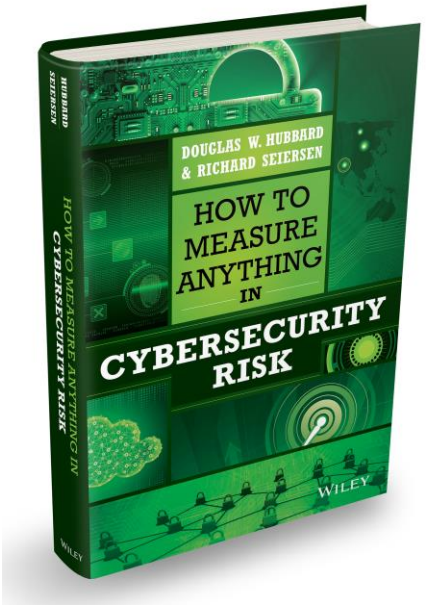
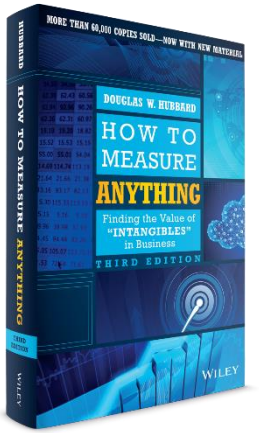
How to Measure Anything in Cybersecurity Risk

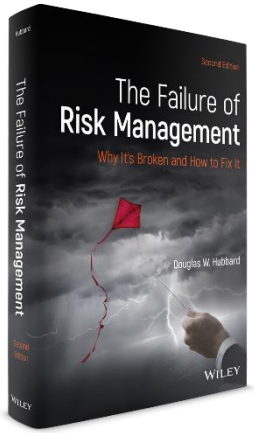
Hubbard Decision Research
2 South 410 Canterbury Ct
Glen Ellyn, Illinois 60137
www.hubbardresearch.com

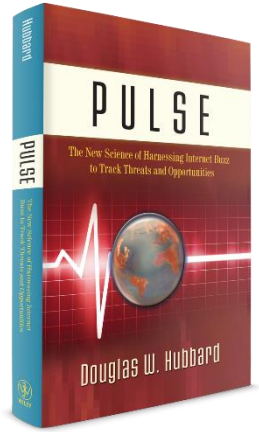


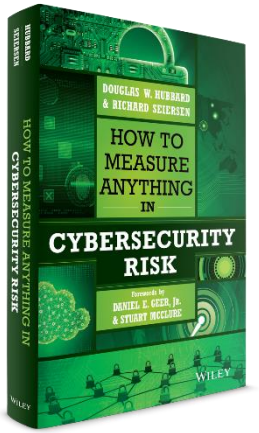
Introduction

My Books



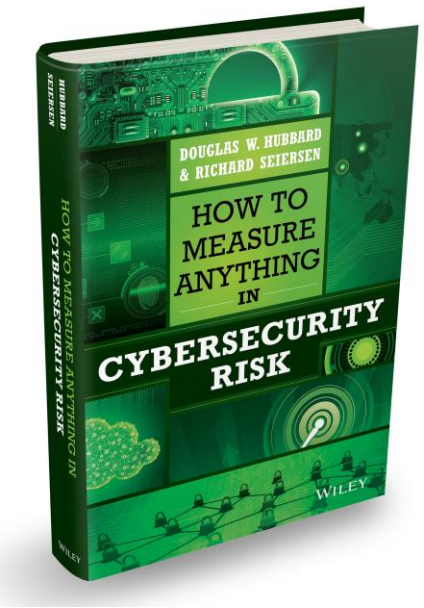






Introduction

How to Measure Anything in Cybersecurity Risk



" For thorough and practical guidance on using probability analysis for cybersecurity decision making, consult the book, How to Measure Anything in Cybersecurity "

Cite: CIS RAM Version 1.0 Center for Internet Security, Risk Assessment Method For Reasonable Implementation and Evaluation of CIS Controls

Introduction

Applied Information Economics

AIE was applied initially to IT business cases. But over the last 20 years it has also been applied to other decision analysis problems in all areas of Business Cases, Performance Metrics, Risk Analysis, and Portfolio Prioritization.

IT

- Prioritizing IT portfolios
- Risk of software development
- Value of better information
- Value of better security
- Risk of obsolescence and optimal technology upgrades
- Value of infrastructure
- Performance metrics for the business value of applications

Business

- Movie / film project selection
- New product development
- Pharmaceuticals
- Medical devices
- Publishing
- Real estate

Engineering

- Risks of major engineering projects
- Risk of mine flooding

Government & Non Profit

- Environmental policy
- Sustainable agriculture
- Procurement methods
- Grants management

Military

- Forecasting battlefield fuel consumption
- Effectiveness of combat training to reduce roadside bomb / IED casualties
- R&D portfolios

Introduction

The Biggest Cybersecurity Risk

Question: What is your single biggest risk in cybersecurity?

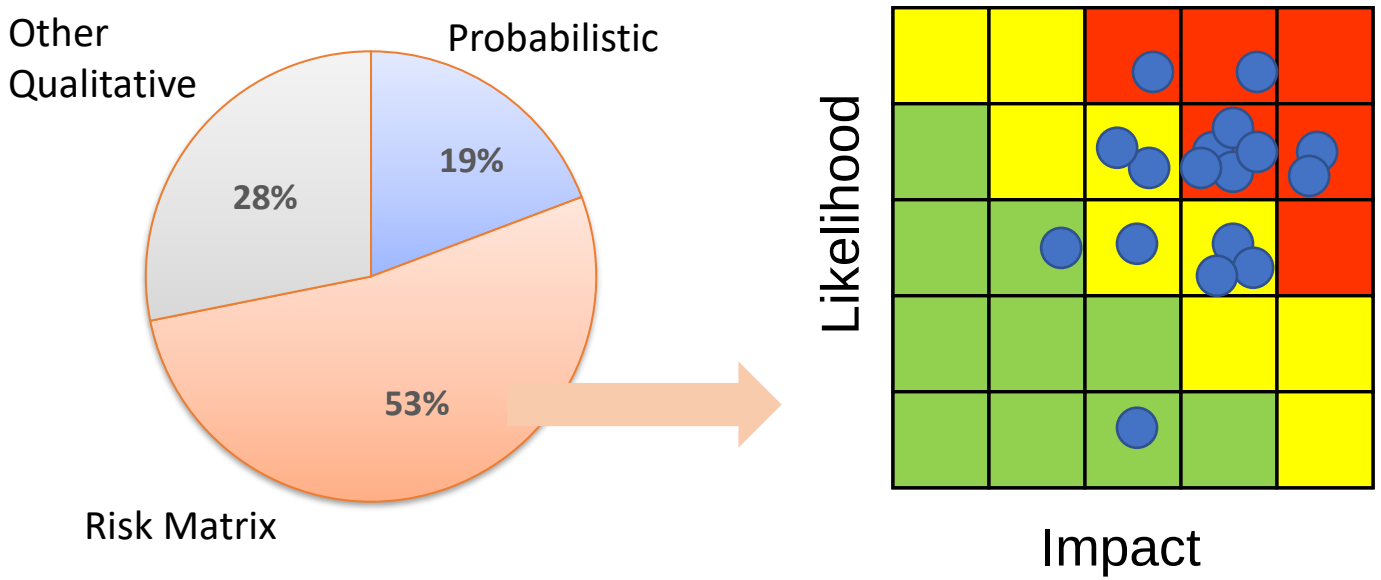
Answer: How you measure cybersecurity risk.

(This also applies to risk in general.)

Do “Scores” and “Scales” Work?

The Current Most Popular Method

Share of Methods Used in Cybersecurity Risk Assessment

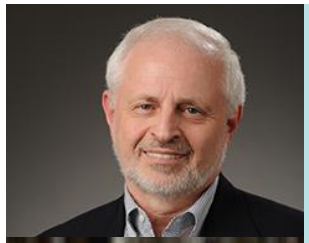


Source: HDR 2015 Survey of Cybersecurity Risk Methods (173 Responses)

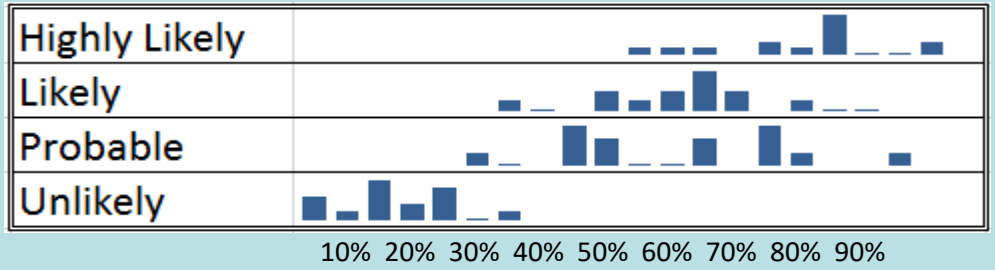
Do “Scores” and “Scales” Work?

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



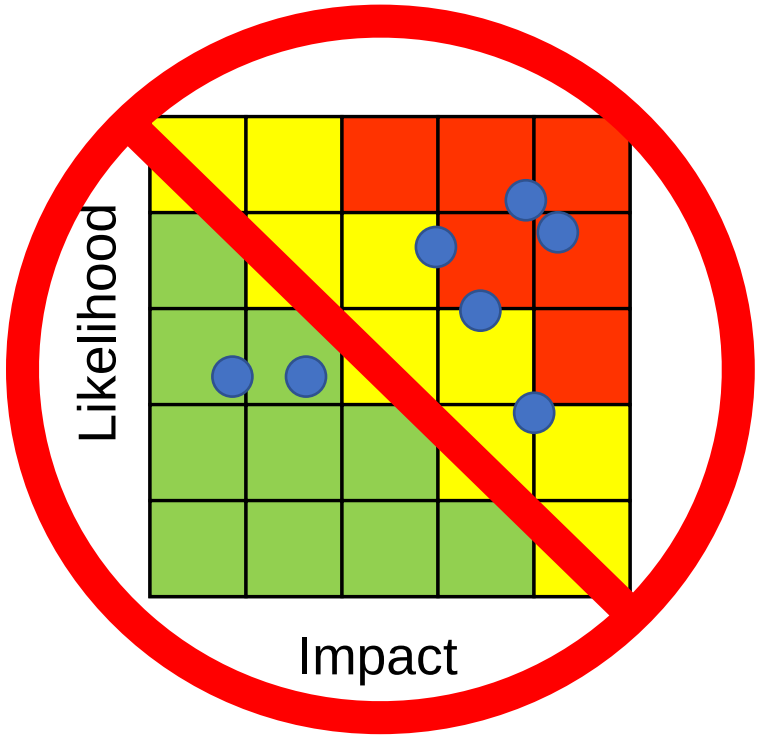
Craig R. Fox showed how arbitrary features of how scales are partitioned effects responses.

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” Work?

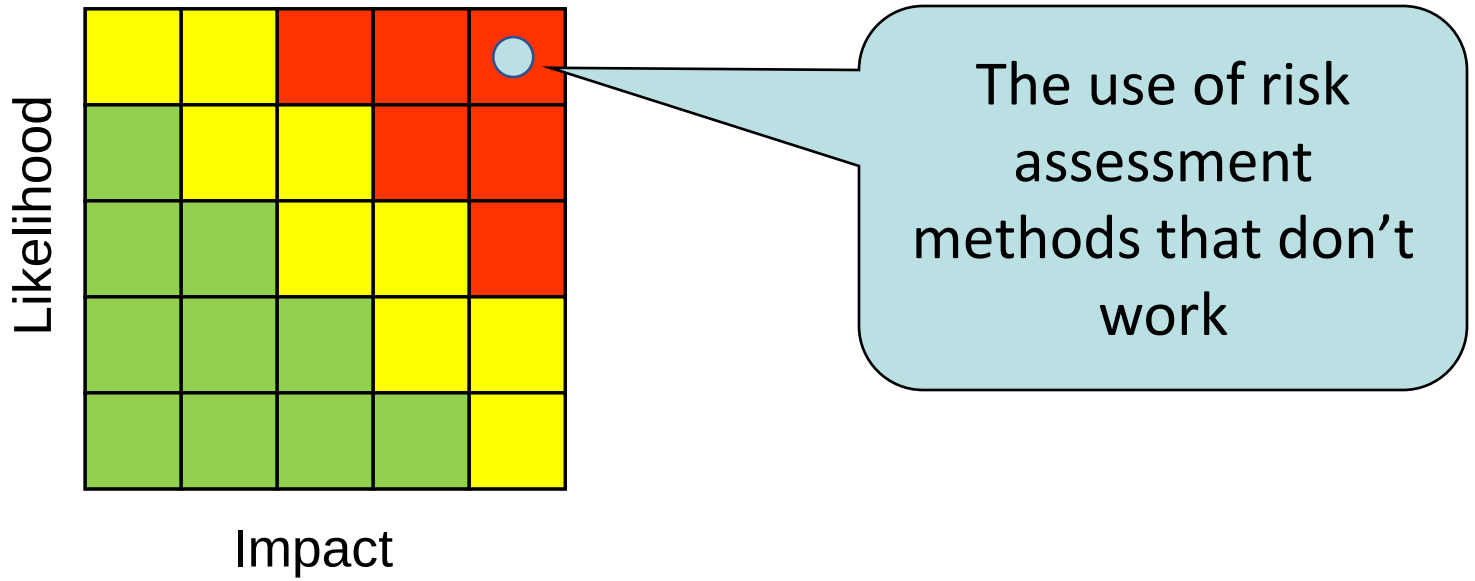
Summarizing Research on Risk Matrices



- “Risk Matrices should not be used for decisions of any consequence.”
  - Bickel et al. “The Risk of Using Risk Matrices”, *Society of Petroleum Engineers, 2014*
- “…they can be ‘worse than useless’”
  - Tony Cox “What’s wrong with Risk Matrices” investigates various mathematical consequences of ordinal scales on a matrix.

Do “Scores” and “Scales” Work?

The Only Risk Matrix You Need



The Analysis Placebo

Confidence in decision making methods is detached from performance

Organizational Behavior and Human Decision Processes
107, no. 2 (2008): 97– 105.

Journal of Behavioral Decision Making 3, no. 3 (July/ September 1990):
153– 174.

Law and Human Behavior 23 (1999): 499– 516.

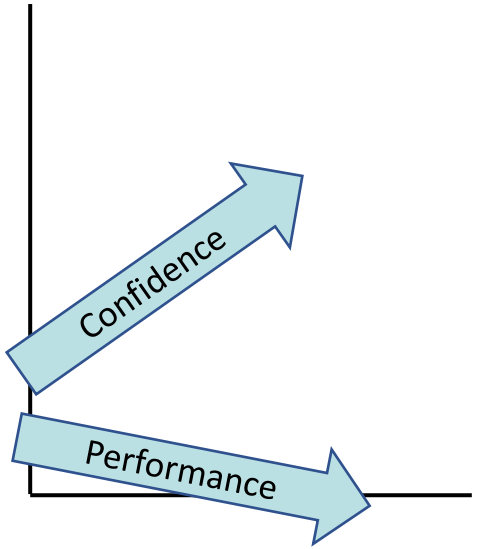
Organizational Behavior and Human Decision Processes 61, no. 3 (1995):
305– 326.

Interaction with Others Increases Decision Confidence but Not Decision Quality: Evidence against Information Collection Views of Interactive Decision Making

Heath and Gonzalez

Abstract

We present three studies of *interactive decision making,* where decision makers interact with others before making a final decision alone. Because the theories of lay observers and social psychologists emphasize the role of information collection in interaction, we developed a series



The Meta Decision

The Primary Strategy: Go Meta

> "Intelligence analysts should be self-conscious about their reasoning processes. They should think about how they make judgments and reach conclusions, not just about the judgments and conclusions themselves."
>
> Dick Heuer, *The Psychology of Intelligence Analysis*

The Meta-Analysis/Meta-Decision Criterion

- Is there clinical trial evidence that all or part of the method improves estimates and decisions – or makes them worse?
- Does it answer the right question?
- Is it practical?

The Meta Decision

How to Build a Method That Works

- Start with components that work.
- Don’t rely on anecdotes, testimonials or claims of “best practices” as evidence of working.
- If you can’t answer “What is the probability of losing more than X in the next 12 months due to event Y?” then you aren’t doing risk analysis.

A Cybersecurity Survey

2015 Survey: Interesting Connection

Those who said they could "compute the probability of various levels of losses" had about *half the rate of data breaches* as those who could not.

Does your organization compute the probability of various levels of losses?	Average Annual Data Breach Rate
Yes	4.5%
No	9%

173 responses total

A single survey might still be inconclusive – but it is consistent with other research about the improvement from using quantitative methods.

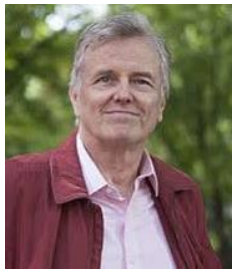
Experts vs. Algorithms

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

To many experts, when assessing probabilities many events ". . .are perceived as so unique that past history does not seem relevant to the evaluation of their likelihood." Tversky, Kahneman, *Cognitive Psychology* (1973)

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 political experts in a 20 year study covering elections, policy effects, wars, the economy and more.

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

What Measuring Risk Looks Like

Is Risk Analysis Actually Supporting Decisions?

- If risks and mitigation strategies were quantified in a meaningful way, decisions could be supported.
- In order to compute an ROI on mitigation decisions, we need to quantify likelihood, monetary impact, cost, and effectiveness.

	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**

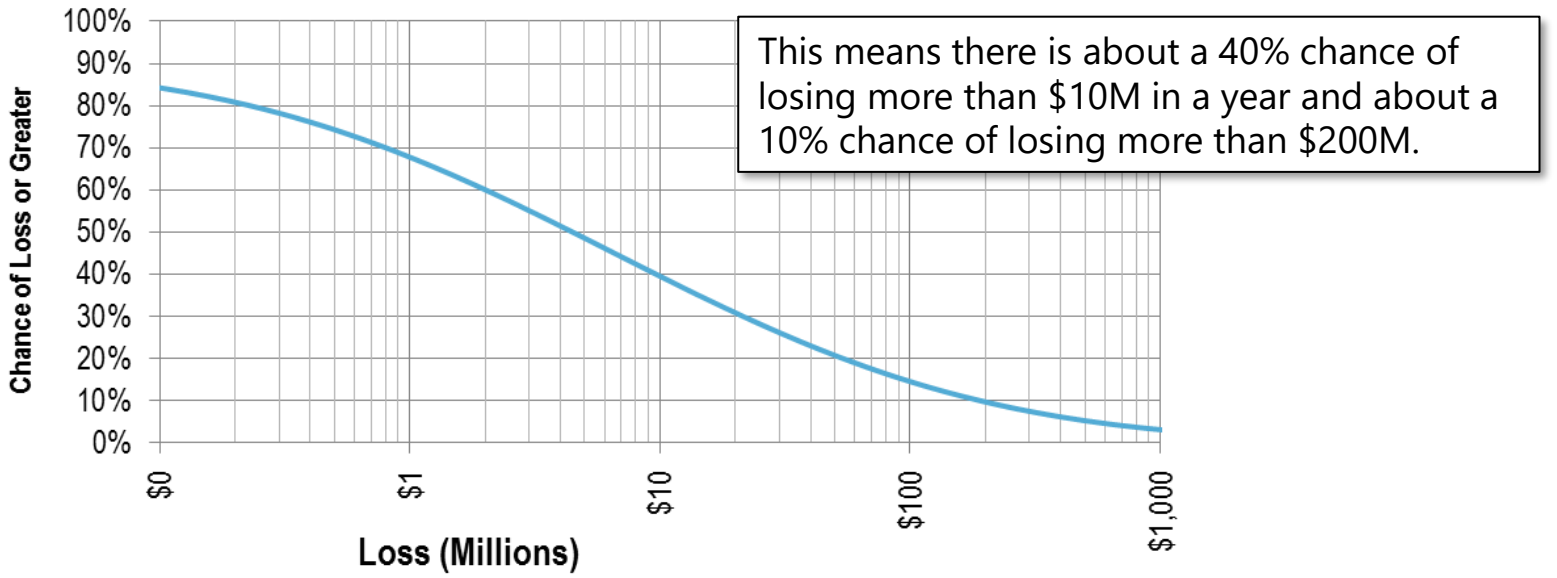
What Measuring Risk Looks Like

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.

What Measuring Risk Looks Like

A Simple “One-For-One Substitution”

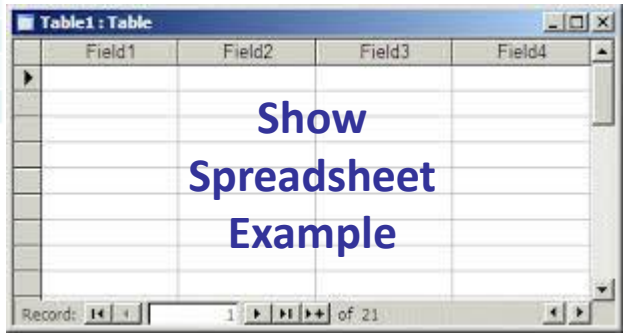
Each of these examples can be found on

www.howtomeasureanything.com/cybersecurity

Event	Event Probability (per Year)	Impact (90% Confidence Interval)		Random Result (zero when the event did not occur)
		Lower Bound	Upper Bound	
AA	.1	$50,000	$500,000	0
AB	.05	$100,000	$10,000,000	$8,456,193
AC	.01	$200,000	$25,000,000	0
AD	.03	$100,000	$15,000,000	0
AE	.05	$250,000	$30,000,000	0
AF	.1	$200,000	$2,000,000	0
AG	.07	$1,000,000	$10,000,000	$2,110,284
AH	.02	$100,000	$15,000,000	0
ZM	.05	$250,000	$30,000,000	0
ZN	.01	$1,500,000	$40,000,000	0
			Total:	$23,345,193

Each “Dot” on a risk matrix can be better represented as a row on a table like this

The output can then be represented as a Loss Exceedance Curve.



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"

"Cybersecurity is too complex to model."

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

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

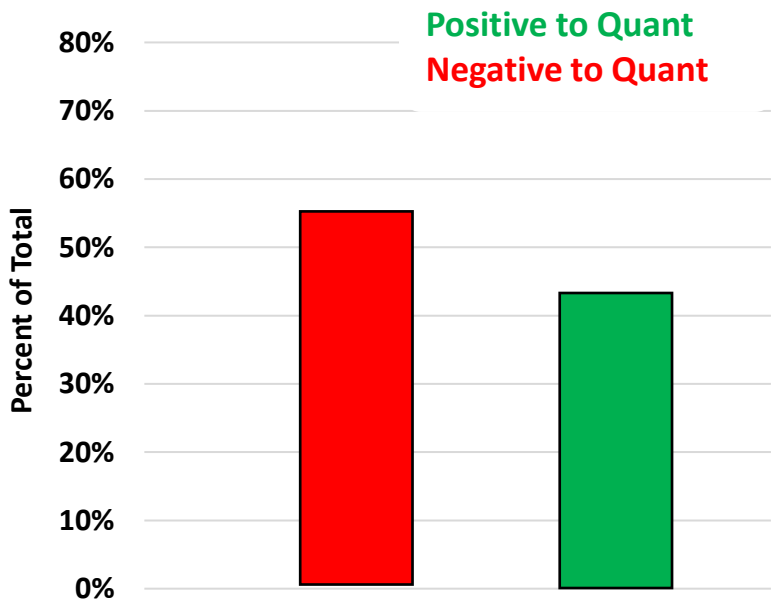
The implied (and unjustified) conclusion from each of these is….

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

So Why Don't We Use More Quantitative Methods?

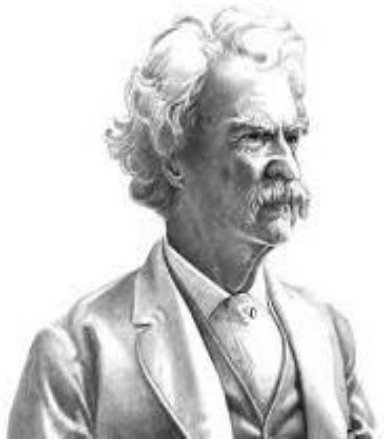
The Main Obstacle to Quantitative Methods

Another finding in the same survey: Strong opinions against "quant" are associated with poor stats understanding.



"It's not what you don't know that will hurt you, it's what you know that ain't so."

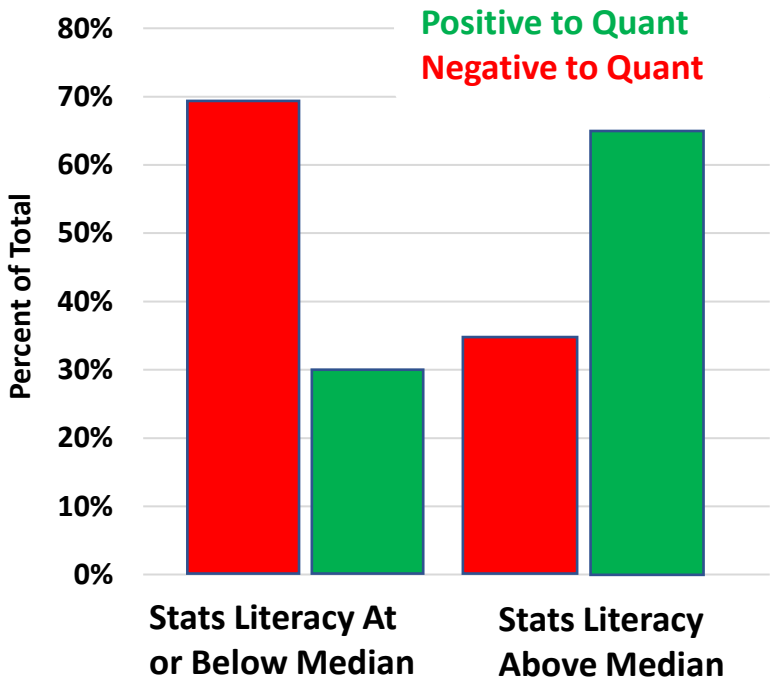
Mark Twain



So Why Don't We Use More Quantitative Methods?

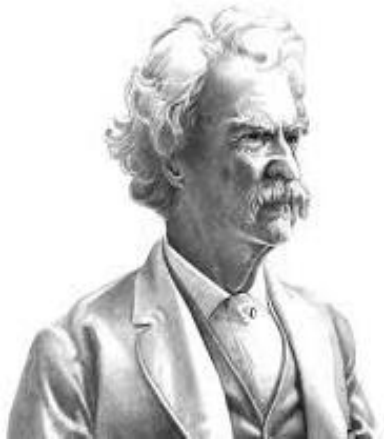
The Main Obstacle to Quantitative Methods

Another finding in the same survey: Strong opinions against "quant" are associated with poor stats understanding.



"It's not what you don't know that will hurt you, it's what you know that ain't so."

Mark Twain

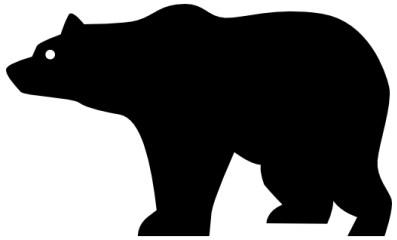


Algorithm Aversion

A Major Fallacy Regarding Comparing Methods

- Don't make the classic "Beat the Bear" fallacy.

Exsupero Ursus



A common form of the ***Exsupero Ursus*** fallacy:
"The quantitative model must have

1) All the variables
2) All the data
3) All the right distributions and correlations
4) All the above

If not, default to a measurably worse method.

The Three Misconceptions Behind Any Perceived "Immeasurable"

The Illusions of Immeasurability

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 Three Misconceptions Behind Any Perceived “Immeasurable”

The Concept of Measurement

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

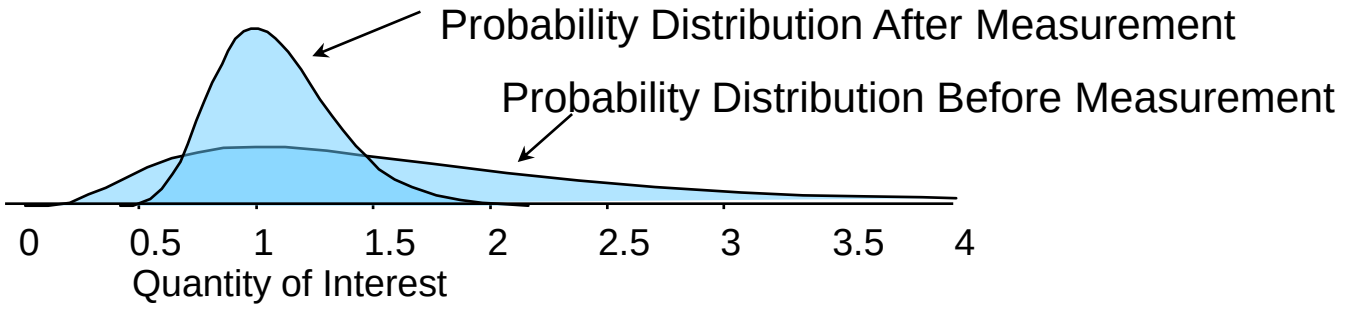
Many procedures of empirical observation are misunderstood.

The Concept of Measurement

It’s not a point value

It’s not a point value.

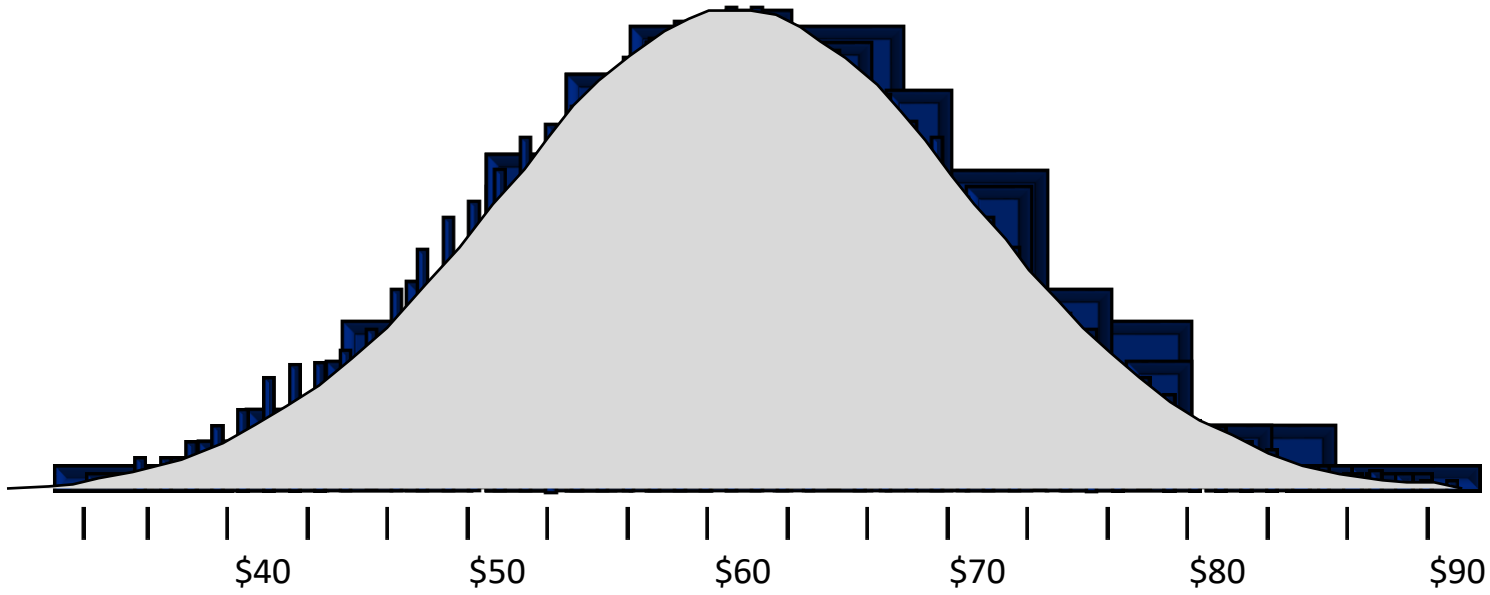
- Measurement: a quantitatively expressed reduction in uncertainty based on observation.
- You can quantify your current uncertainty – additional observations reduce it.
- Even marginal reductions in uncertainty can be extremely valuable.



The Concept of Measurement

Constructing a Distribution

- Uncertainty about “either/or” events are expressed as “discrete” probabilities (e.g. “35%).
- Uncertainty about continuous values can still be thought of as sets of discrete probabilities.



The Three Misconceptions Behind Any Perceived “Immeasurable”

The Object of Measurement

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

The Importance of Defining a Measurement

- If a thing seems like an immeasurable “intangible” it may just be ill-defined.
- Often, if we can define what we mean by a certain “intangible” we find ways to measure it.
- Examples: Brand image, Security, Safety, etc.

The Object of Measurement

Clarifying the Problem

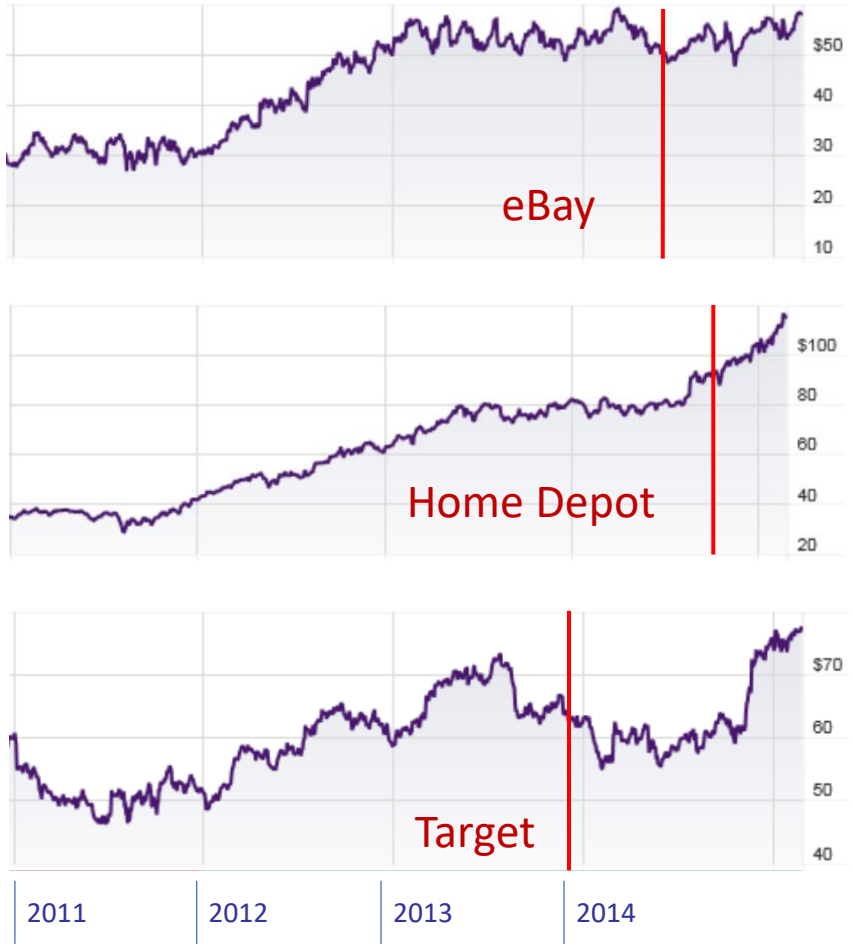
1. Why do you care? (What decision could depend on the outcome of this measurement?)
2. What do you see when you see more of it? (Describe it in terms of observable consequences, then units of measure.)
3. How much do you know about it now?
4. At what point will the value make a difference?
5. How much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

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

The Object of Measurement

Measurement Challenge: Reputation Damage

- One of the perceived most difficult measurements in cybersecurity is damage to reputation.
- Trick: *There is no such thing as a "secret" damage to reputation!*
- How about comparing stock prices after incidents? (That's all public!)
- So what is the *REAL* damage?
  - Legal liabilities,
  - Customer outreach
  - "Penance" projects (security overkill)
- The upshot, damage to reputation actually has available information and easily observable measured costs incurred to *avoid* the bigger damages!



The Three Misconceptions Behind Any Perceived “Immeasurable”

The Method of Measurement

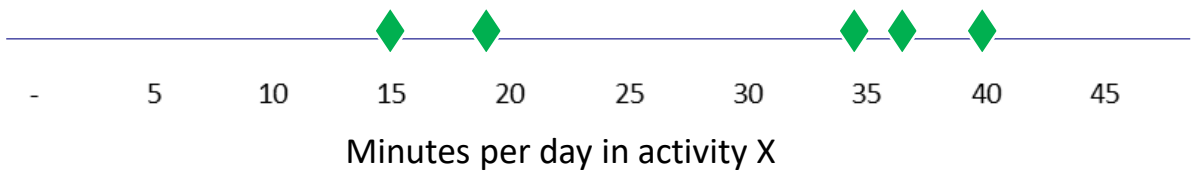
CONCEPT of Measurement	The definition of measurement itself is widely misunderstood.
OBJECT of Measurement	The thing being measured is not well defined.
METHOD of Measurement	Many procedures of empirical observation are misunderstood.

The Method of Measurement

Testing Measurement Intuition

A Sample of 5

- Suppose you are extremely uncertain about how much time per day is spent in some activity in a company of 10,000 people.
- Imagine you randomly sample 5 people out of a company and they spend an amount of time in this activity as shown by the data points below.
- Is this statistically significant?
- Is it possible to estimate the chance the median time spent per person per day is between 15 and 40 minutes?



Minutes per day in activity X

The Method of Measurement

The Urn of Mystery

THE *URN OF MYSTERY* PROBLEM

- There is a warehouse full of urns.
- Each urn is filled with over a million marbles, each of which are red or green.
- The proportion of red marbles in each urn is unknown – it could be anything between 0% and 100% and all possibilities are equally likely.

Questions:
If you randomly select a single marble from a randomly selected urn, what is the chance it is red?
If the marble you draw is red, what is the chance the majority of marbles are red?
If you draw 8 marbles and all are green, what is the chance that the next one you draw will be red?

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

Useful Assumptions About Measurement

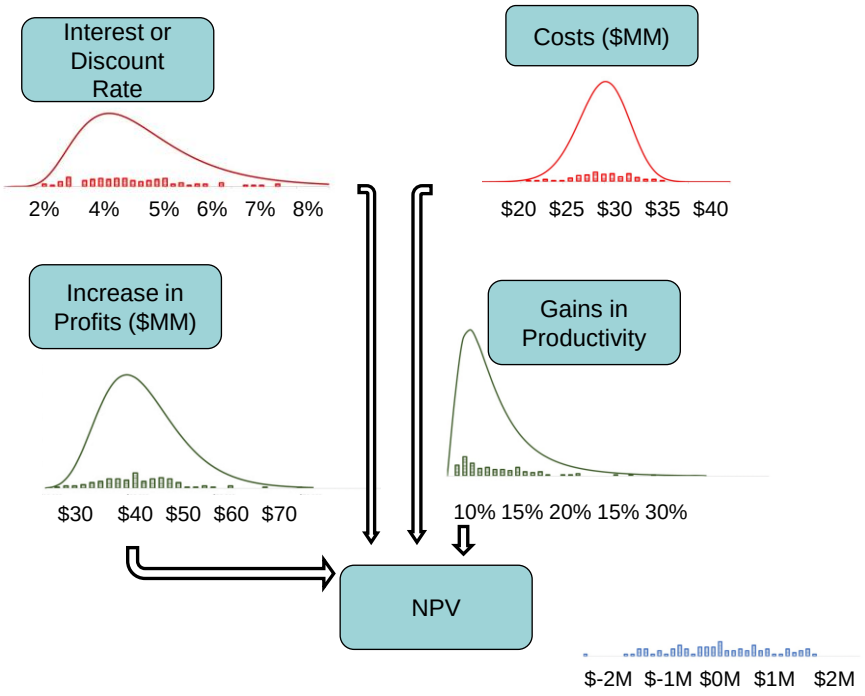
If your measurement is challenged with limited or messy data, consider the following:

- It's been measured before.
- You have more data than you think.
- You need less data than you think.

"It's amazing what you can see when you look"
Yogi Berra

The Method of Measurement

Monte Carlo: How to Model Uncertainty in Decisions



What Published Research Says
(See sources slide for details)

- Psychologists showed that simple decomposition greatly reduces estimation error for estimating the most uncertain variables.
- In the oil industry there is a correlation between the use of quantitative risk analysis methods and financial performance.
- Data at NASA from over 100 space missions showed that Monte Carlo simulations and historical data beat softer methods for estimating cost and schedule risks.

The Method of Measurement

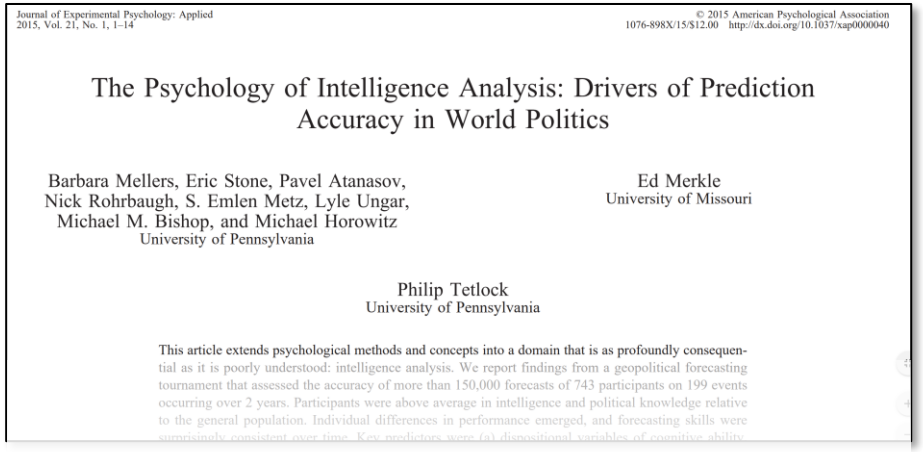
Improving Expert Judgement

- Calibration of experts for overconfidence and inconsistency is a start.
- Decomposition tends to further improve expert estimates.
- We can leverage these facts for making improved models even without other recorded, empirical data (adding that comes next).

The Method of Measurement

Improving Expert Forecasts

- Tetlock also looked at what improved *forecasting.*
- He tracked 743 individuals who made at least 30 forecasts each over a 2-year period.
- He determined factors that made the biggest difference in the performance of forecasting.



Journal of Experimental Psychology: Applied
2015, Vol. 21, No. 1, 1–14

1076-898X/15/$12.00 http://dx.doi.org/10.1037/xap0000040

The Psychology of Intelligence Analysis: Drivers of Prediction Accuracy in World Politics

Barbara Mellers, Eric Stone, Pavel Atanasov, Nick Rohrbaugh, S. Emlen Metz, Lyle Ungar, Michael M. Bishop, and Michael Horowitz
University of Pennsylvania

Ed Merkle
University of Missouri

Philip Tetlock
University of Pennsylvania

This article extends psychological methods and concepts into a domain that is as profoundly consequential as it is poorly understood: intelligence analysis. We report findings from a geopolitical forecasting tournament that assessed the accuracy of more than 150,000 forecasts of 743 participants on 199 events occurring over 2 years. Participants were above average in intelligence and political knowledge relative to the general population. Individual differences in performance emerged, and forecasting skills were

Probabilistic Training

- Subjects were trained in basic inference methods, using reference classes, and avoiding common errors and biases.

Teams and Belief Updating

- Teams deliberated more and individuals were willing to update beliefs based on new information.

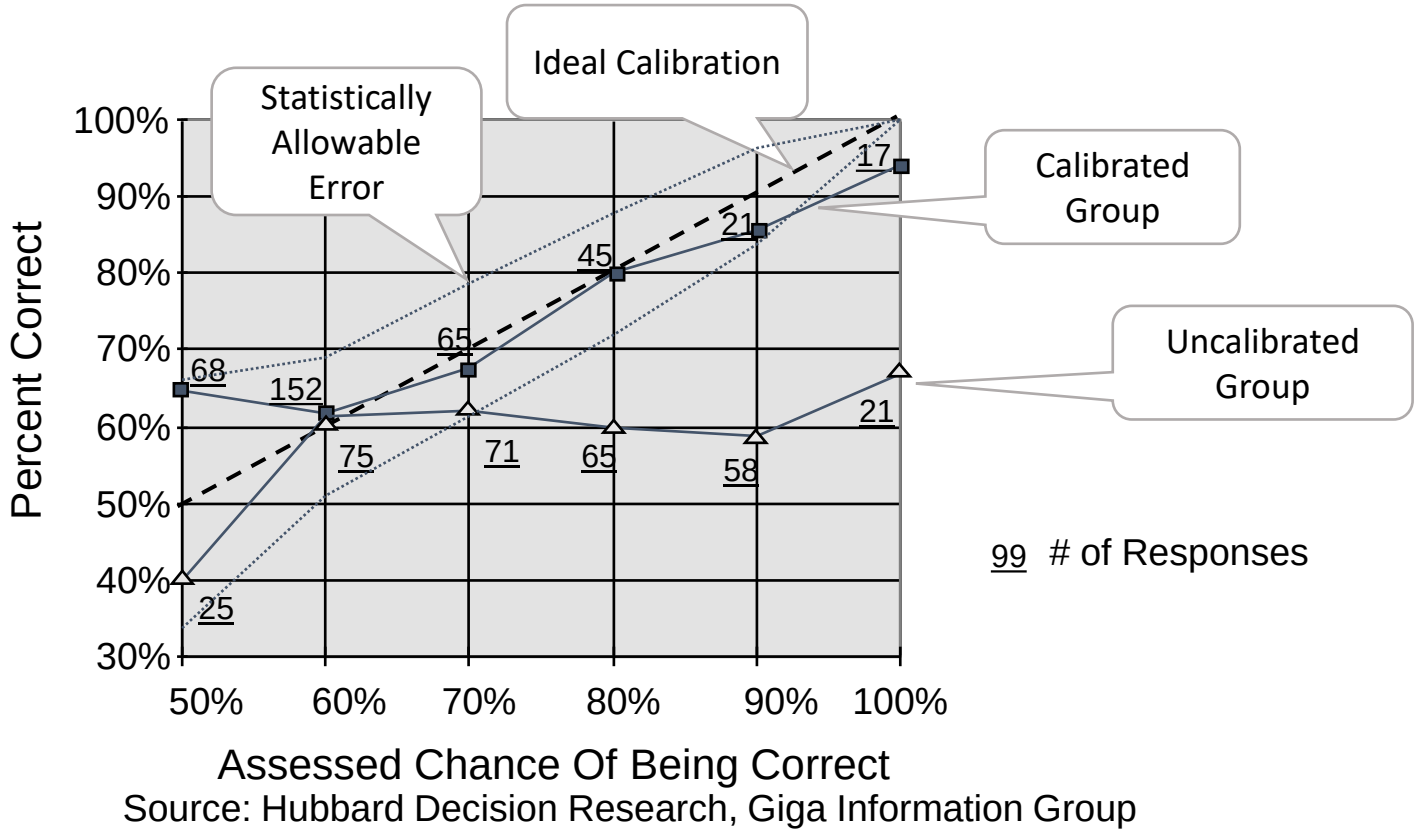
Selecting the Best

- Brains matter. Both topic expertise and overall IQ were the best predictors of performance.

The Method of Measurement

Training Experts to Give Calibrated Probabilities

Training can “calibrate” people so that of all the times they say they are 90% confident, they will be right 90% of the time.



Source: Hubbard Decision Research, Giga Information Group

The Method of Measurement

Quantifying Risk Tolerance

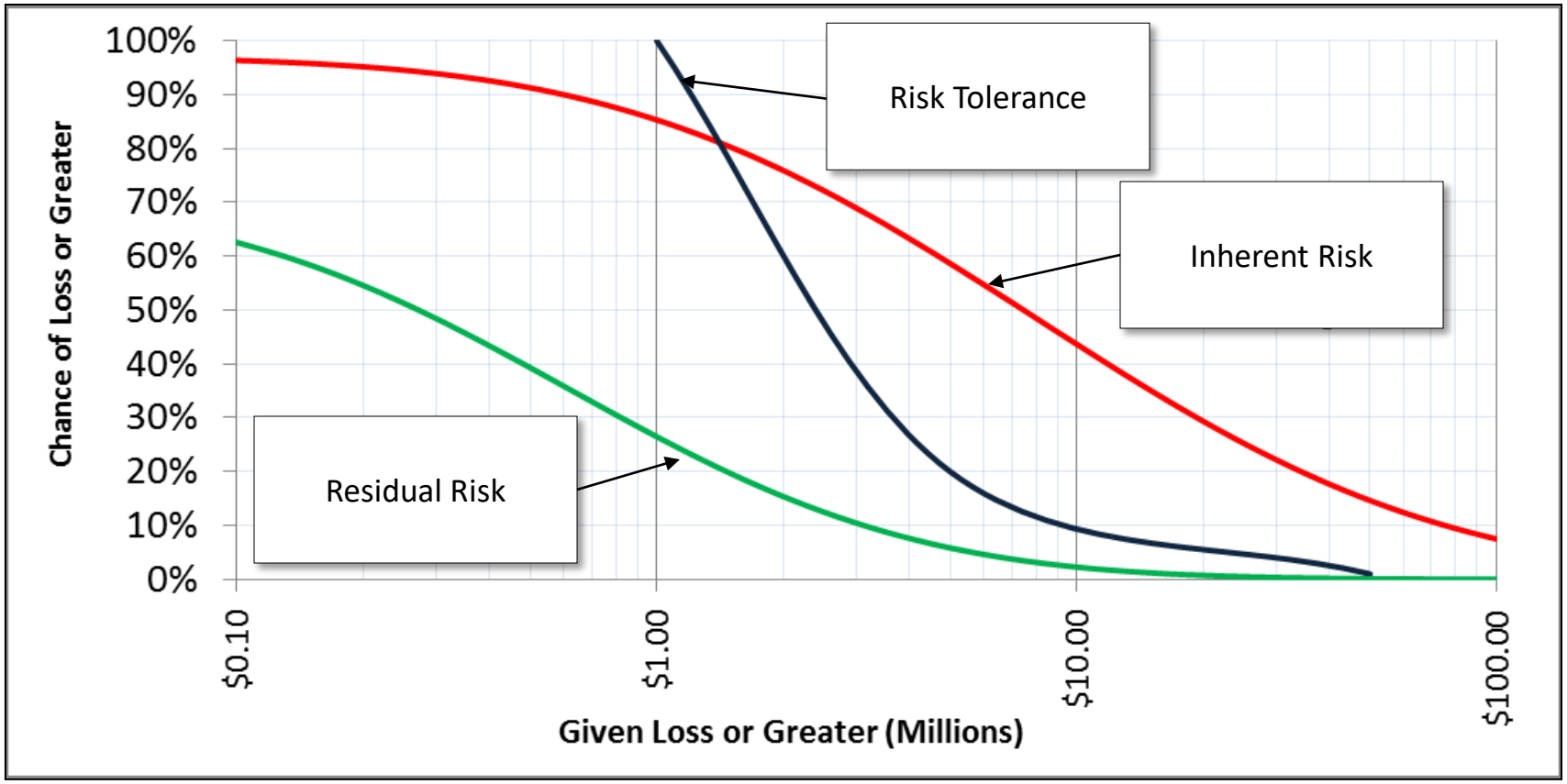
Studies have shown risk aversion changes due to what should be irrelevant external factors including:

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	↑

The Method of Measurement

Loss Exceedance Curves: Before and After

How do we show the risk exposure after applying available mitigations?



The Method of Measurement

Which Decomposition?

- Decomposition can improve models but not all decompositions are of equal value.
- Uninformative decompositions: Dwelling on speculations that you actually have no information about.
- Example: Assessing skill levels of unknown future attackers, speculating whether the risk is more the Russian mafia, Anonymous or China)

The Method of Measurement

Informative Decompositions

Informative decompositions use what you know or data you can get to improve estimates in models.

Informative Decompositions:

- **Systems**: You have fairly detailed knowledge of your applications, what data they have and the hardware it runs on. Some of the parameters of these systems would change your estimate of a risk.
- **Types of Impacts**: You separate confidentiality, integrity and availability events. You have an idea of business volumes like sales and other processes. If a breach or outage occurred, you can describe something about the consequences.
- **Staff**: You have knowledge of the number of employees, device loss rates, and some knowledge of what data they may have.
- **Vendors & Customers**: You know who the parties you interact with and you have some knowledge about them.
- **Insurance:** Any cyber-insurance will have detailed language regarding limitations, exclusions, etc.

The Method of Measurement

Bayesian Methods

- "Bayesian" methods in statistics use new information to update prior knowledge.

Bayes Theorem: $$P(X|Y) = \frac{P(X)P(Y|X)}{P(Y)} = \frac{P(X)P(Y|X)}{\sum_i P(Y|X_i)\,P(X_i)}$$

P(X) = the probability of X

P(X|Y) = the probability of X given the condition Y

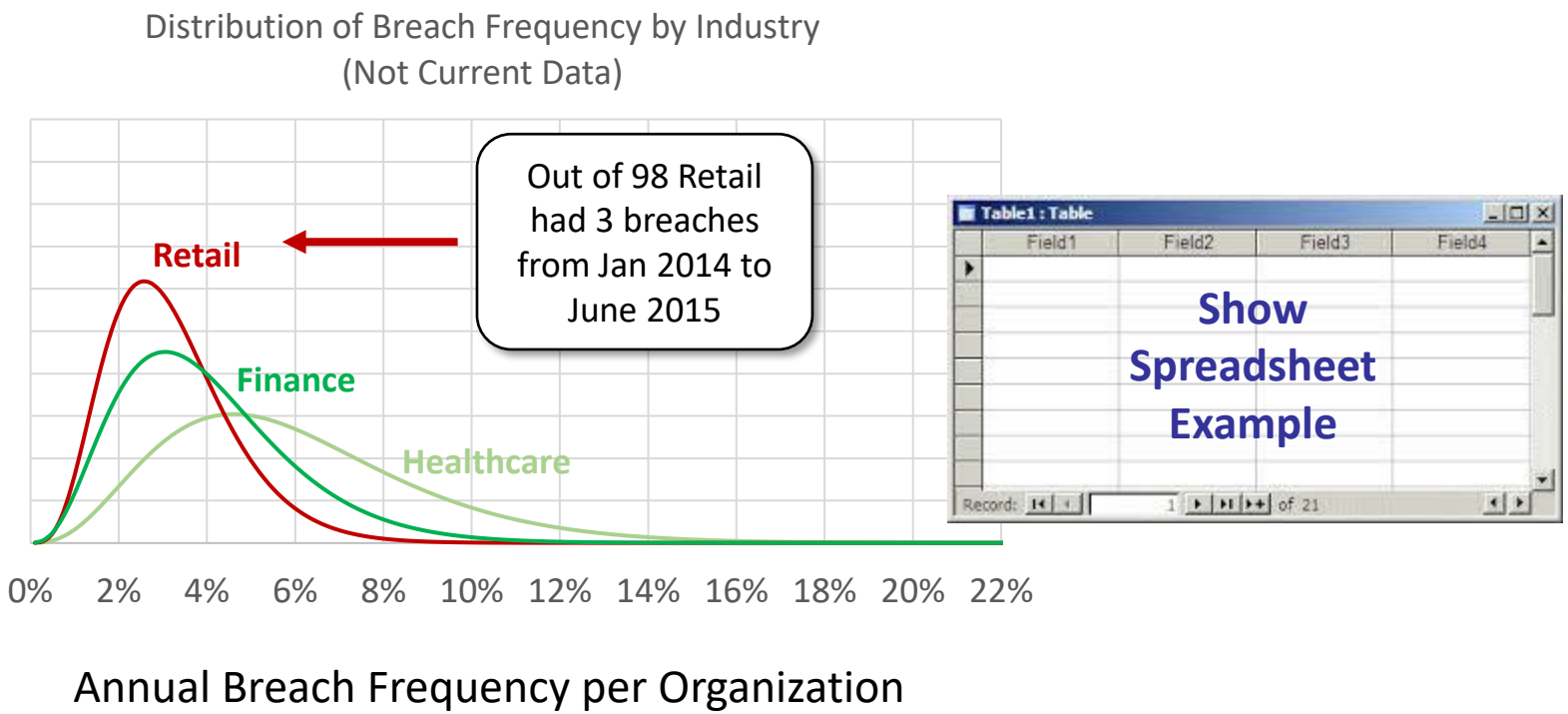
$\sum_i P(Y \mid X_i)\,P(X_i)$ = the sum of the probability of Y under each possible condition

- The Simplest Measurement Method — It turns out that calibrated people are already mostly "instinctively Bayesian".
  - Assess your initial subjective uncertainty with a calibrated probability
  - Gather and study new information
  - Give another subjective calibrated probability assessment

The Method of Measurement

Estimating Breach Rate w/History

- You have relatively few examples of major, reported breaches in each industry.
- There is a statistical method for estimating the frequency of breaches based on small samples.
- Spreadsheet for this at www.howtomeasureanything.com/cybersecurity



The Method of Measurement

The Rule of Succession

Danny Kahneman

A reference class is a population from which you draw observations of events to determine their frequency. Your "reference class" is much larger than you.

You can start by making as few assumptions as possible – your "baseline" uses only your reference class

Pierre-Simon Laplace
1749-1827

- Laplace's "rule of succession": Given a population of reference class, like company-years, where some number of events occurred:
  - Chance of X (per year, per draw, etc.) =(1+hits)/(2+hits+misses)

Summary

Final Thoughts

It's Been Measured Before	• Important topics have often been measured already.
You Have More Data Than You Think	• Define a reference class – don't commit the reference class fallacy.
You Need Less Data Than You Think	• Question your intuition about how and whether messy and incomplete data is.

Example Spreadsheets for many of the calculations mentioned can be found at www.howtomeasureanything.com

Do’s and Don’ts

- Stop using risk matrices and “high, medium, low” as assessments of risk.

- Start using previously proven components:
  - probabilistic methods including Monte Carlo
  - calibrated experts
  - historical observations
  - quantified risk tolerance

Questions?

Contact:

Doug Hubbard

Hubbard Decision Research

dwhubbard@hubbardresearch.com

www.hubbardresearch.com

630 858 2788