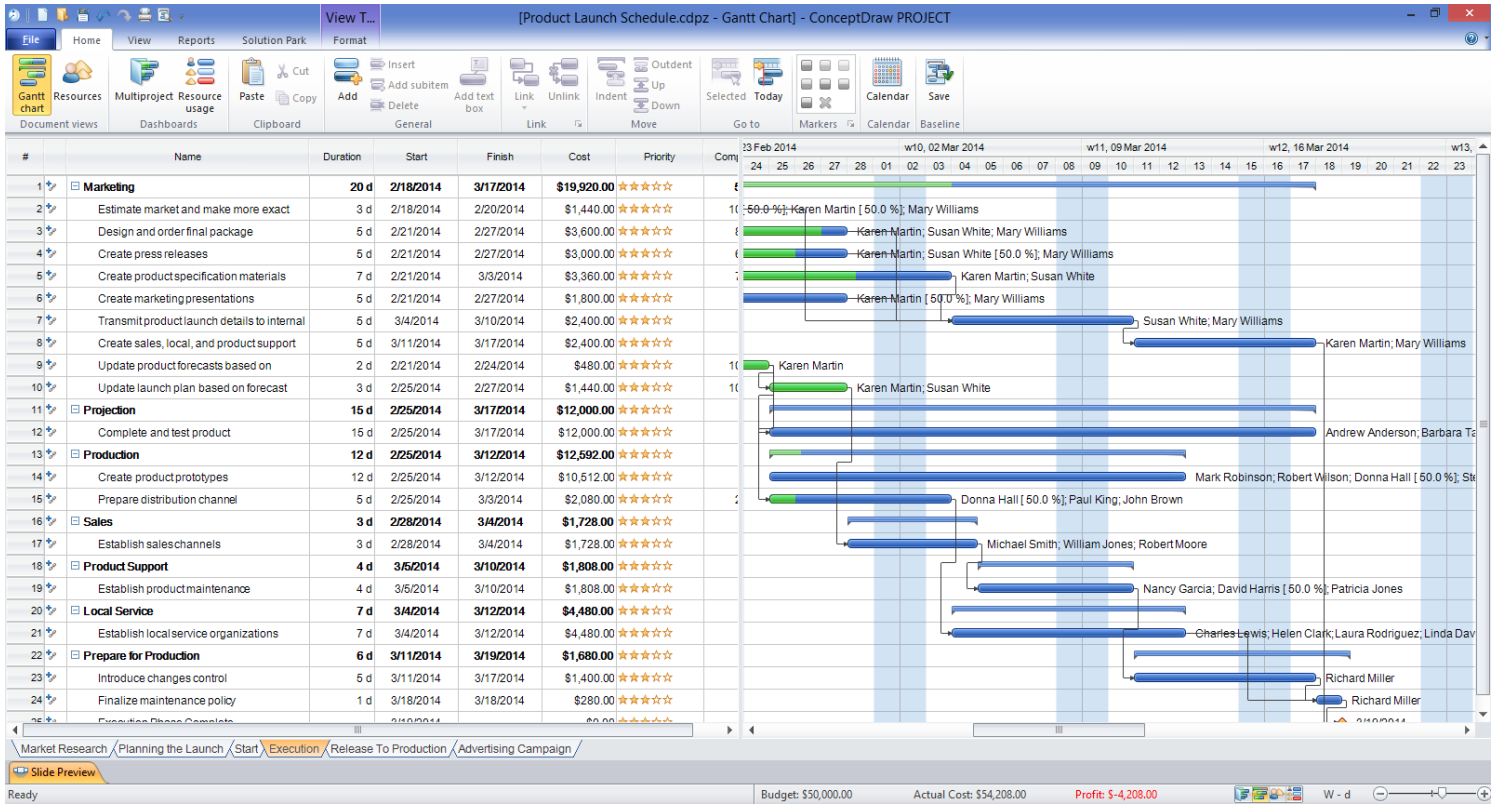
Abstract

When missions experience cost growth, cost estimators are often criticized for underestimating the cost of missions in the early conceptual design stage. The final spacecraft and instrument payload configuration at launch, however, can be significantly different as the project evolves, thereby leading to cost "growth" as

The Concept of Measurement

The Need for Simulation in Project Plans

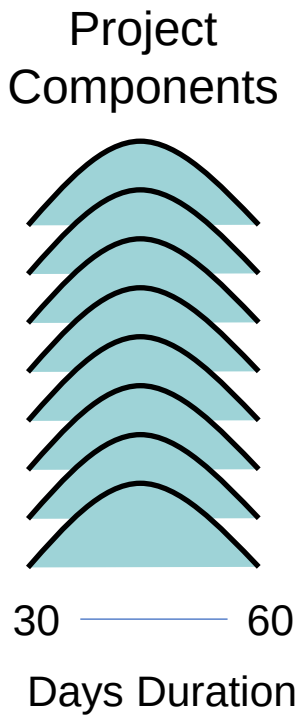
- A project plan even just moderately more complicated requires a simulation for even the most basic risk analysis.
- Furthermore, the simulation must include inputs like benefits and discrete/conditional events to support project decisions comprehensively.



The Concept of Measurement

Aggregating Uncertainty in Project Plans

Some project planning methods estimate durations based on “expected” time values. This is a classic “Flaw of Averages” problem (Sam Savage, Stanford).



- Eight project tasks all have to be completed before going to the next phase of the project.
- The duration of each task is an average of 45 days but could be within +/- 15 days of that.
- What is the chance the start of the next phase will be after 45 days?
- Answer: 99.6%

The Value of Information & The Measurement Inversion

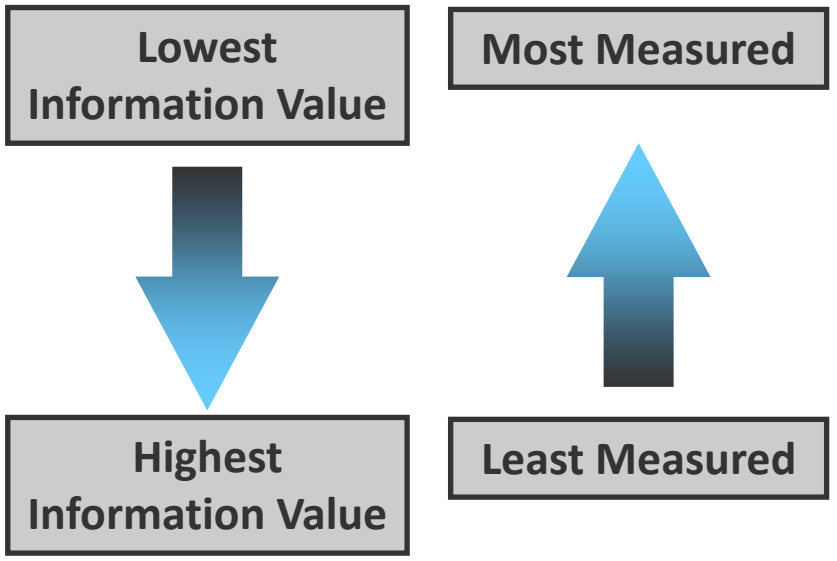
$$EVI = \sum_{i=1}^{k} p(r_i) \max\left[\sum_{j=1}^{z} V_{1,j}\, p(\Theta_j|r_i), \sum_{j=1}^{z} V_{2,j}\, p(\Theta_j|r_i), \ldots \sum_{j=1}^{z} V_{l,j}\, p(\Theta_j|r_i),\right] - EV^*$$

Or more simply: The chance of being wrong times the cost of being wrong.

The economic value of measuring a variable is usually inversely proportional to the measurement effort.

HDR has observed a "Measurement Inversion" in nearly every industry, profession and type of decision model we've every made.

The cure for starts with knowing which variables are the highest information value.



The Concept of Measurement

The Measurement Inversion

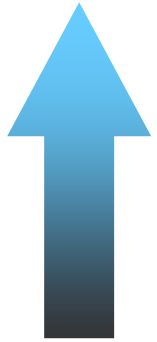
Examples of Measurement Inversions from Information Technology Projects

Lowest Information Value

Highest Information Value

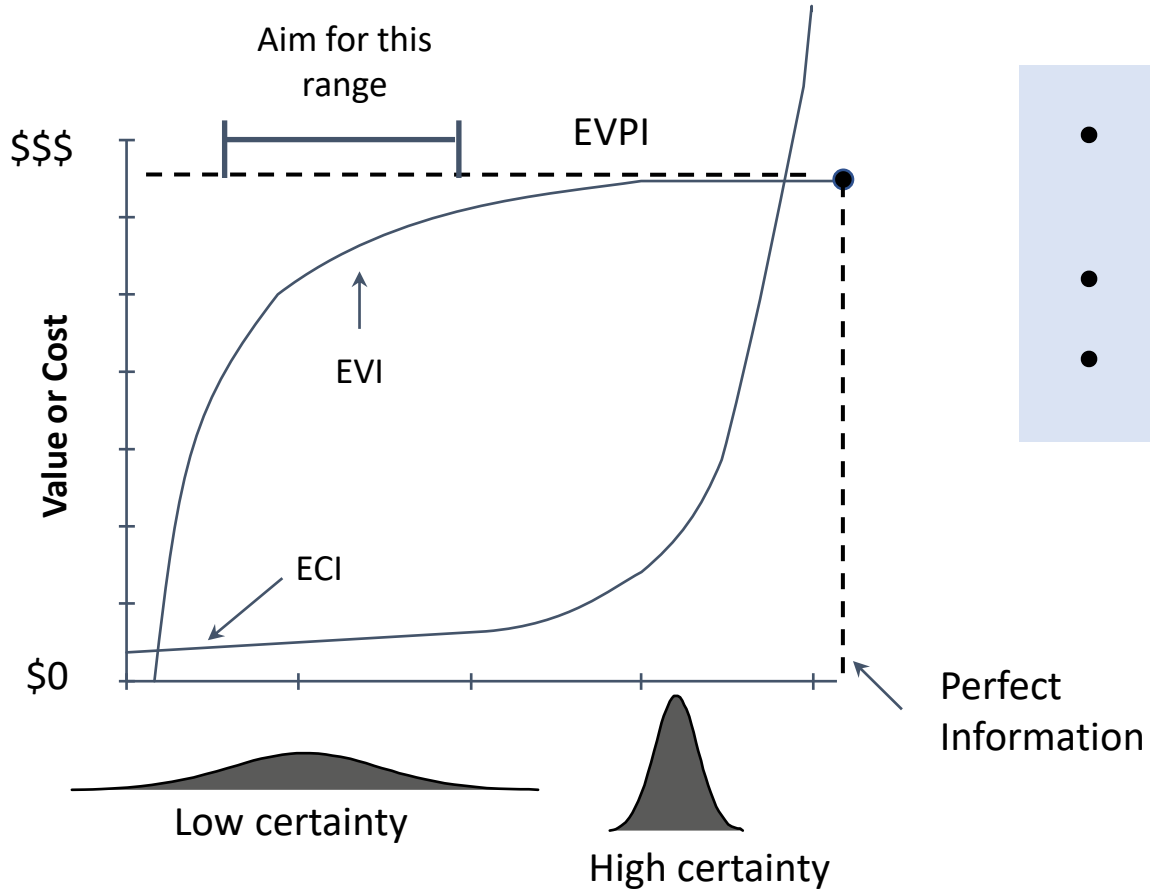
- Initial cost
- Long-term costs
- Cost saving benefit other than labor productivity
- Labor productivity
- Revenue enhancement
- Technology adoption rate
- Project completion

Most Measured



Least Measured

Information Value vs. Information Cost



- EVPI – Expected Value of Perfect Information
- ECI – Expected Cost of Information
- EVI – Expected Value of Information

If you know almost *nothing*, almost *anything* will tell you *something*.

Measurement Misconceptions

CONCEPT of Measurement	The definition of measurement itself is widely misunderstood.
OBJECT of Measurement	The thing being measured is not well defined.
METHOD of Measurement	Many procedures of empirical observation are misunderstood.

The Object of Measurement

Clarifying the Problem

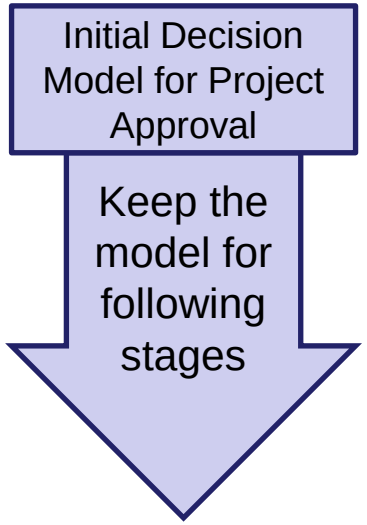
1. Why do you care? (What decision could depend on the outcome of this measurement?)
2. What do you see when you see more of it? (Describe it in terms of observable consequences, then units of measure.)
3. How much do you know about it now?
4. At what point will the value make a difference?
5. How much is additional information worth?

If you can answer the first three, you can usually compute the last two.

The Object of Measurement

Project Management Decisions

- We can think through three categories of decisions. Your specific decision will depend on where you are in the project.
- Your initial business case is not “throw away.” It should stay with the project for its lifespan.

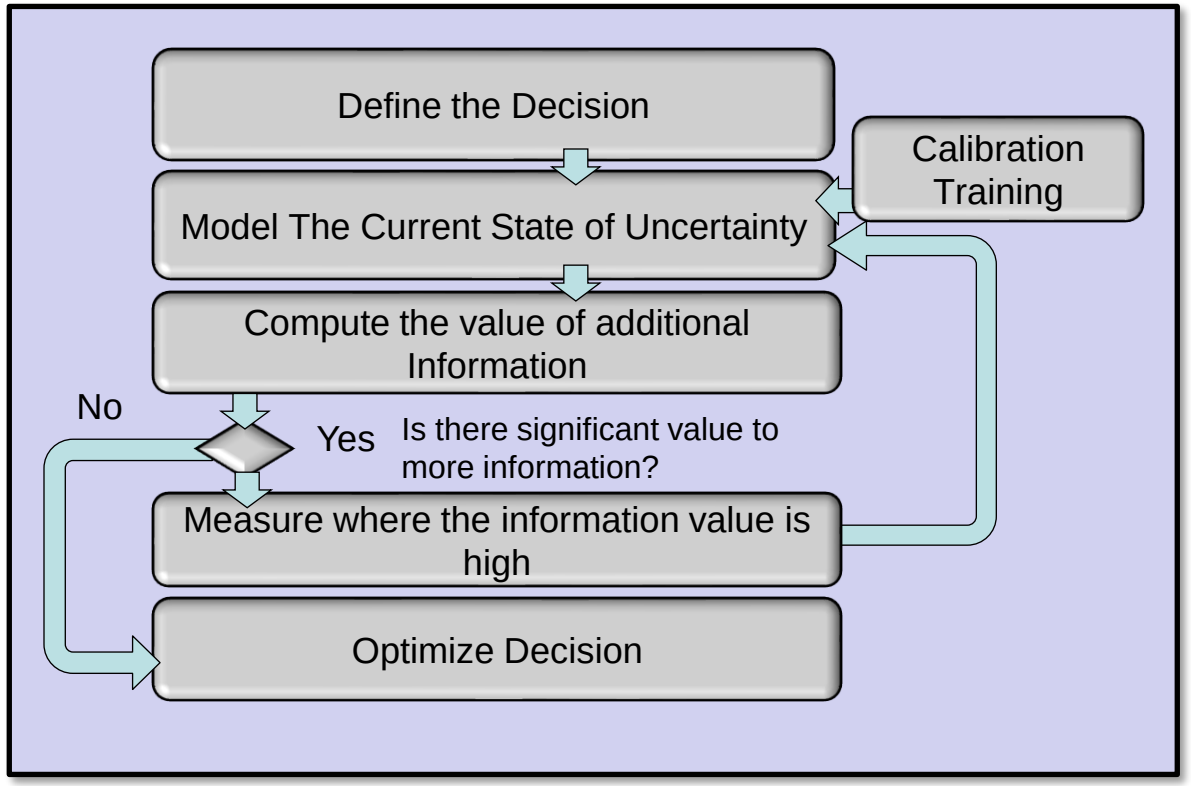


When	Examples of Decisions
Before	Should I engage in this specific effort?
During	Should I change course including: stop the project, reduce the features, change the managers, etc.
After	Do I need to do more? (Also informs future project approval decisions.)

The Object of Measurement

A General Procedure for Measurement

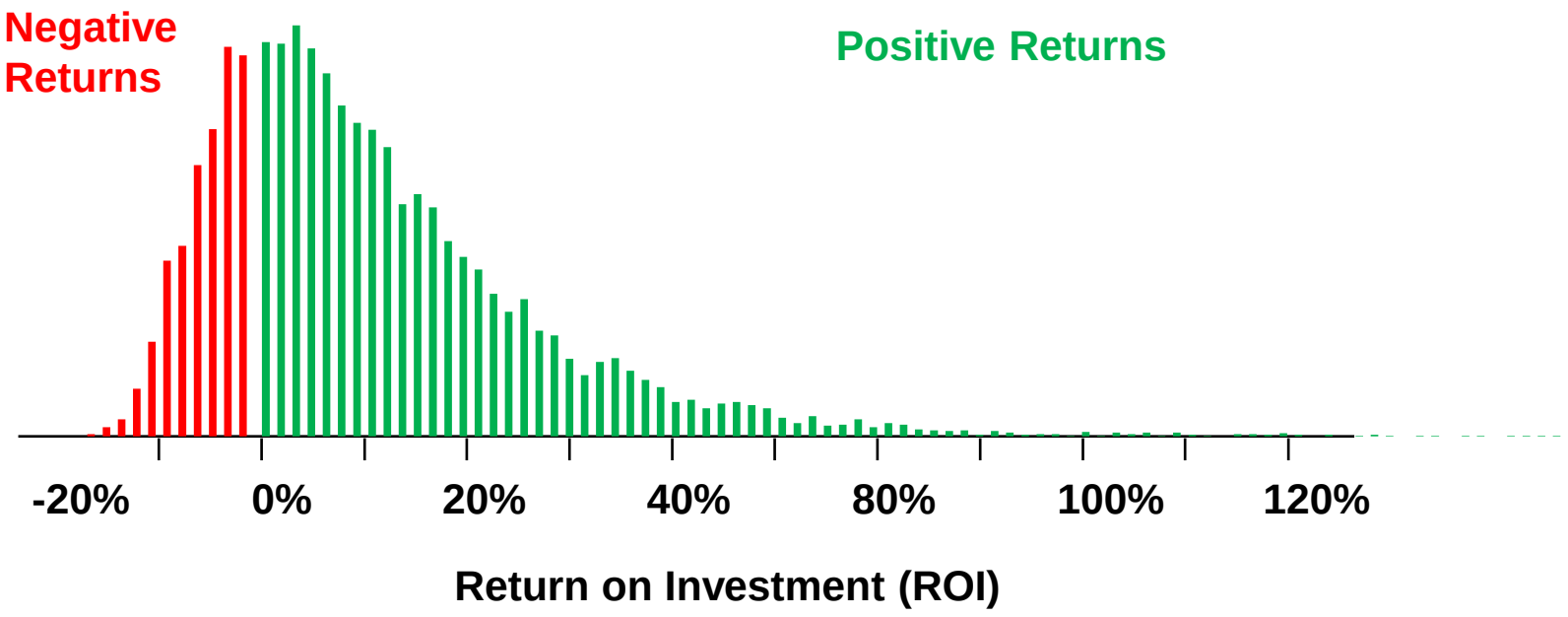
AIE quantifies and then optimizes decisions by focusing measurements where it matters most.



The Object of Measurement

Optimizing the Decision

- When the inputs to a decision model are uncertain, the output should be uncertain – this is what simulations are for.
- Is this a “good” distribution or a “bad” one? How would you know?



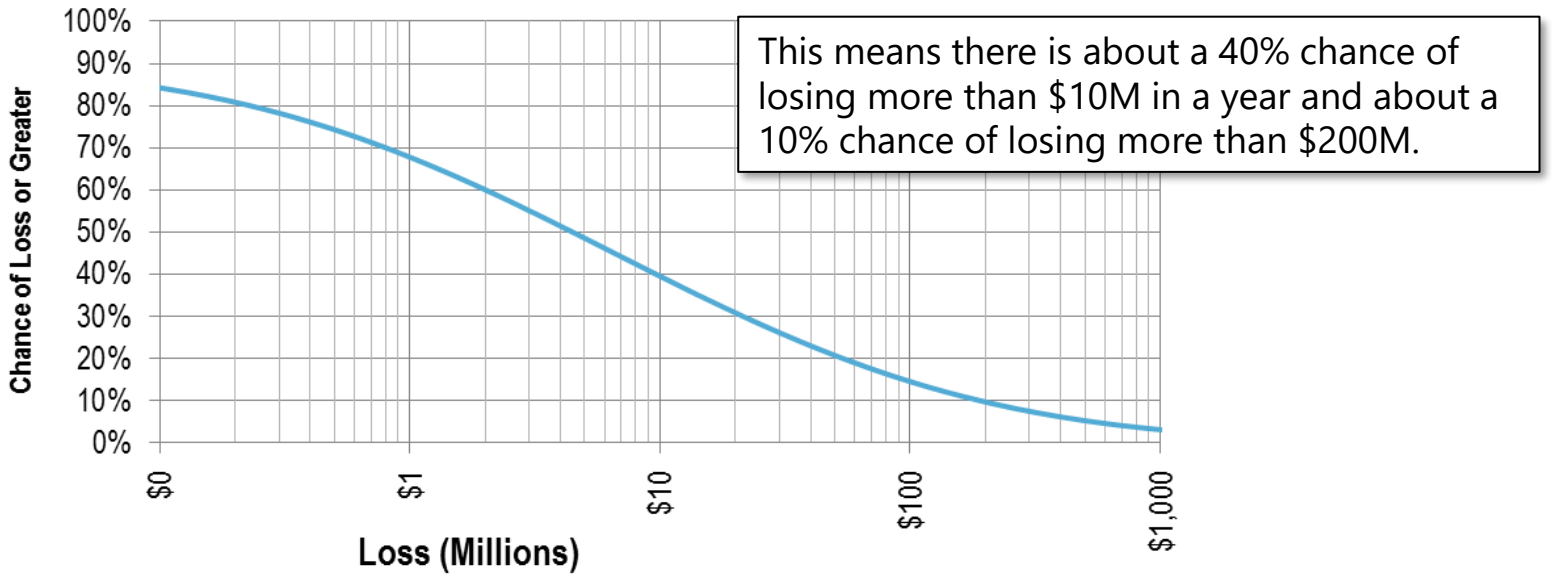
What Measuring Risk Looks Like

The Loss Exceedance Curve

What if we could measure risk more like an actuary? For example, “The probability of losing more than $10 million due to security incidents in 2016 is 16%.”

What if we could prioritize security investments based on a “Return on Mitigation”?

	Expected Loss/Yr	Cost of Control	Control Effectiveness	Return on Control	Action
DB Access	$24.7M	$800K	95%	2,832%	Mitigate
Physical Access	$2.5M	$300K	99%	727%	Mitigate
Data in Transit	$2.3M	$600K	95%	267%	Mitigate
Network Access Control	$2.3M	$400K	30%	74%	Mitigate
File Access	$969K	$600K	90%	45%	Monitor
Web Vulnerabilities	$409K	$800K	95%	-51%	Track
System Configuration	$113K	$500K	100%	-77%	Track

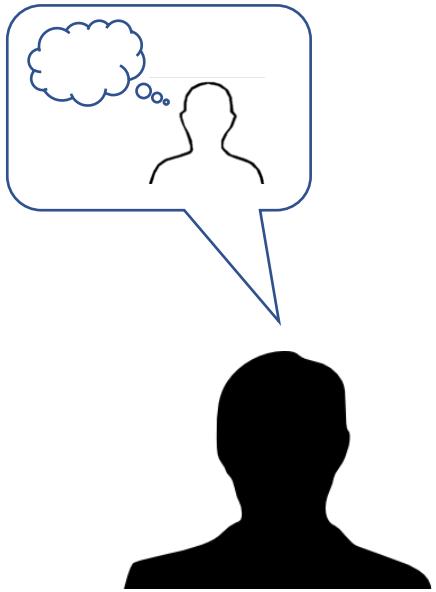


This means there is about a 40% chance of losing more than $10M in a year and about a 10% chance of losing more than $200M.

The Method of Measurement

Why Does Our Risk Tolerance Change?

Decision makers are also inconsistent regarding their own aversion to risk.



Neuron Vol. 47, (2005): 763–770

The Neural Basis of Financial Risk Taking

Camelia M. Kuhnen and Brian Knutson

Journal of Personality and Social Psychology
2001, Vol. 81, No. 1, 146–159

Copyright 2001 by the American Psychological Association, Inc.
0022-3514/01/$5.00 DOI: 10.1037//0022-3514.81.1.146

Fear, Anger, and Risk

Jennifer S. Lerner
Carnegie Mellon University

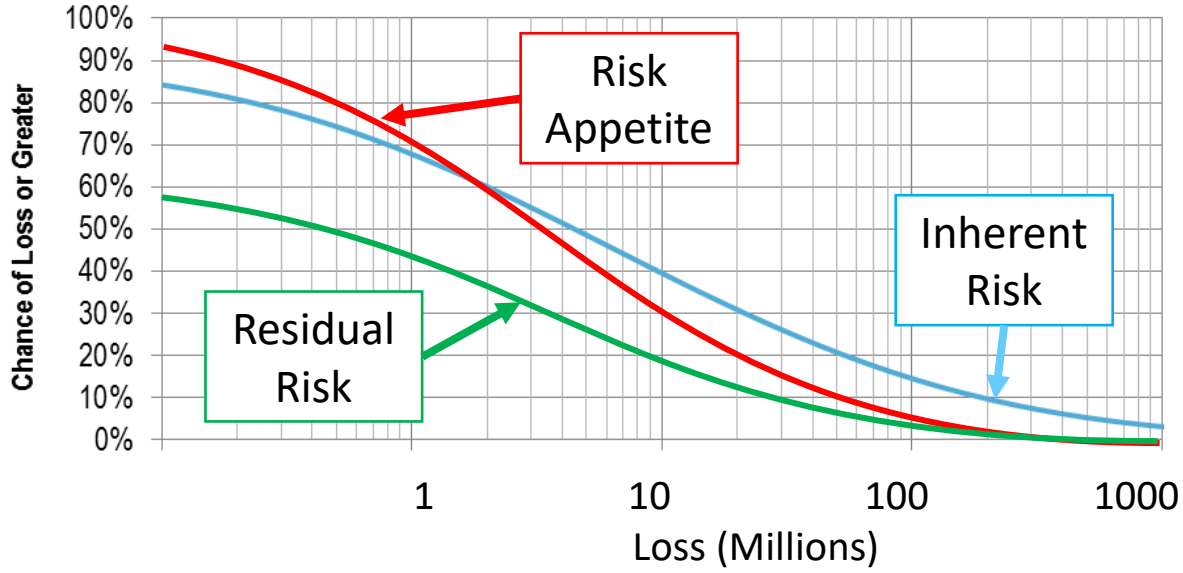
Dacher Keltner
University of California, Berkeley

Factor	Risk Aversion
Being around smiling people	↓
Recalling an event causing fear	↑
Recalling an event causing anger	↓
A recent win in an unrelated decision	↓
A recent loss in an unrelated decision	↑

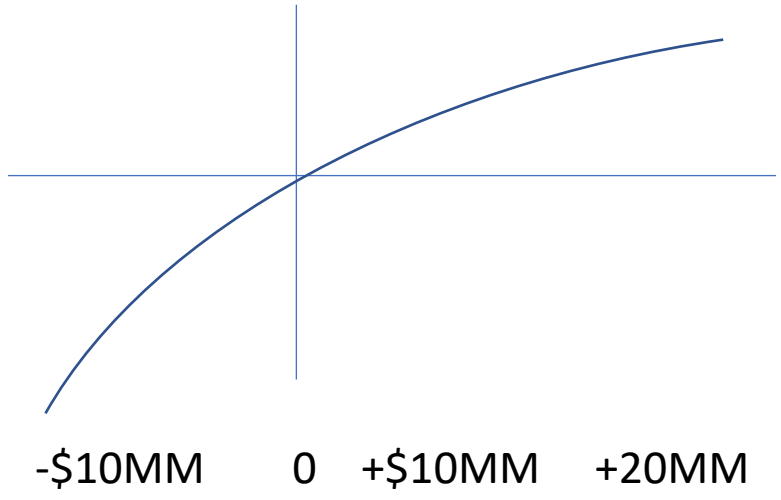
Quantifying Risk Tolerance

Monetizing risk and risk reduction is a basis for computing "Return on Mitigation"

Loss Exceedance Curve



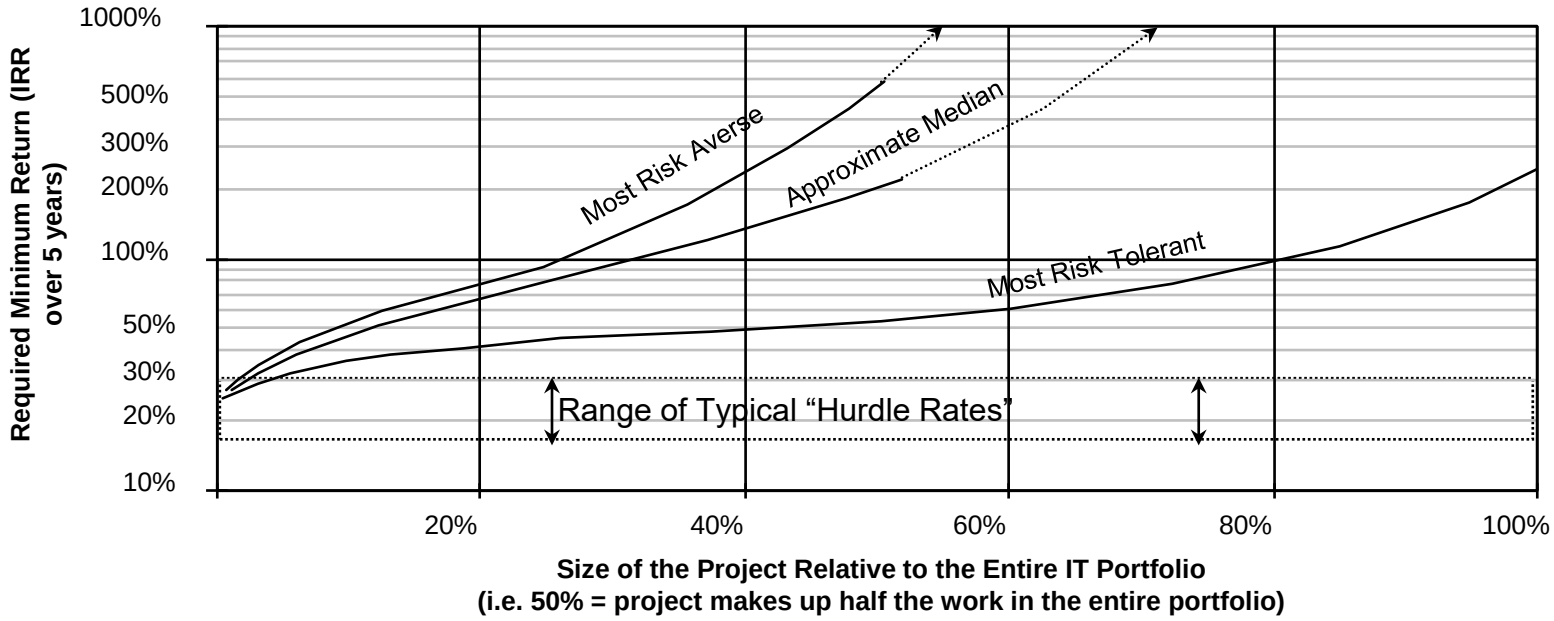
Expected Utility



The Object of Measurement

Expert Inconsistency in Estimates & Risks

- Adjusting for risk causes some previously-acceptable projects to be rejected.
- Also, some low return but low risk projects would now be acceptable.
- More projects with “intangible” benefits are now economically justified.
- The net result: A completely reshuffled deck of IT project approvals.

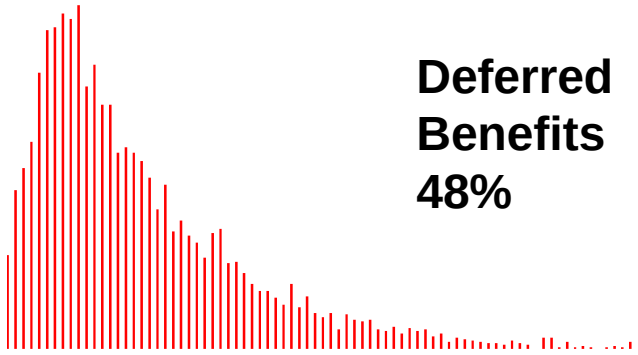


The Object of Measurement

One Decision During a Project: Managing Scope

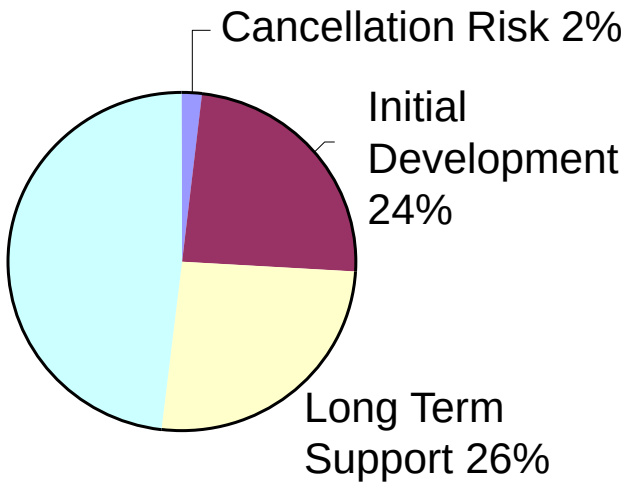
- One reason for “scope creep” may be that the true cost of adding additional features to software in development is greatly underestimated.
- If costs are computed at all, they usually consider only initial development.

Actual Case: Cost of Adding Feature Which Extends Delivery by One Month
(Avg. Proportions in Simulation Shown in Pie Chart)



Deferred Benefits 48%

$ Cost of Delay

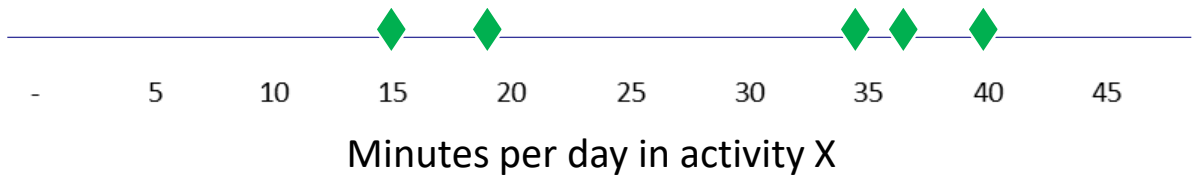


Measurement Misconceptions

CONCEPT of Measurement	The definition of measurement itself is widely misunderstood.
OBJECT of Measurement	The thing being measured is not well defined.
METHOD of Measurement	Many procedures of empirical observation are misunderstood.

A Sample of 5

- You randomly sample 5 people out of a company of 10,000 people and they spend an amount of time in a specific activity as shown by the data points below.
- Can this be a statistically significant estimate of time spent in this activity?
- Is it possible to estimate the chance the median time spent per person per day is between 15 and 40 minutes?



The Method of Measurement

Intuitions About Samples Are Wrong

- There are widely held misconceptions about probabilities and statistics – especially if they vaguely remember some college stats.
- These misconceptions lead many experts to believe they lack data for assessing uncertainties or they need some ideal amount before anything can be inferred.

"Our thesis is that people have strong intuitions about random sampling…these intuitions are wrong in fundamental respects...[and] are shared by naive subjects and by trained scientists"
Amos Tversky and Daniel Kahneman, Psychological Bulletin, 1971

The Method of Measurement

The “Math-less” Statistics Table

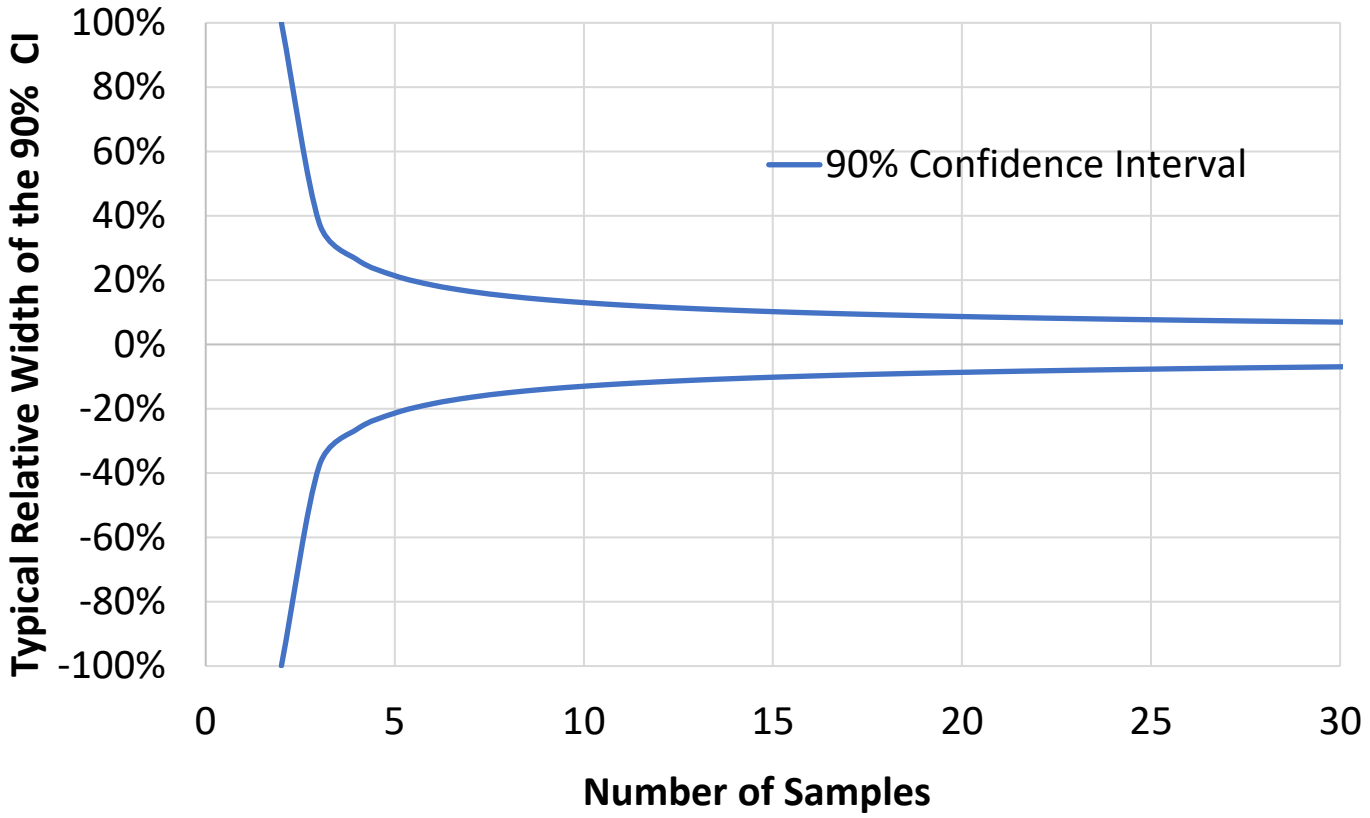
Approximate 90% Confidence Interval	
Sample Size	N^{th} largest & smallest sample value
5	1^{st}
8	2^{nd}
11	3^{rd}
13	4^{th}
16	5^{th}
18	6^{th}
21	7^{th}
23	8^{th}

- Simple Measurement Takeaway - This table makes estimating a 90% confidence interval of a population median easy.
- The Rule of Five: There is a 93.75% chance that the median of any population is between the smallest and largest values in a random sample of five.
- This table expands on the Rule of Five. If you take 16 random samples of something, the 5^{th} largest and 5^{th} smallest values of that sample set approximate a 90% confidence interval.

The Method of Measurement

How Much Samples Can Tell Us

The graph below shows the average of relative reduction in uncertainty as sample sizes increase by showing the 90% CI getting narrower and narrower with each sample according to the student-t method.



With a few samples, there is still high uncertainty but…

… each new sample reduces uncertainty a lot and the first few samples reduce uncertainty the most when initial uncertainty is high.

As number of samples increases, the 90 % CI get much narrower, but each new sample reduces uncertainty only slightly and beyond about 30 samples you need to quadruple the sample size to cut the error in half.

Practical Lessons

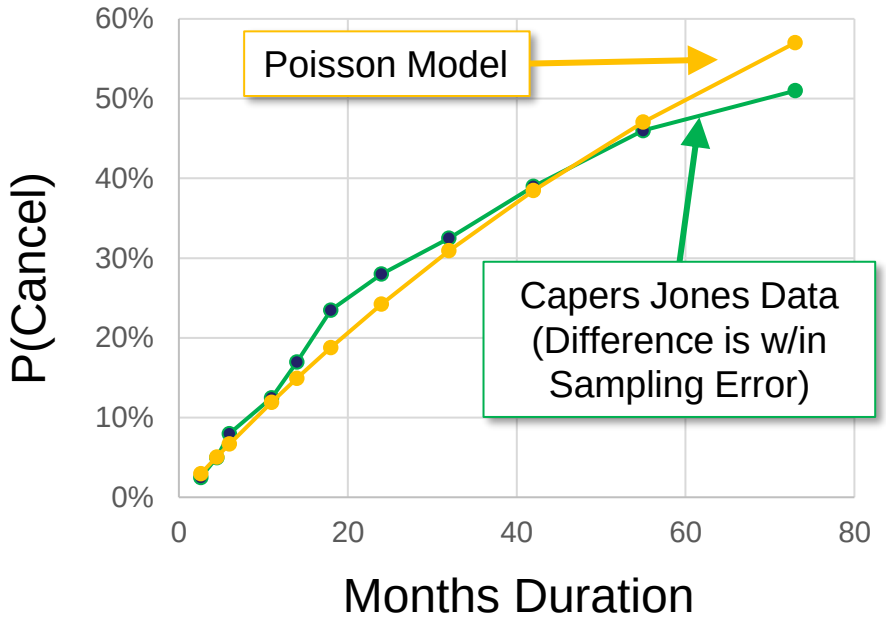
Here are a few key things I've learned measuring the "immeasurable"

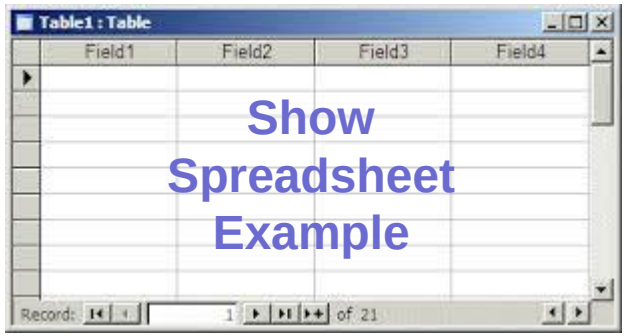
- It's been measured before.
- You *have more* data than you think and you *need less* data than you think.
- Your probably need *different* data than you think.
- Decision makers can always define observable consequences.
- Decision makers understand it just fine if explained well.
- The best investment in most portfolios was better measurements of investments.

The Method of Measurement

The Distribution of Canceled Projects vs. Duration

- Duration alone is the single best predictor of project cancellation. (Duration is also affected by several other factors normally associated with project failure – sponsorship, proper project management, quality and avoiding rework, etc.)
- Duration seems to follow a “Poisson” distribution – as if “cancellation events” are evenly distributed through time and longer projects are more likely to contain one of these events.





The Method of Measurement

You Have More Data Than You Think

You have a lot of data in the organization. Even samples of some of this data can be revealing indicators of communication, involvement, etc.

- Calendars
- Working group discussion threads
- Project deliverables
- Frequency and cost of errors/rework
- Utilization of various systems
- Turnover
- Milestones on schedule

The Method of Measurement

What Project Management Should Really Manage

Project completion is consistently a high value measurement, followed by adoption and benefits. Therefore:

- Manage cancellation by managing duration - Project management should be about managing duration to avoid uncontrolled cancelation events, and ensuring the utility of the outcome.
- Manage user adoption and benefits – User involvement will generally increase project benefits, which is also a responsibility of the project manager.
- Fail early - Someone, if not project managers, needs to be in the position of asking, “Do we still need this?” based on changing external factors and evolving knowledge of the project difficulty.

If the Risk is significant (it usually is), consider doing the following:

- Reduce the size and functionality of the proposed system - focus on fewer high-return features.
- Define “Independently Justifiable Phases” (IJP) and quicker, iterative development methods (Agile, Lean, etc.).
- Wait until specific uncertainties in the environment subside - e.g. major mergers, reengineering, etc.
- Wait to tackle big projects until proper skills are developed and methods are in place.
- “Off the Shelf” looks better when risk is considered.
- Invest more on a proper economic analysis of the largest project investments - this should reduce uncertainty about critical quantities.
- Include deferred benefits in any estimate of scope creep costs.

The Method of Measurement

Value of Quantitative Analysis for Projects

- The cost of analysis routinely comes in below 1% and has always been under 2% of the investment size - including initial training.

- Considering the risk of bad project approval decisions, this would be entirely appropriate (and some types of projects exceed this).

- Quantitative analysis is not necessarily more time consuming that some qualitative methods. (One of the reasons this analysis is efficient is we conduct a Value of Information Analysis - we only measure what is economically justified).

- Using the standard information value calculation for the value of AIE analysis, quantitative analysis itself was the best investment of all the IT investments we analyzed - *very conservative* measures of payoffs put $20 to every $1 spent on AIE.

The Method of Measurement

What to Do Next

Things you can do now or very soon:

- Drop the use of “scores” and “matrices” – define actual observables in units of measure and quantify risk with probabilities.
- Identify the specific decisions you are trying to support.
- Build a decision model/business case for the project and keep it for the life cycle of the project and what it builds.
- Get calibrated so you can quantify your uncertainty.

Things to strive toward (the effort is easily justified for large projects or even small but frequent projects):

- Learn to model the uncertainty of a decision in a simulation – evolve the model with more detail over time.
- Learn to compute the value of information.
- Learn a few more simple statistical methods – especially models for what experts usually estimate.

Thank you for Your Time!

Questions?

Doug Hubbard

Hubbard Decision Research

dwhubbard@hubbardresearch.com

www.hubbardresearch.com

Measure What Matters.

Make Better Decisions.

Supplementary Material

Hubbard Decision Research
2 South 410 Canterbury Ct
Glen Ellyn, Illinois 60137
www.hubbardresearch.com

Selected Sources

- Tsai C., Klayman J., Hastie R. “Effects of amount of information on judgment accuracy and confidence” *Org. Behavior and Human Decision Processes,* Vol. 107, No. 2, 2008, pp 97-105.
- Heath C., Gonzalez R. “Interaction with Others Increases Decision Confidence but Not Decision Quality: Evidence against Information Collection Views of Interactive Decision Making” *Organizational Behavior and Human Decision Processes,* Vol. 61, No. 3, 1995, pp 305-326.
- Andreassen, P.” Judgmental extrapolation and market overreaction: On the use and disuse of news” *Journal of Behavioral Decision Making*, vol. 3 iss. 3, pp 153-174, Jul/Sep 1990.
- Williams M. Dennis A., Stam A., Aronson J. “The impact of DSS use and information load on errors and decision quality” *European Journal of Operational Research,* Vol. 176, No. 1, 2007, pp 468-81.
- Knutson et. al. “Nucleus accumbens activation mediates the influence of reward cues on financial risk taking” *NeuroRepor*t, 26 March 2008 - Volume 19 - Issue 5 - pp 509-513.
- A small study presented at Cognitive Neuroscience Society meeting in 2009 by a grad student at U. of Michigan showed that simply being briefly exposed to smiling faces makes people more risk tolerant in betting games.
- Risk preferences show a strong correlation to testosterone levels – which change daily (Sapienza, Zingales, Maestripieri, 2009).
- Recalling past events that involved fear and anger change the perception of risk (Lerner, Keltner, 2001).

The Method of Measurement

Measuring and Removing Inconsistency

Methods that statistically “smooth” estimates of experts show reduced error in several studies for many different kinds of problems.