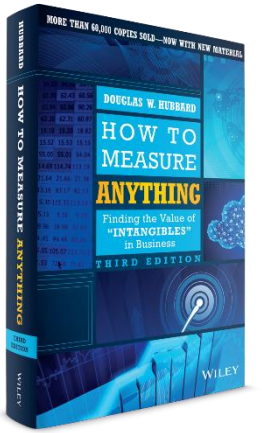
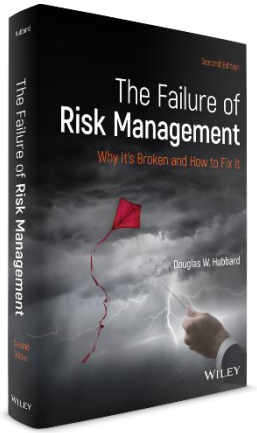
Risk Management
with
Applied Information Economics

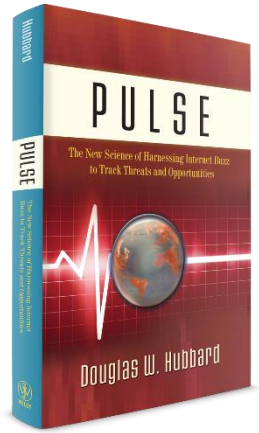
Hubbard Decision Research
2 South 410 Canterbury Ct
Glen Ellyn, Illinois 60137
www.hubbardresearch.com

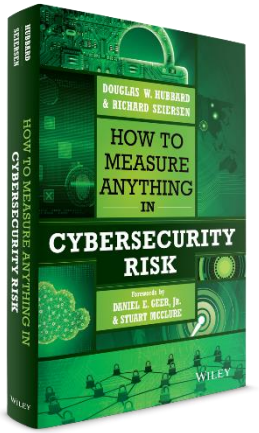
Introduction

My Books









Introduction

Applied Information Economics

Applied Information Economics (AIE)

Information Technology

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

Business Investments

- Prioritizing R&D in aerospace, biotech, pharma, medical devices and more
- Publishing
- Real estate
- Movie/film project selection

Engineering

- Power and road infrastructure upgrades
- Mining Risks

Government & Non-Profit

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

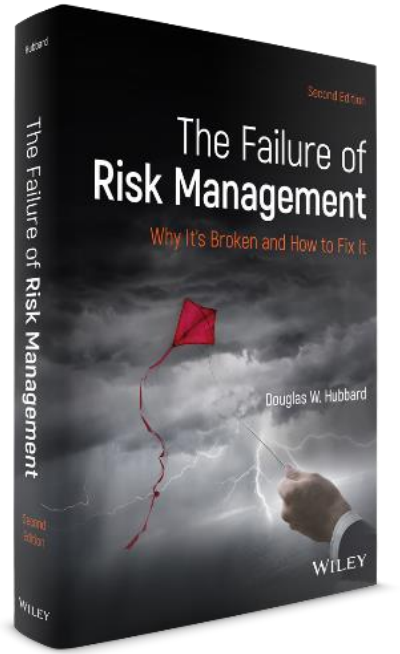
Military

- Forecasting battlefield fuel consumption
- Effectiveness of combat training to reduce roadside bomb/IED casualties
- Methods for testing equipment

Introduction

Topics for Today

- The Meta-Decision
- Getting Started
- Obstacles
- Simple Math



Introduction

The Biggest Risk

Question: What is your single biggest risk?

Answer: How you measure risk.

Top Five Risks?

- Large surveys by major consulting firms have produced very different rankings of top risks.
- These are self-reported and none of them ask exactly how they assess risks.
- The survey HDR conducted jointly with KPMG Netherlands examined how they assess risks among other topics.

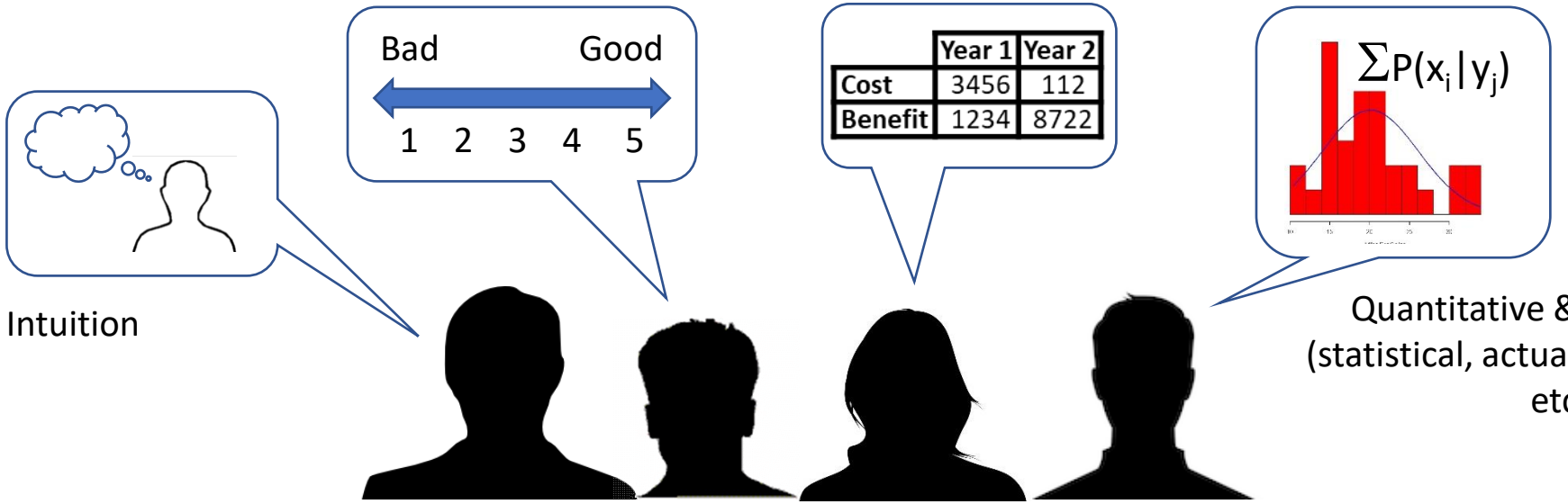
Protiviti	Aon	EIU
Disruptive technologies	Damage to reputation	Weak demand
Internal resistance to change	Economic slowdown	Market instability within own industry
Cyber threats	Increasing competition	Difficulty raising financing
Regulatory changes	Regulatory changes	Labor (skills shortage, strikes, etc.)
Timely identification and escalation of risks	Cyber threats	Exchange rate fluctuation

Introduction

Types of Measurement Methods

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

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

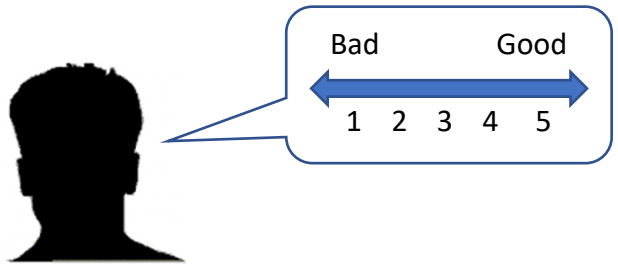


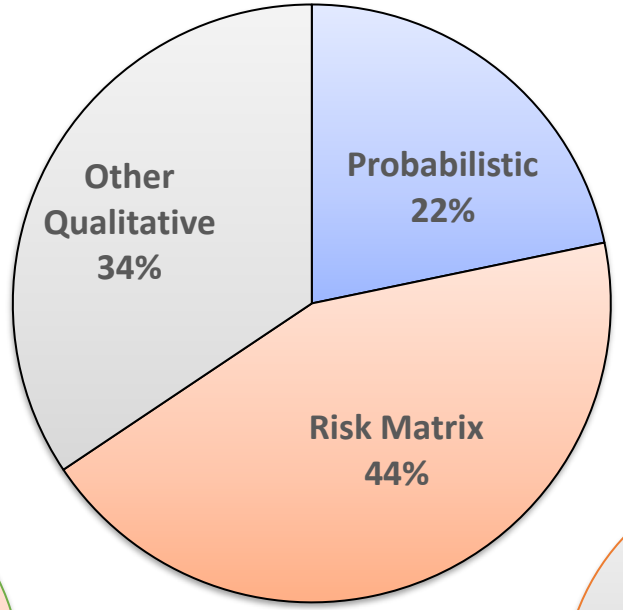
Expert Intuition

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

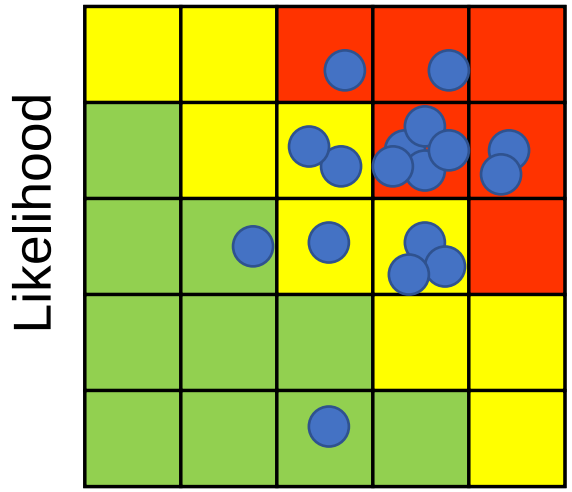
Do “Scores” and “Scales” Work?

The Current Most Popular Method

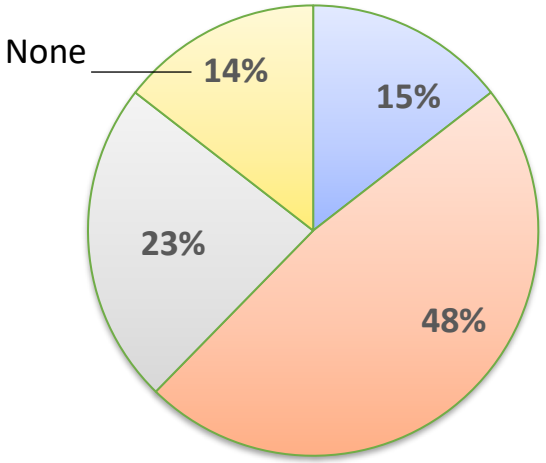


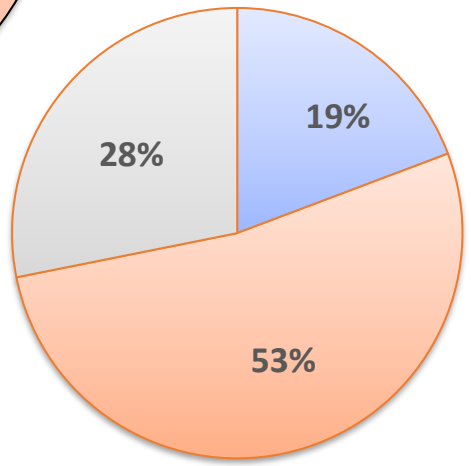


Enterprise Risk Management



Project Management

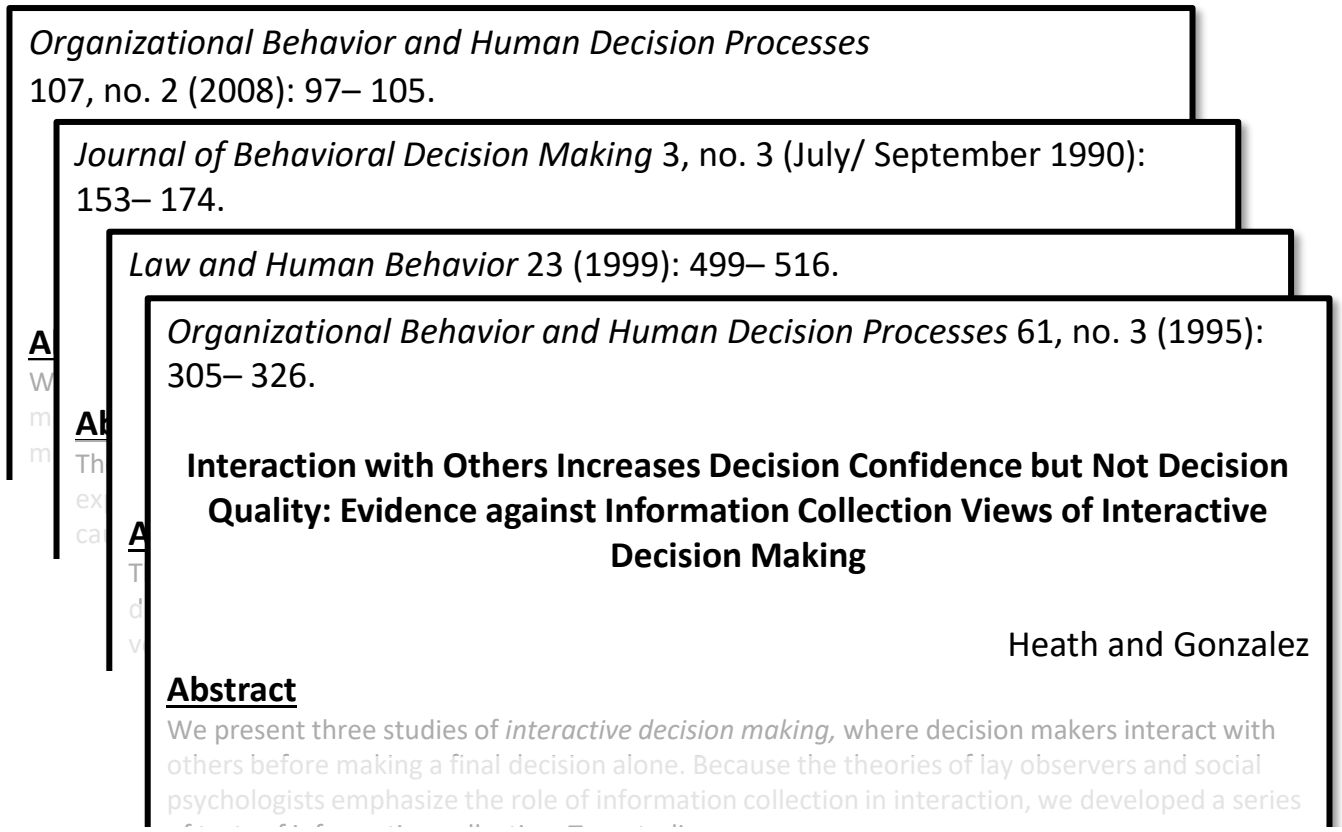


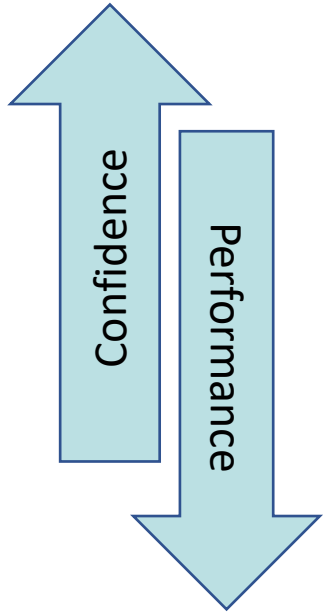


Cybersecurity

The Analysis Placebo

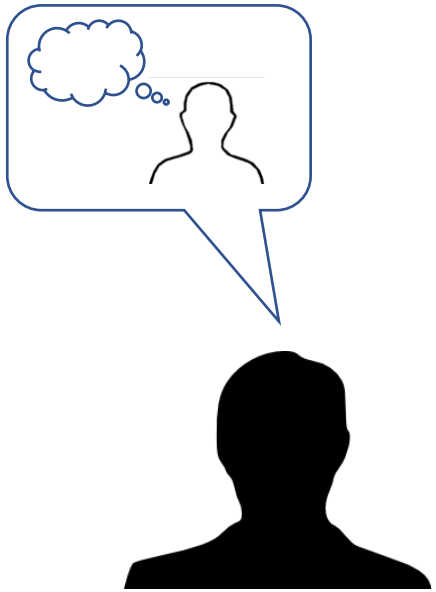
Confidence in decision making methods is detached from performance





Deciding How to Decide

- Why experience alone may not be enough to make the meta-decision



To learn from experience, you need feedback.

And that feedback has to be *CONSISTENT…*

…IMMEDIATE…

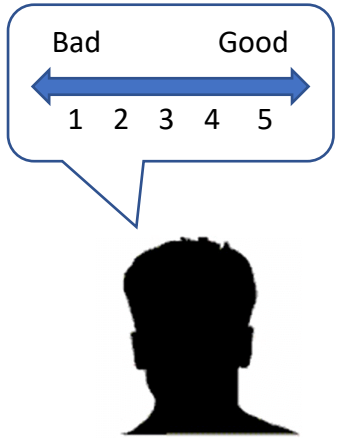
…and UNAMBIGUOUS.

Daniel Kahneman

Gary Klein

Do “Scores” and “Scales” Work?

Unintended consequences of simple scoring methods



Climatic Change (2012) 113:181–200
DOI 10.1007/s10584-011-0330-3

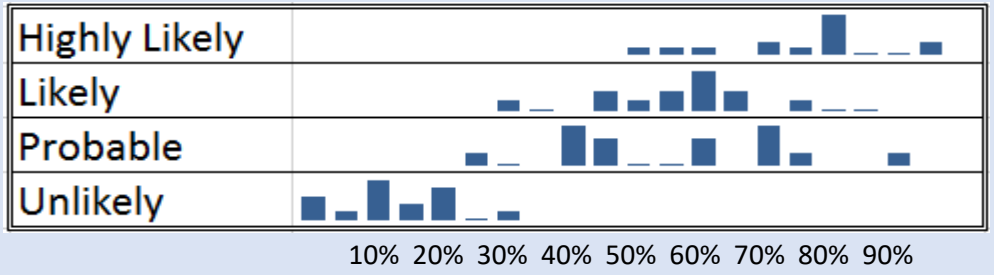
Effective communication of uncertainty in the IPCC reports

David V. Budescu · Han-Hui Por · Stephen B. Broomell

Received: 21 June 2010 / Accepted: 19 October 2011 / Published online: 23 November 2011

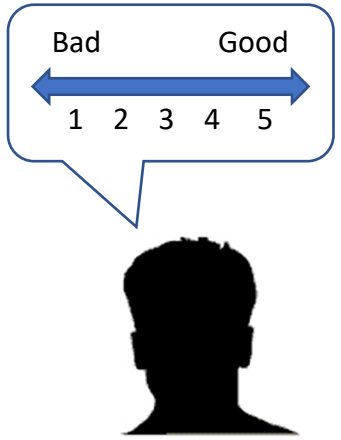
Abstract The Intergovernmental Panel on Climate Change (IPCC) publishes periodical

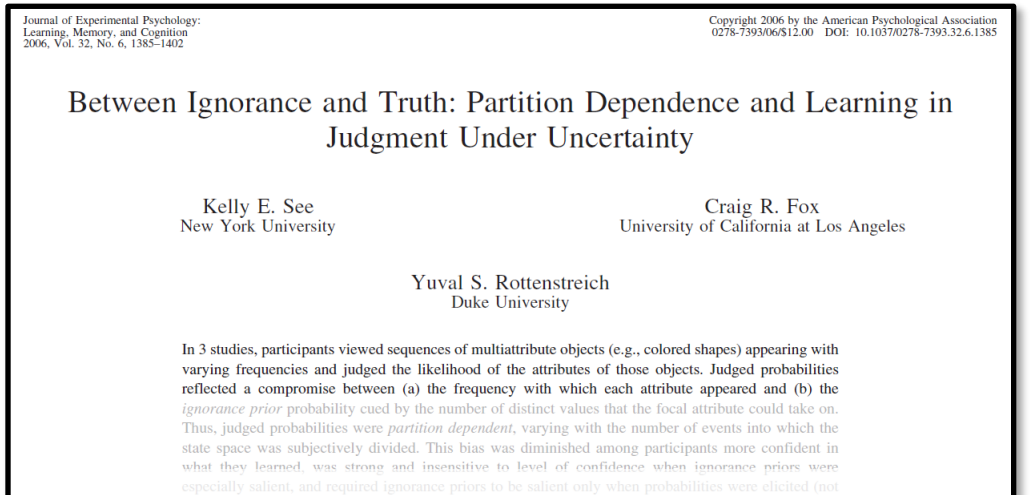
David Budescu and Dick Heuer (separately) researched the “illusion of communication” regarding interpretations of verbal labels for probabilities.



Do “Scores” and “Scales” Work?

Unintended consequences of simple scoring methods





Journal of Experimental Psychology:
Learning, Memory, and Cognition
2006, Vol. 32, No. 6, 1385–1402

0278-7393/06/$12.00 DOI: 10.1037/0278-7393.32.6.1385

Between Ignorance and Truth: Partition Dependence and Learning in Judgment Under Uncertainty

Kelly E. See
New York University

Craig R. Fox
University of California at Los Angeles

Yuval S. Rottenstreich
Duke University

In 3 studies, participants viewed sequences of multiattribute objects (e.g., colored shapes) appearing with varying frequencies and judged the likelihood of the attributes of those objects. Judged probabilities reflected a compromise between (a) the frequency with which each attribute appeared and (b) the *ignorance prior* probability cued by the number of distinct values that the focal attribute could take on. Thus, judged probabilities were *partition dependent*, varying with the number of events into which the state space was subjectively divided. This bias was diminished among participants more confident in what they learned, was strong and insensitive to level of confidence when ignorance priors were especially salient, and required ignorance priors to be salient only when probabilities were elicited (not

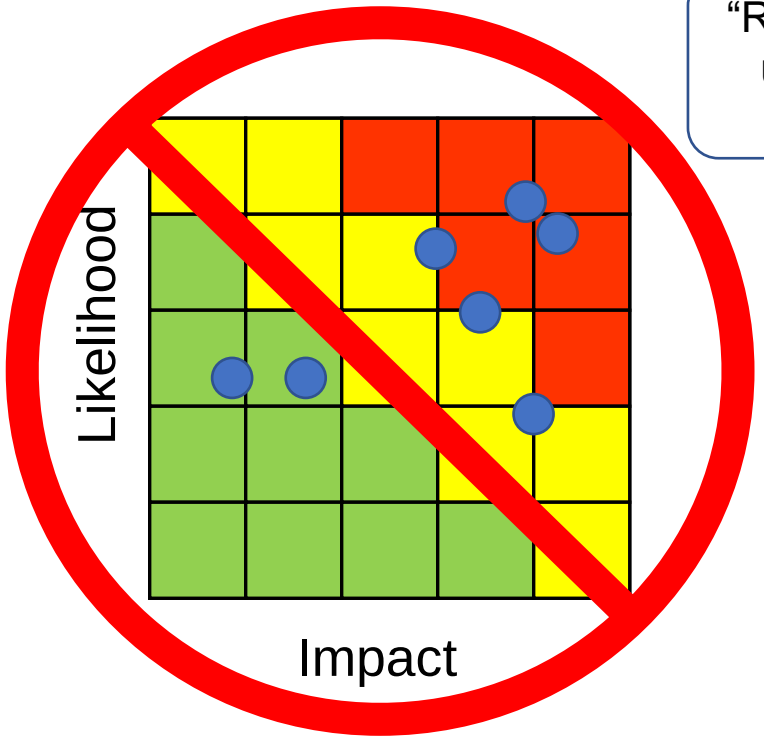
Craig R. Fox showed how arbitrary features of how scales are partitioned effects responses.

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?

The Ubiquitous Risk Matrix



“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?

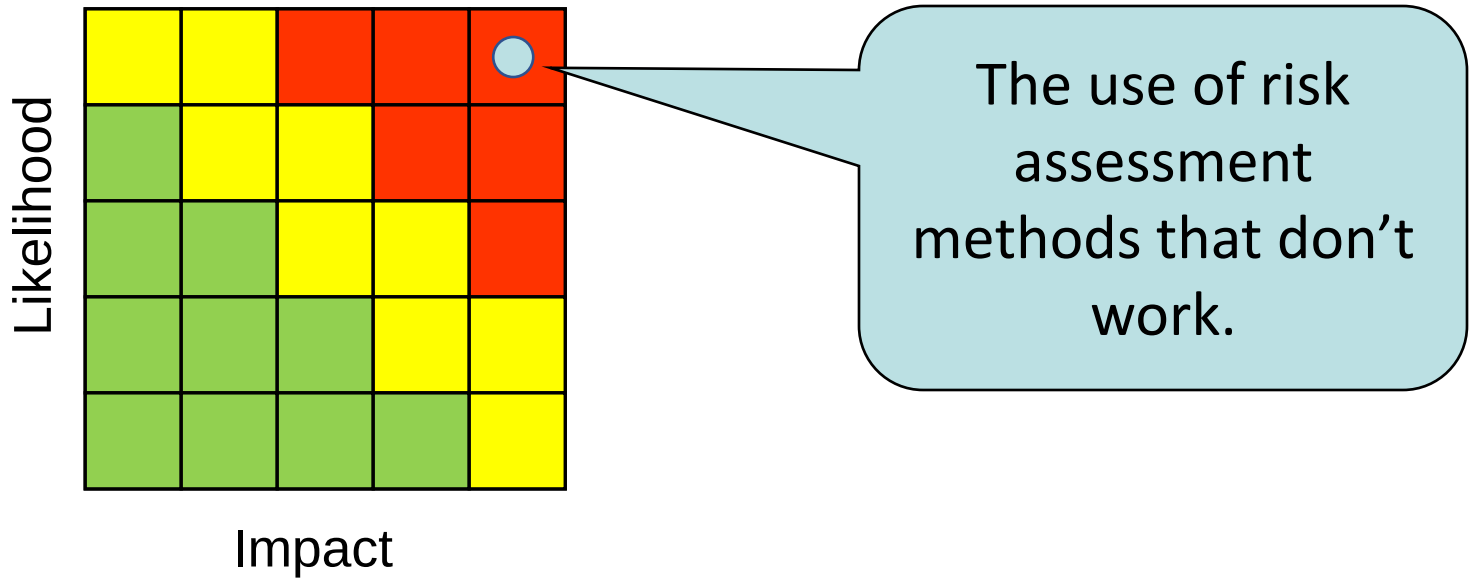
L. A. Cox, Jr.

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



The Meta Decision

How to Build a Method That Works

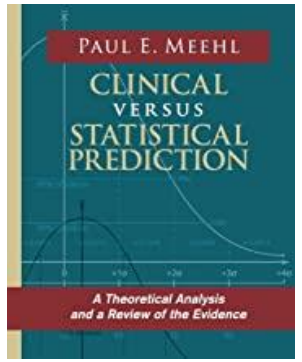
- Start with components that work.
- Don’t rely on anecdotes, testimonials or claims of “best practices” as evidence of working.
- If you can’t answer “What is the probability of losing more than X in the next 12 months due to event Y?” then you aren’t doing risk analysis.

Experts vs. Algorithms

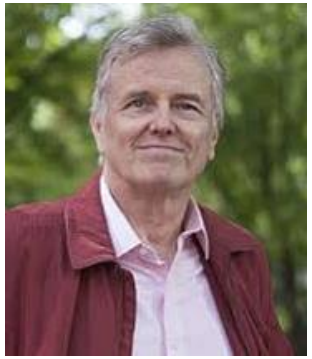
What the research says about statistical methods vs. Subject Matter Experts

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

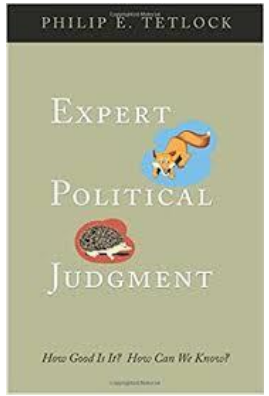
"There is no controversy in social science which shows such a large body of qualitatively diverse studies coming out so uniformly in the same direction as this one."



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.

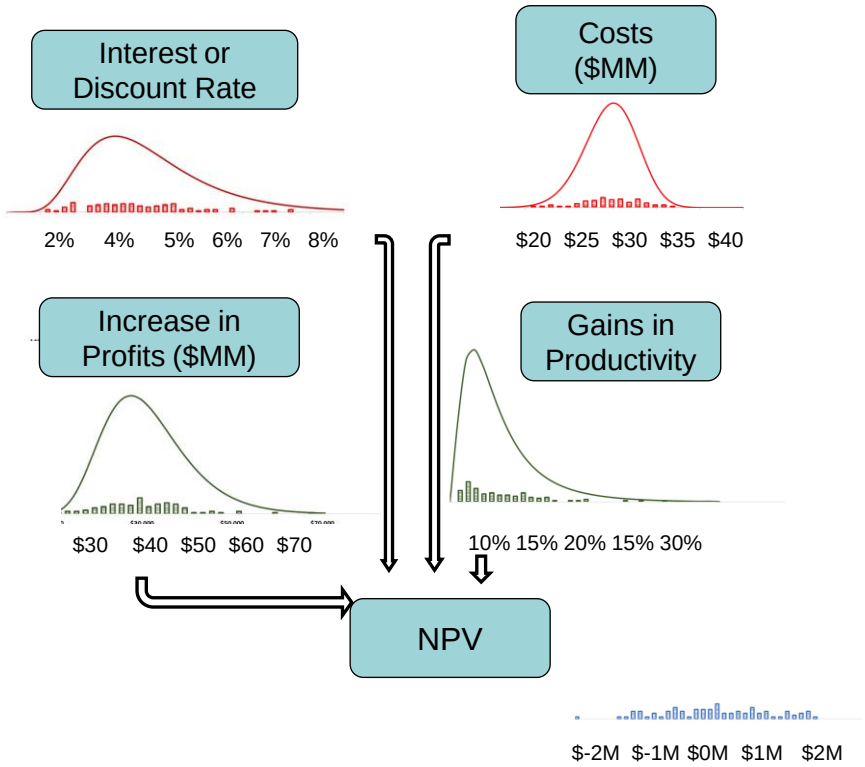


"It is impossible to find any domain in which humans clearly outperformed crude extrapolation algorithms, less still sophisticated statistical ones."



Monte Carlo Simulations

How to Model Uncertainty in Decisions



- We use Monte Carlo simulations to do the math with uncertain quantities.
- Research (I cite in the books) shows that people who build Monte Carlos *estimate better*.
- This allows us to do *real* risk analysis (i.e. compute the probability and magnitude of negative outcomes).

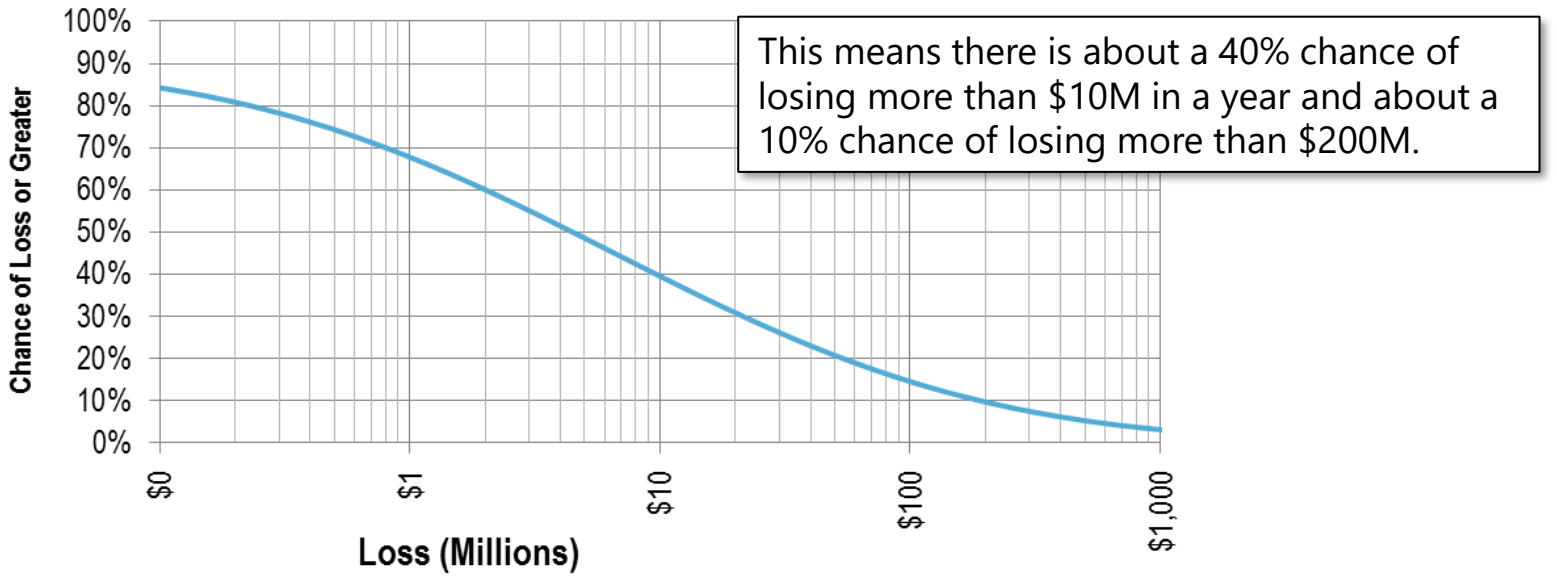
What Measuring Risk Looks Like

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

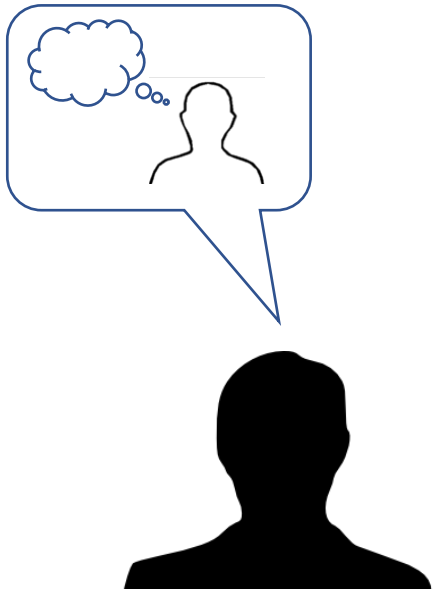


This means there is about a 40% chance of losing more than $10M in a year and about a 10% chance of losing more than $200M.

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

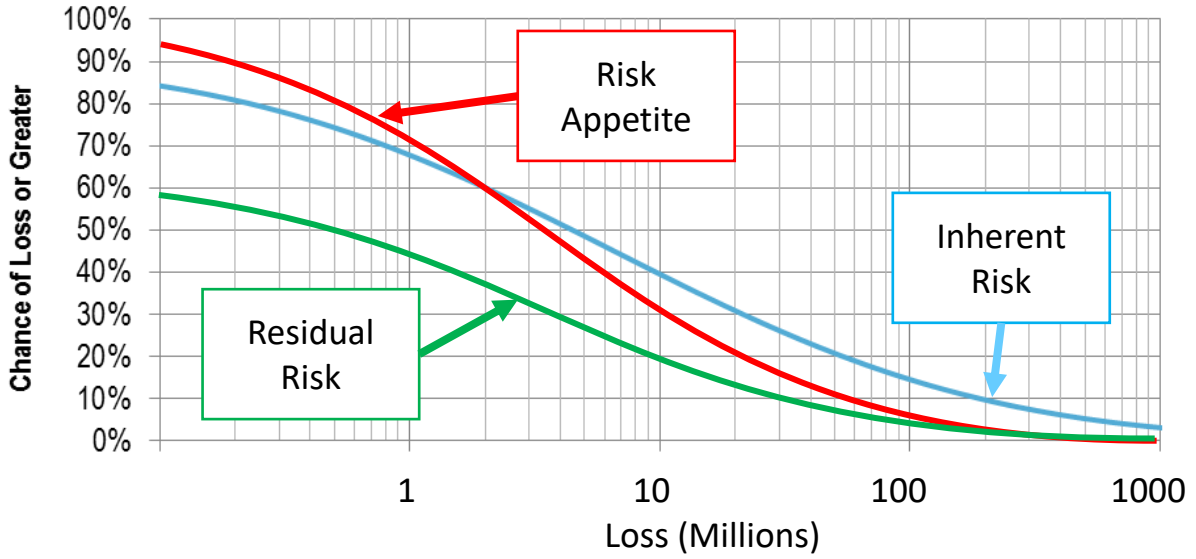
Dacher Keltner
University of California, Berkeley

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	↑

A Version of Risk Tolerance

The Loss Exceedance Curve

Unambiguous risk lets us have unambiguous risk tolerance.



What Measuring Risk Looks Like

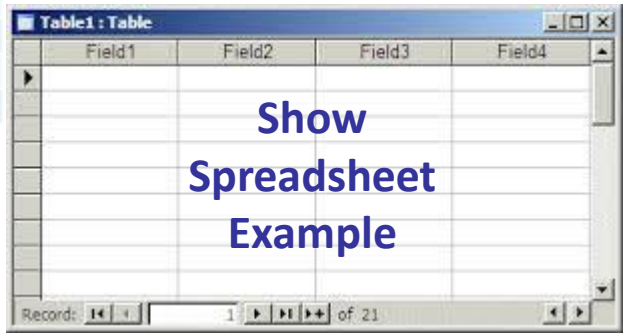
A Simple “One-For-One Substitution”

Each of these examples can be found on
https://www.howtomeasureanything.com/riskmanagement/

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."

"Risk management is too complex to model."

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

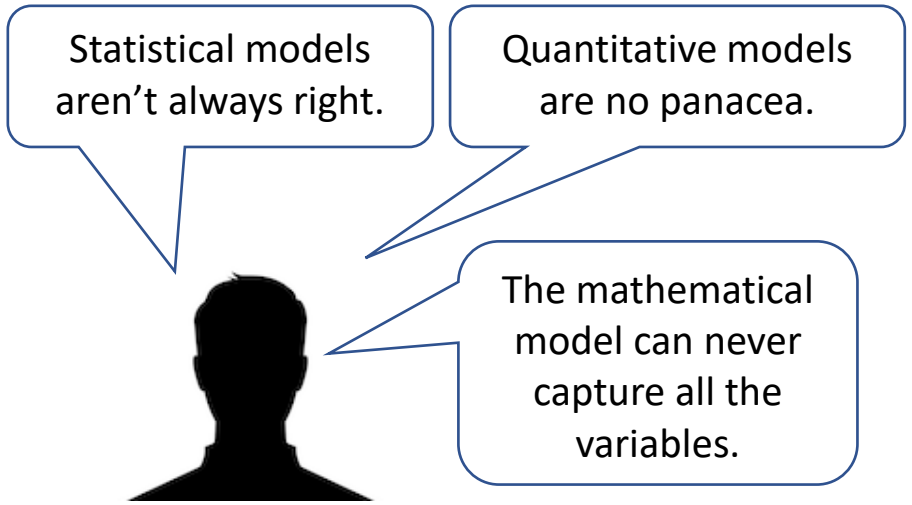
"How do you know you have all the variables?"

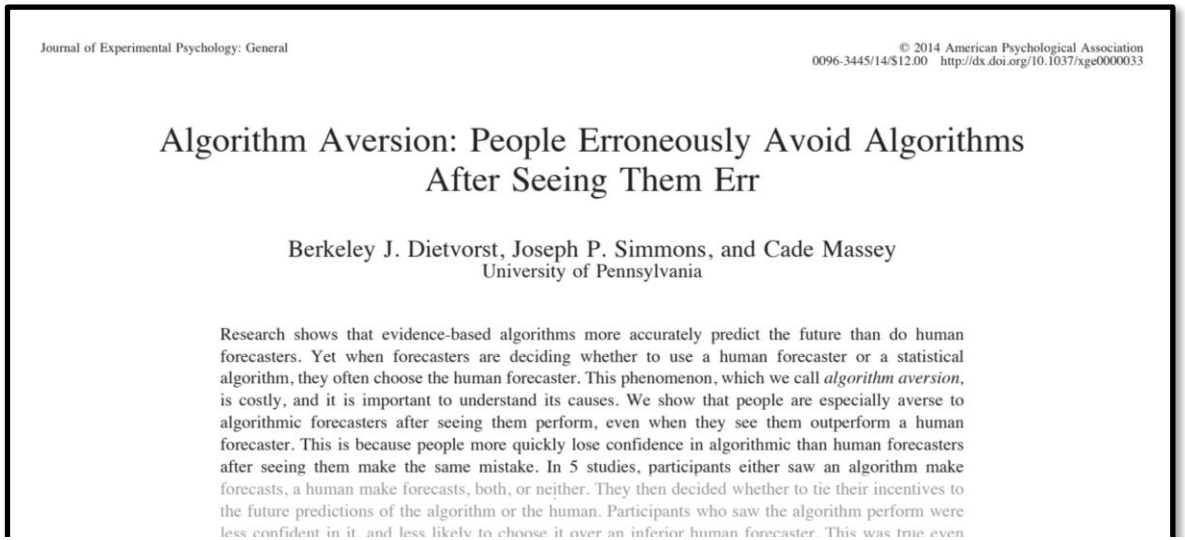
The implied (and unjustified) conclusion from each of these is….

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

Irrational Bias Against Algorithms

A Double Standard





Journal of Experimental Psychology: General

© 2014 American Psychological Association
0096-3445/14/$12.00 http://dx.doi.org/10.1037/xge0000033

Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err

Berkeley J. Dietvorst, Joseph P. Simmons, and Cade Massey
University of Pennsylvania

Research shows that evidence-based algorithms more accurately predict the future than do human forecasters. Yet when forecasters are deciding whether to use a human forecaster or a statistical algorithm, they often choose the human forecaster. This phenomenon, which we call *algorithm aversion*, is costly, and it is important to understand its causes. We show that people are especially averse to algorithmic forecasters after seeing them perform, even when they see them outperform a human forecaster. This is because people more quickly lose confidence in algorithmic than human forecasters after seeing them make the same mistake. In 5 studies, participants either saw an algorithm make forecasts, a human make forecasts, both, or neither. They then decided whether to tie their incentives to the future predictions of the algorithm or the human. Participants who saw the algorithm perform were less confident in it, and less likely to choose it over an inferior human forecaster. This was true even

Don't commit the classic "Beat the Bear" fallacy.

Exsupero Ursus

Statistical Literacy vs. Attitudes About Quant

The Main Obstacle to Quantitative Methods

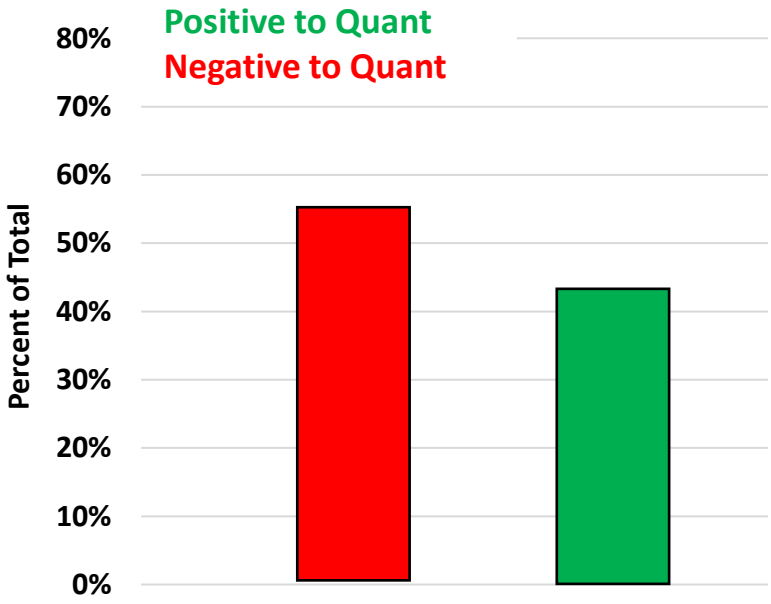
"Our thesis is that people have strong intuitions about random sampling…these intuitions are wrong in fundamental respects."

Daniel Kahneman and Amos Tversky, *Psychological Bulletin*, 1971

2016 HDR Cybersecurity Survey

Statistical Literacy vs.

Attitudes About Statistical Methods



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

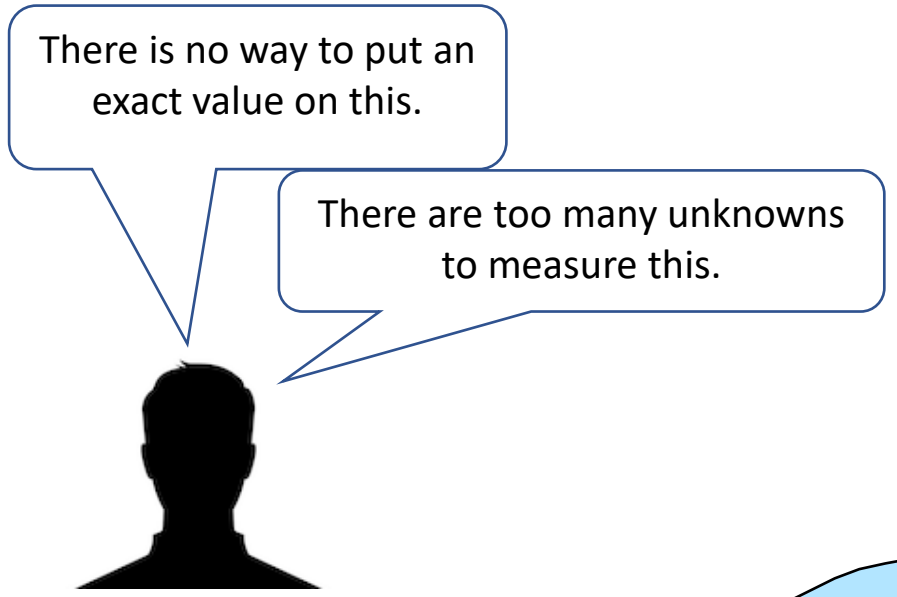
CONCEPT of Measurement	The definition of measurement itself is widely misunderstood.
OBJECT of Measurement	The thing being measured is not well defined.
METHOD of Measurement	Many procedures of empirical observation are misunderstood.

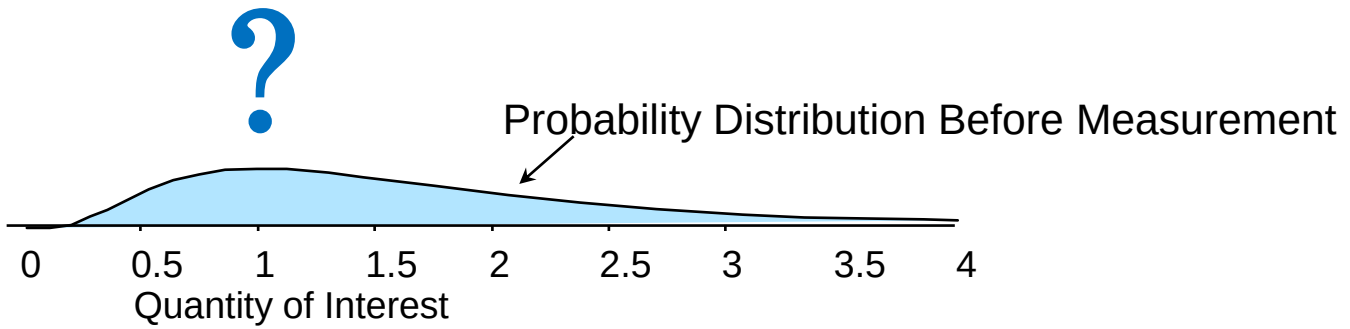
The Concept of Measurement

- What Measurement Really Means

It's not a point value.

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



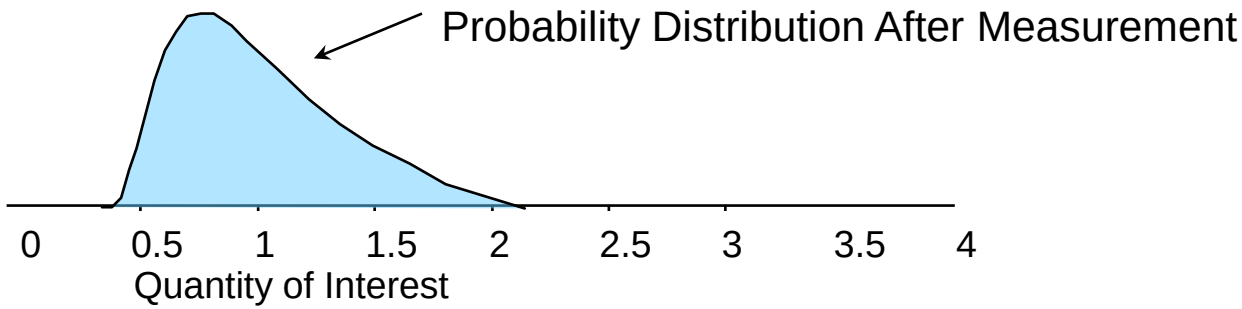


The Concept of Measurement

- What Measurement Really Means

It's not a point value.

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



The Concept of Measurement

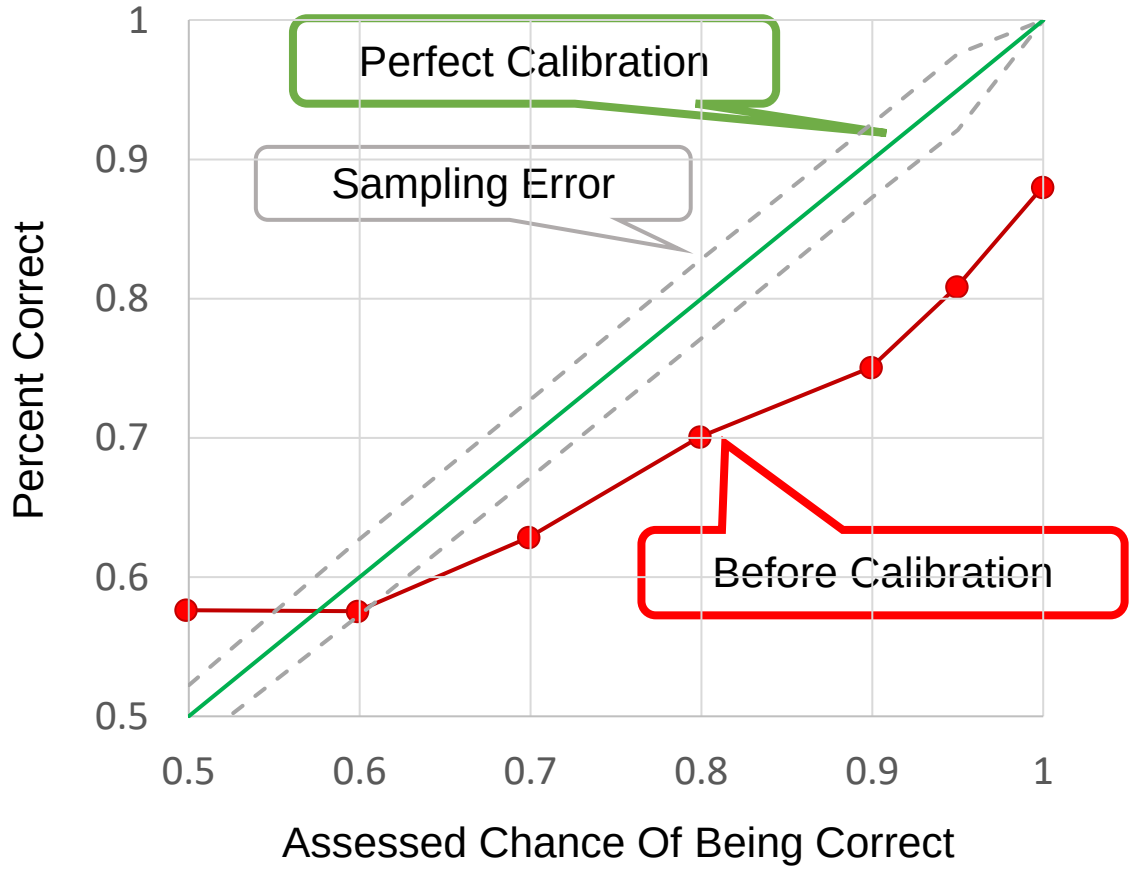
What the research says about Subject Matter Experts

"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

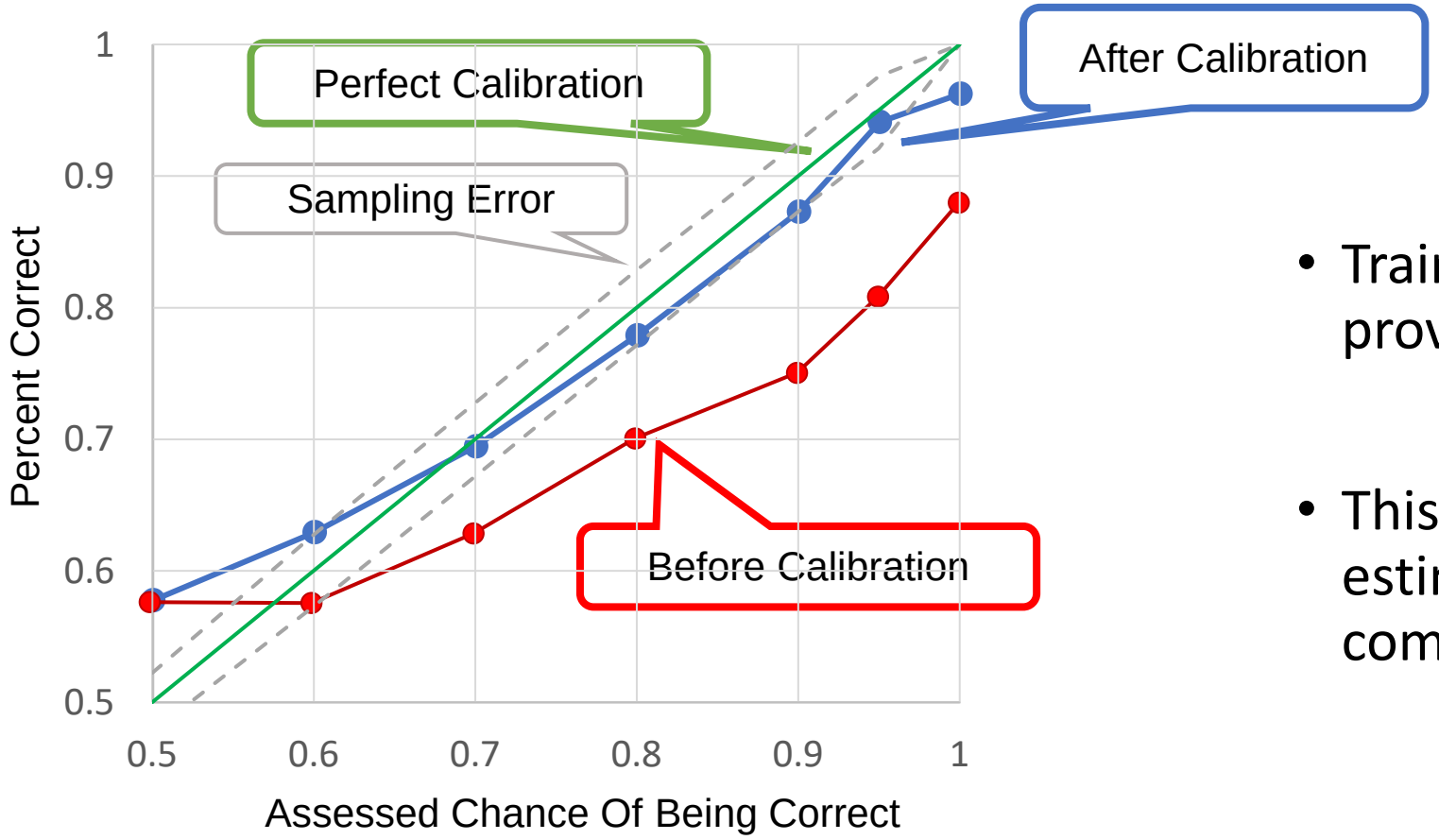
- Decades of studies show that most managers are statistically "overconfident" when assessing their own uncertainty.
- Studies also show that measuring *your own* uncertainty about a quantity is a general skill that can be taught with a ***measurable*** improvement.

Measuring Overconfidence



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

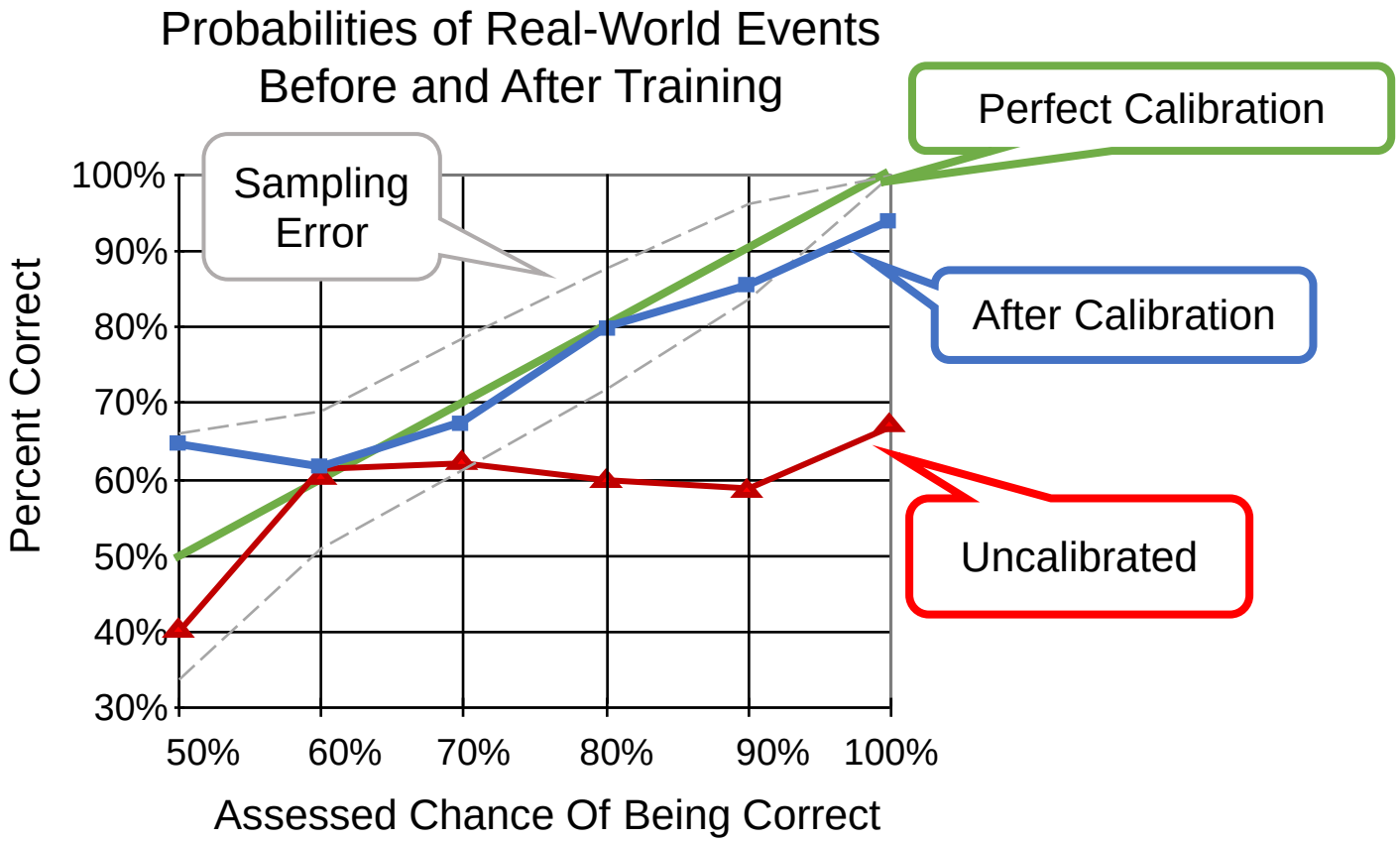
Measuring Calibration Training



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

Training vs Real World

- I conducted a calibration training experiment with 16 IT Industry Analysts and 16 CIOs to test if calibrated people were better at putting odds on uncertain future events.
- The analysts were calibrated and all 32 subjects were asked to predict 20 IT Industry events.
- Example: Steve Jobs will be CEO of Apple again, by Aug 8, 1997 - True or False? Are you 50%, 60%...90%, 100% confident?



Combining Experts With Bayes

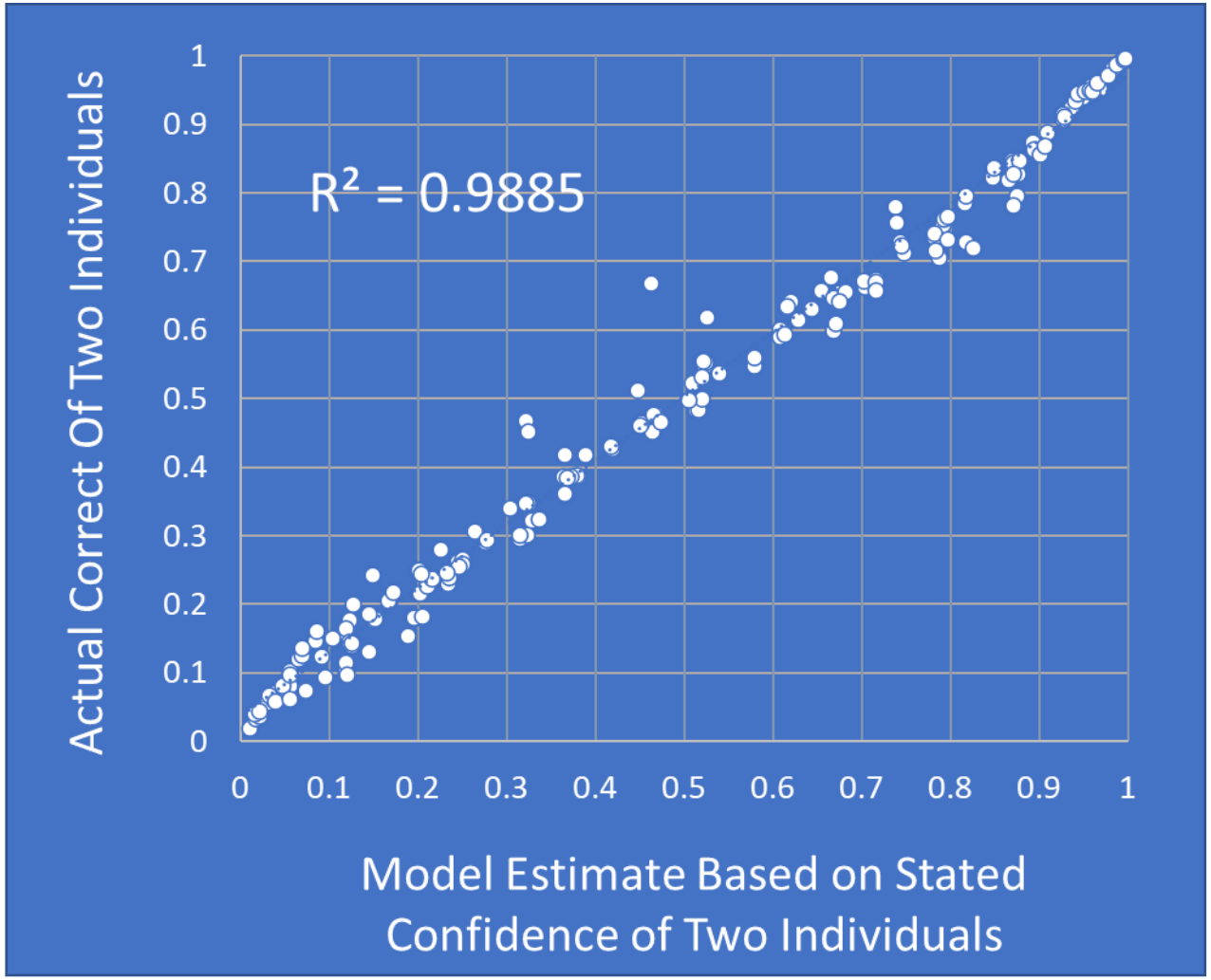
Bayes Theorem: $P(X|Y) = \frac{P(X)P(Y|X)}{P(Y)} = \frac{P(X)P(Y|X)}{\sum_i P(Y|X_i)\,P(X_i)}$

P(X) = the probability of X

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

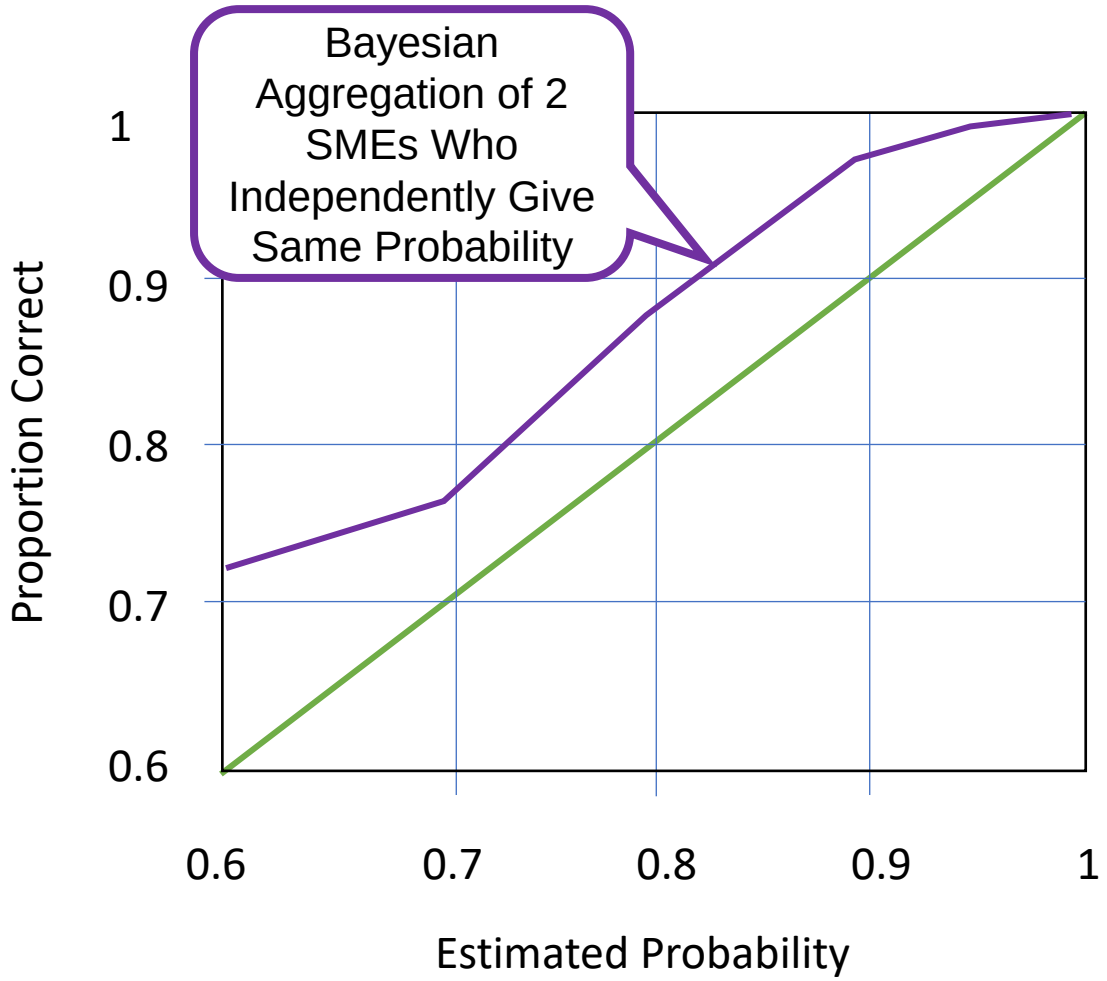
$\sum_i P(Y \mid X_i)\,P(X_i)$ = the sum of the probability of Y under each possible condition

$$\frac{P(X|C_1 \dots C_n)}{1 - P(X|C_1 \dots C_n)} = \left(\frac{1 - P(X)}{P(X)}\right)^{n-1} \prod_{i=1}^{n} \frac{P(X|C_i)}{1 - P(X|C_i)}$$



- I generated over 380,000 random pairs of individuals who responded to the same question.
- When we look at all the combinations of probabilities that two people put on a claim being true, the Bayesian model which estimates team performance based on individual performance is a good predictor of actual team performance.

Aggregating Subject Matter Experts

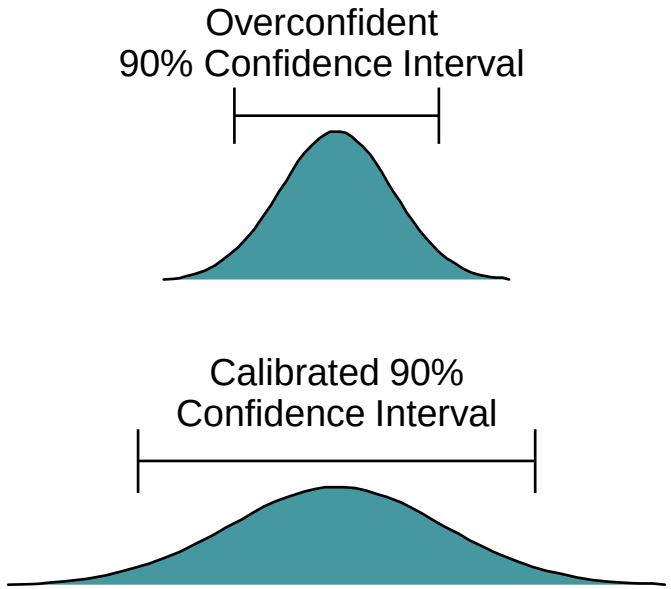


- There has been a lot of research on how to combine experts.
- Just averaging multiple experts is *not* the best method.
- A method based on Bayesian statistics shows that two experts should have less uncertainty than either expert alone.
- The math agrees with our data (R^2=.9885).

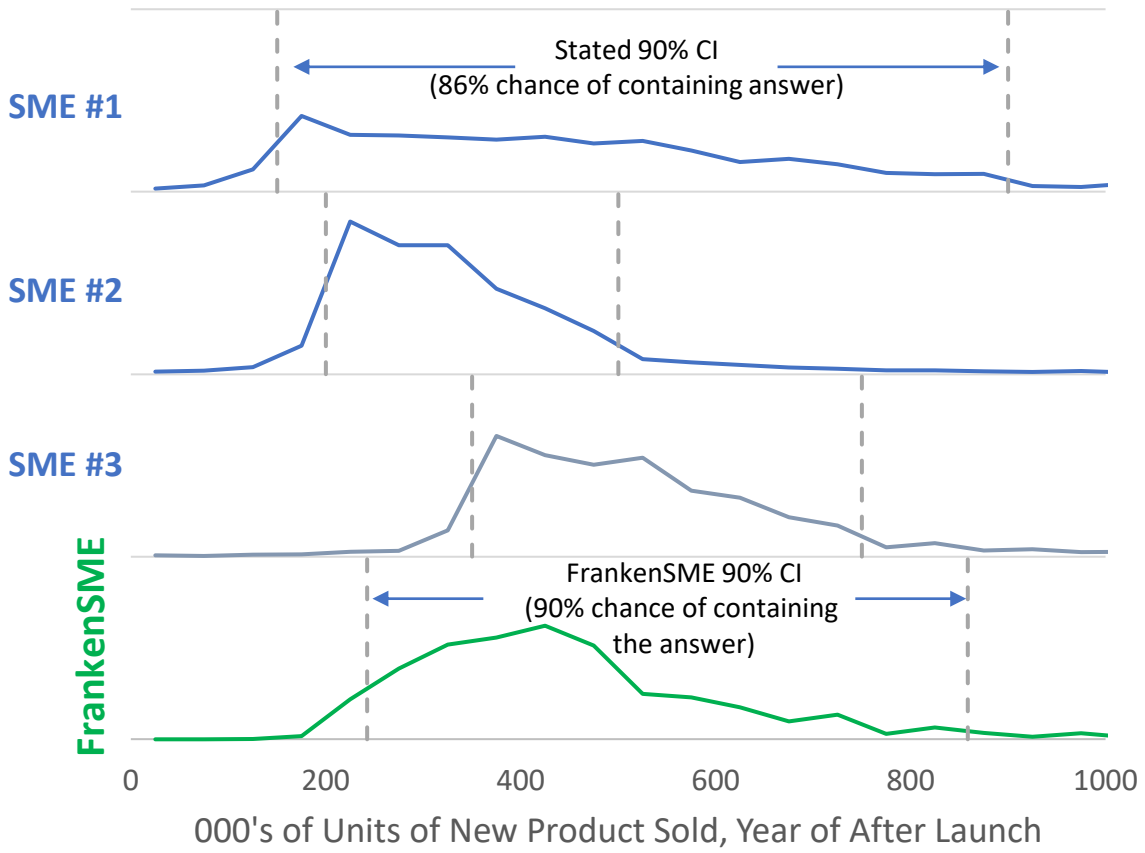
Overconfidence in Ranges

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%



Example of Three SME's, One FrankenSME



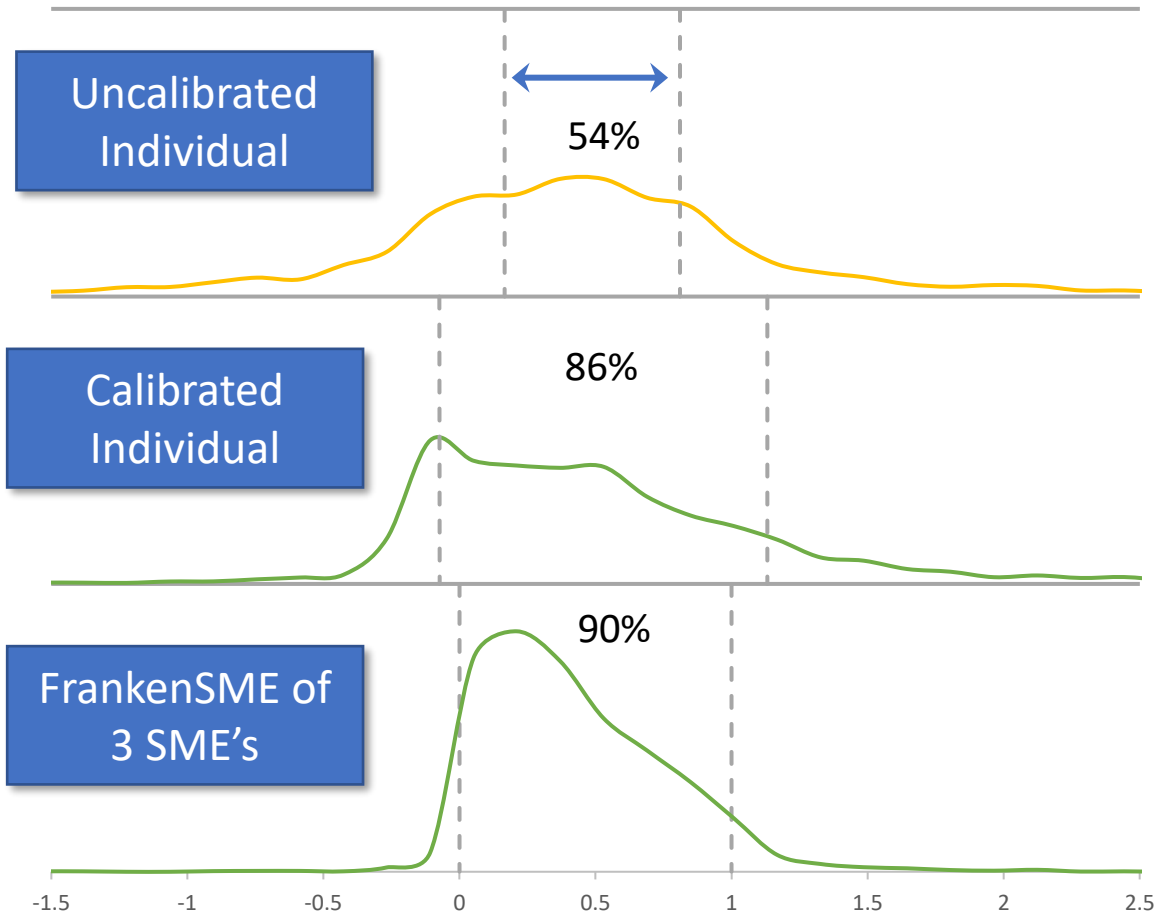
Three calibrated SME's are each asked to estimate the number of units sold of a new product the year after the launch.

The three SMEs give overlapping but not identical ranges.

Just like the binary probabilities, range estimates can be combined to produce a single range.

The solution is not a simple averaging. Averaging several people together actually makes a wider range than simply choosing the SME with the best track record and the best SME isn't as good as FrankenSME.

Uncalibrated, Calibrated & Combined



Prior to Training, SME's provide 90% Confidence Intervals which actually contain only 54% of the answers.

After training, they widen their ranges. 86% of answers fall withing stated intervals. (This data includes SMEs who are unresponsive to training.)

Combining SMEs not only makes the result perfectly calibrated (90% fall w/in 90% CI), but actually has a distribution with less variance than an individual SME.

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.

Risk Identification Problems

- **Incompleteness (Not Collectively Exhaustive):** The focus is usually on completeness but it's not the only problem.
- **Ambiguity:** Risks need to be defined well enough that observable examples can be imagined.
- **Overlap (Not Mutually Exclusive):** Risk can be correlated but shouldn't be double-counted. Are "fines" and "legal" both risks? Is a data breach overlapping risks of civil liability, operational risk, and regulatory fines?
- **Misclassification:** Some "risks" may be inconsistently classified by type of impact, type of cause, or may not actually be risks (this can also lead to overlap). Is failing to meet a growth goal a risk?

Examples of Identified Risks			
Assets	Cost of components	Loss of political support	Reputation
Bad debt	Customer satisfaction low	Machinery failure	Revenue forecast missed
Bankruptcy of suppliers or	Data security	Market acceptance	Seasonal risk
Brand fatigue	Difficult-to-sell product	Market changes	Staff sickness/absence
Business strategy	Environment	Natural disaster	Supply chain failure/delays
Cashflow	Espionage	New markets	Technology advances
Client attrition	Exchange rates	Operational risk	Technology breakdown
Competition: marketing	Failure of utilities e.g.	Patent theft/infringement	Theft
Competition: better intel	Health and safety	Poor management	Time-to-market
Competition: legal action	Lack of office space	Political instability e.g. coup,	Transportation delay/damage
Compliance	Lack of skills/expertise	Profit	Under-resourcing
Copyright theft	Loss of key skills	Recession	Unexpected demand

Causes vs. Consequences

- A common solution to disambiguation is to establish whether you are classifying risks by cause or consequence.
- You could say a data breach is a cause of a loss (although it has causes). A regulatory (compliance related) loss is a consequence.
- There are multiple ways to model this but consistency is always desirable.

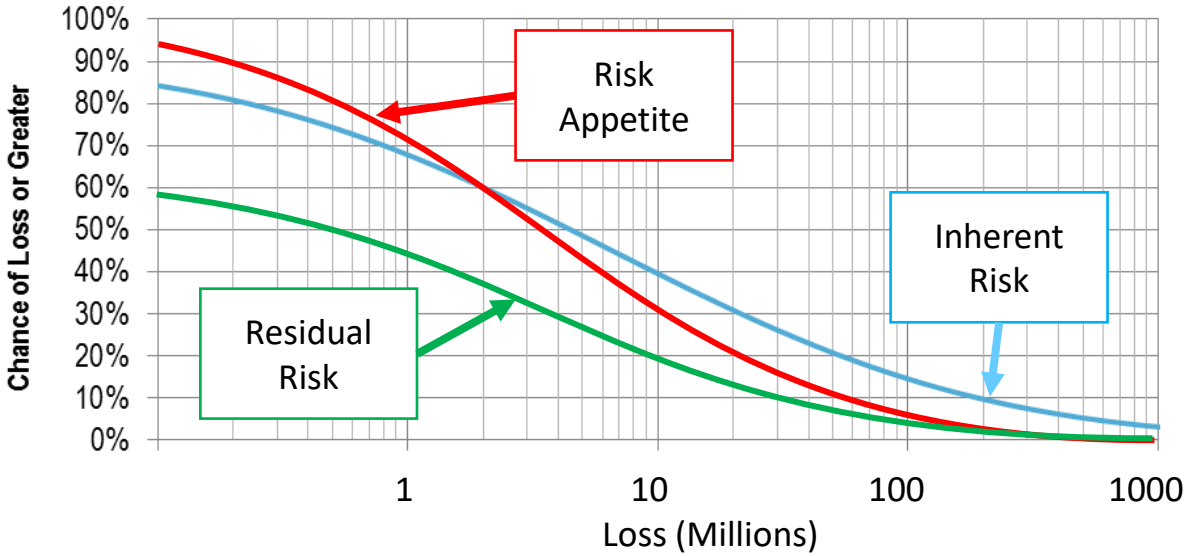
		Consequences			
		Suspended Ops	Regulatory Fine	Loss of Market	Etc.
Causes	Data Breach				
	Weather Event				
	Insider Theft				
	Etc.				

Clarification Test

What should the LEC include?

Is falling short of a sales goal really a risk? Is *not* approving a project really a risk? Is a recurring loss really a risk?
There is no hard rule on this but there are some guidelines:

- If a cost is predictable enough to budget for, it might not be what you want on an LEC.
- It should inform specific, consequential mitigation decisions.
- If you want to model uncertainty about benefits as a risk, you might be ready to adopt actual Decision Analysis.



The Object of Measurement

Clarifying the Problem

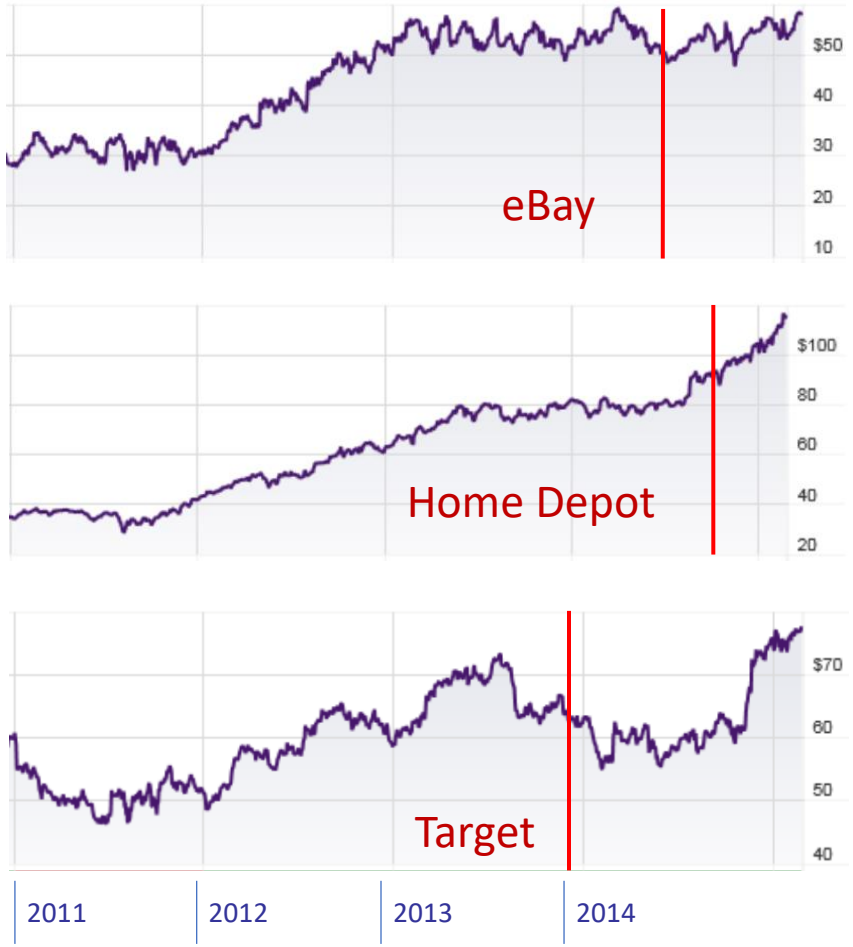
1. Why do you care? (What decision could depend on the outcome of this measurement?)
2. What do you see when you see more of it? (Describe it in terms of observable consequences, then units of measure.)
3. How much do you know about it now?
4. At what point will the value make a difference?
5. How much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

Another Small Sample Example

- THE *URN OF MYSTERY* PROBLEM
- There is a warehouse full of thousands 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

Improving Models with Empirical Data

Simply improving the method of eliciting expert estimates is just a start.

Now we need to inform the model with empirical data and continually update it based on new observations.

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.

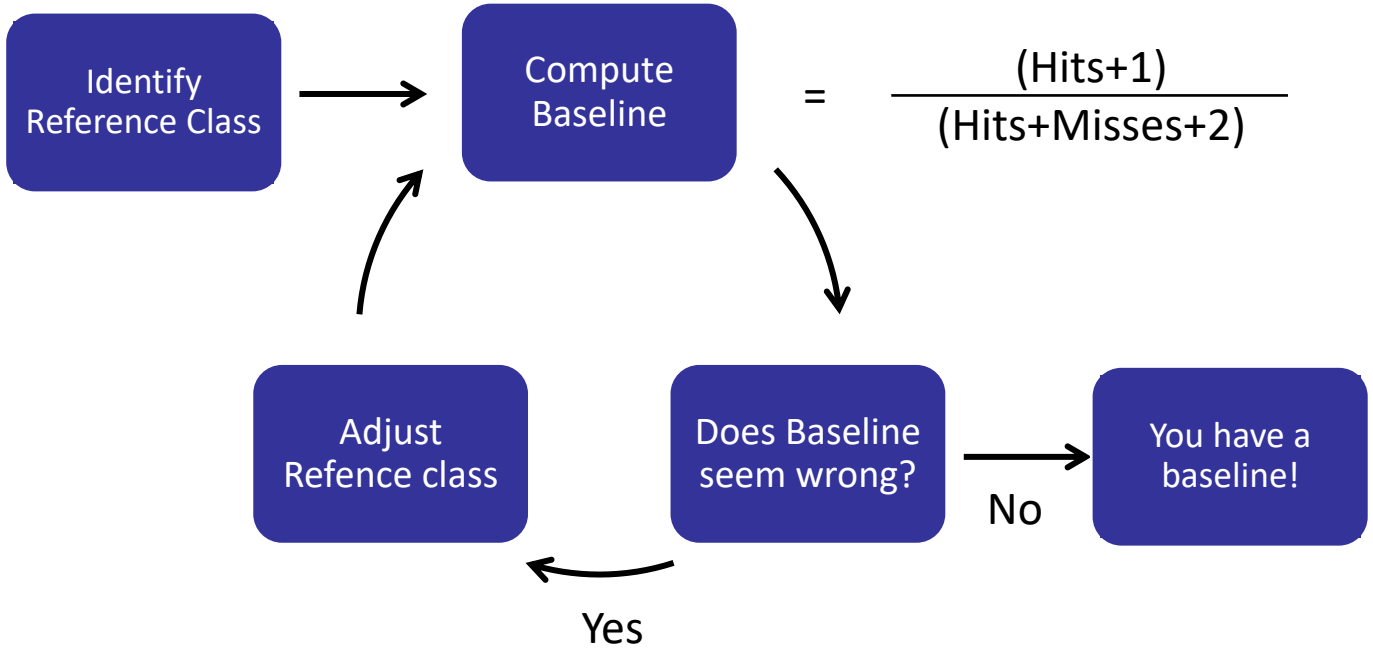
Pierre-Simon Laplace
1749-1827

- 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)

The Method of Measurement

Computing Baseline Probabilities

If the baseline seems too low or too high, it is probably because your reference class is larger than you first thought or because you believe a subset of it is more relevant.



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.

The Method of Measurement

Other Handy "Naïve Estimators"

Mean of a beta distribution is alpha/(alpha+beta).

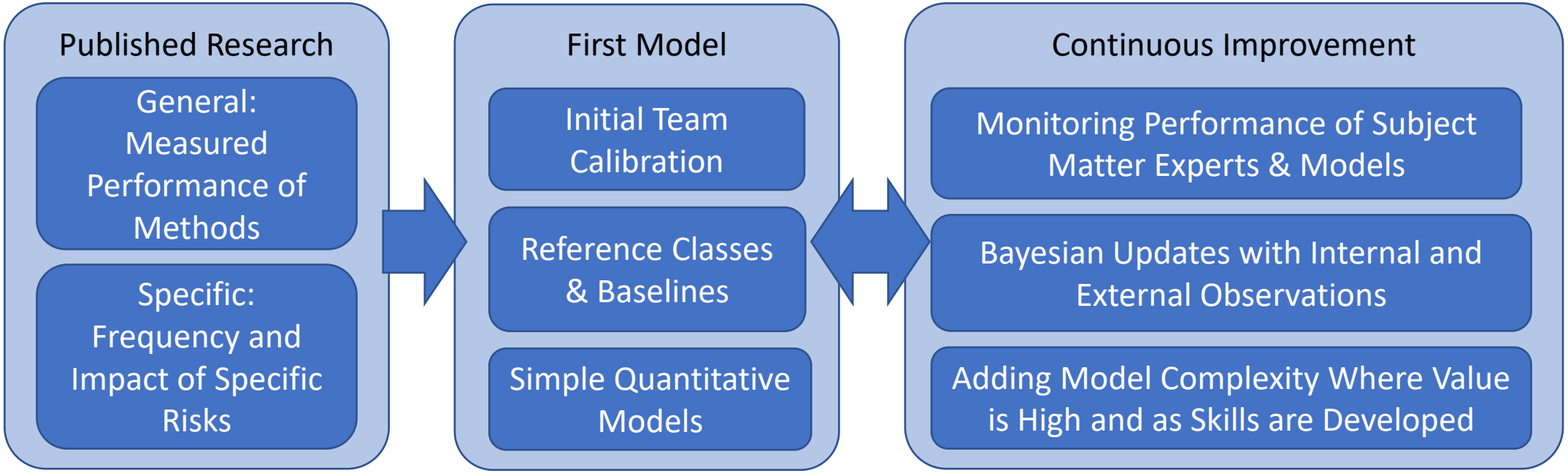
alpha=observed hits +1, beta=observed misses+1

These are all the means of beta distributions to different questions. The alpha and beta are "hits and misses" but with one "free" hit and miss.

- The chance of seeing an event that happened x times in y years in z organizations
  - **=(1+x)/(2+yz)**
- The chance that the next event will be worse than previous events:
  - **=1/(1+n)**

Your Real Job in Risk Management

You are a creator and manager of models – not just a "down in the weeds" estimator/forecaster.



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

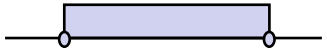
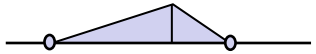
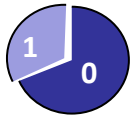
Supplementary Material

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

Basic Distributions

Each of these examples can be found on

www.howtomeasureanything.com/cybersecurity

Distributions*	Upper & Lower Bound	Best Estimate
Normal distribution 	Represents the "90% confidence interval"	Always half-way between upper and lower bound
Lognormal distribution 	Represents the "90% confidence interval"; the absolute lower bound of a lognormal is always 0	Always a function of the upper and lower bound
Uniform distribution 	Represents the absolute (100% certain) upper and lower bounds	NA
Triangular distribution 	Represents the absolute (100% certain) upper and lower bounds	Represents the mode; the most likely value
Binary distribution 	NA	Represents the % chance of the event occurring
Beta distribution 	Generates a value between 0 and 1 based on “hits” and “misses”	The mode of a beta is (hits-1)/(hits+misses-2)

*A “●” means a “hard” stop, an “➔” arrow means unbounded

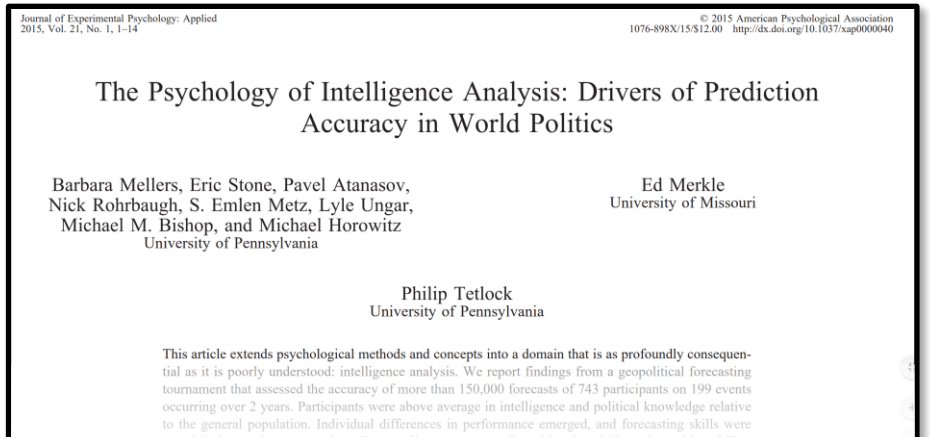
Selected Sources

- 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 Concept 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

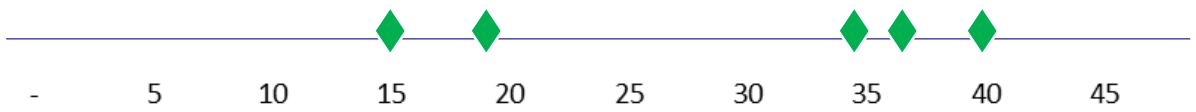
- Brains matter. Both topic expertise and overall IQ were the best predictors of performance.

The Method of Measurement

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?

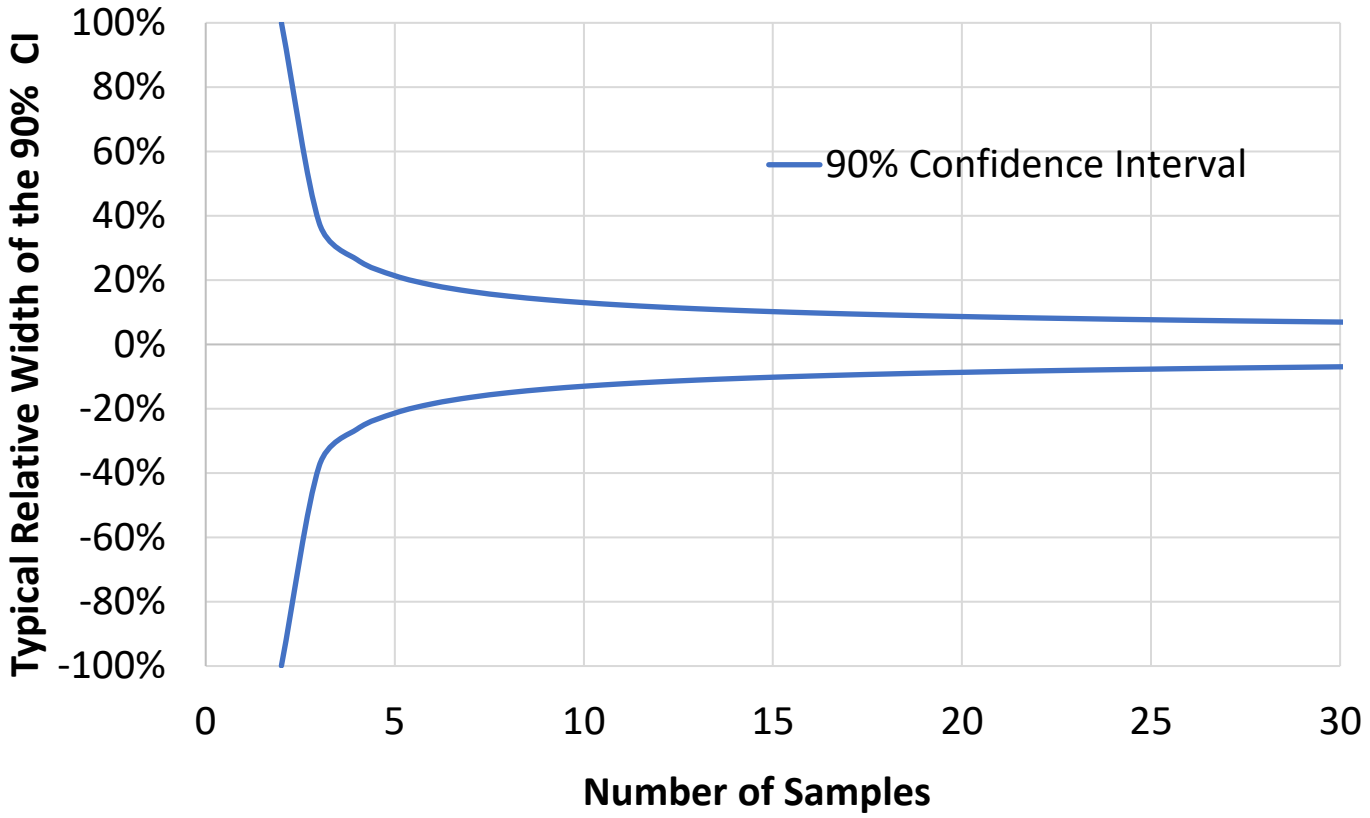


Minutes per day in activity X

The Method of Measurement

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…

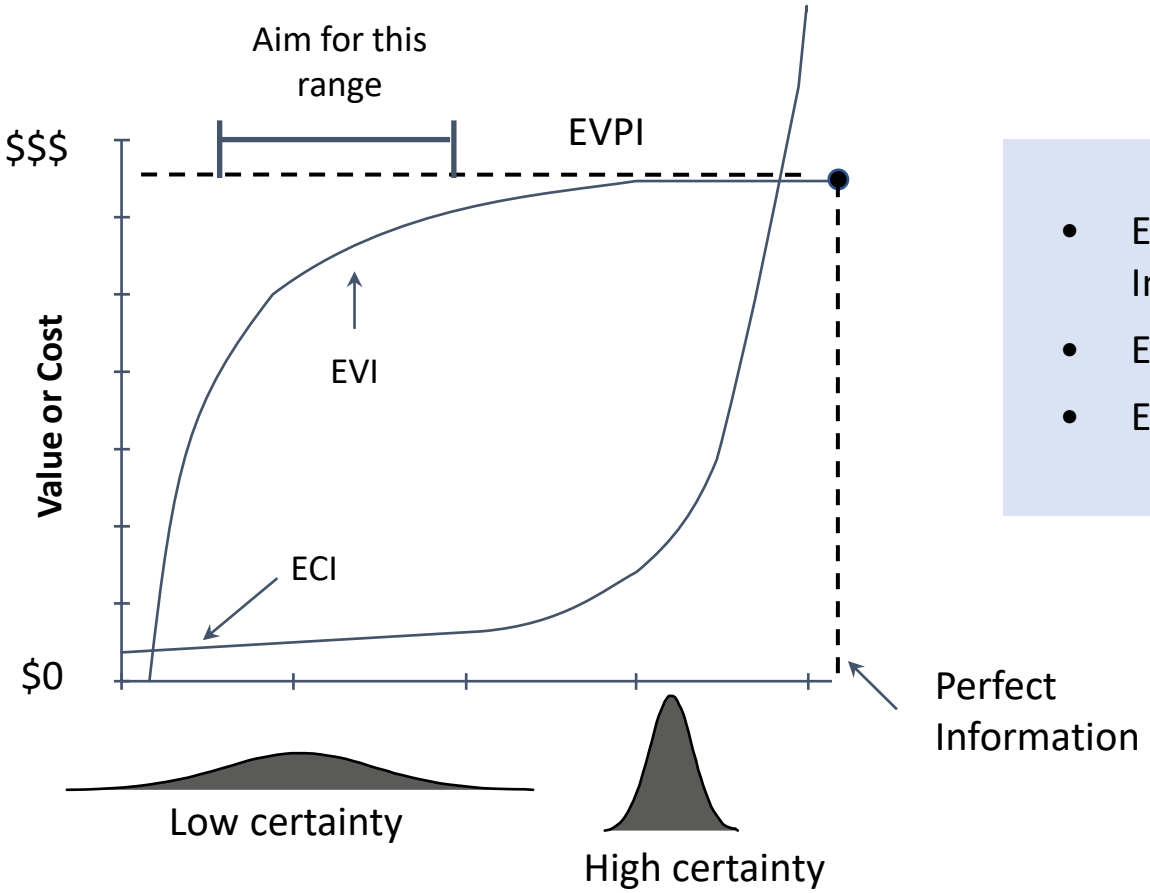
… each new sample reduces uncertainty a lot and the first few samples reduce uncertainty the most when initial uncertainty is high.

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.

The Method of Measurement

The Value of Information

If we can model uncertainty about decisions, we can compute the value of information.



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