Applied Information Economics:
Calibrated Probability Assessment Training

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

Uses of 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

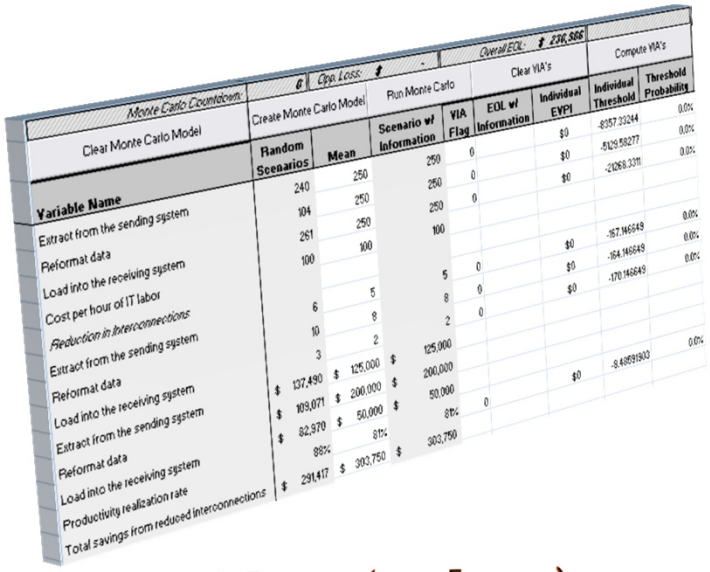
- Environmental policy
- Sustainable agriculture
- Procurement methods
- Grants management

Military

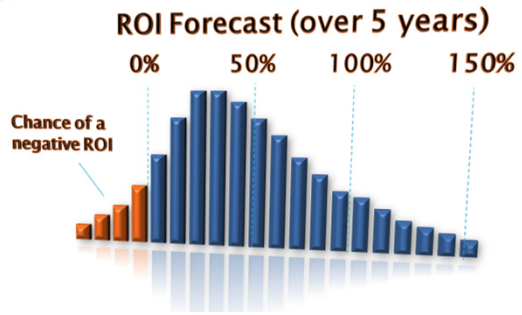
- Forecasting battlefield fuel consumption
- Effectiveness of combat training to reduce roadside bomb / IED casualties
- R&D portfolios

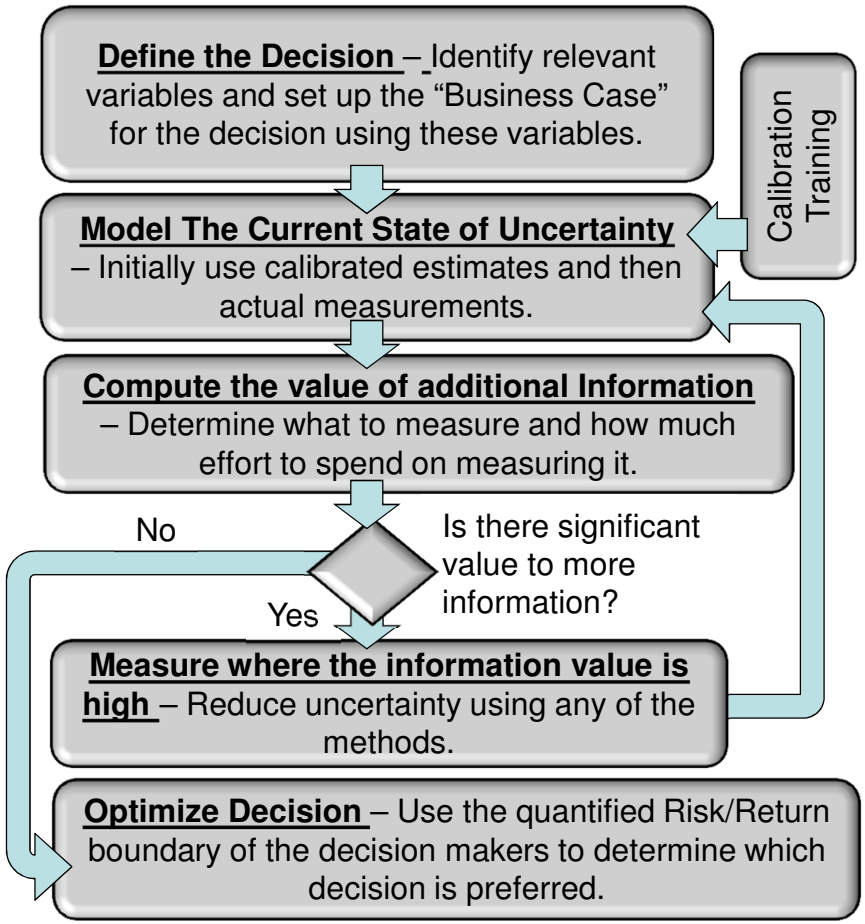
Hubbard Decision Research

Making the Best Bet



Monte Carlo Countdown:	6	Cap Loss:	$ -	Overall EOL:	$ 230,566				
Clear Monte Carlo Model	Create Monte Carlo Model		Run Monte Carlo			Clear VIA's		Compute VIA's	
Variable Name	Random Scenarios	Mean	Scenario w/ Information	VIA Flag	EOL w/ Information	Individual EVPI	Individual Threshold	Threshold Probability
Extract from the sending system	240	250	250	0		$0	-8357.33244	0.0%
Reformat data	104	250	250	0		$0	-5329.58277	0.0%
Load into the receiving system	261	250	100	0		$0	-21268.3311	0.0%
Cost per hour of IT labor	100	100						
Reduction in interconnections			5	0		$0	-167.146649	0.0%
Extract from the sending system	6	5	8	0		$0	-164.146649	0.0%
Reformat data	10	8	2	0		$0	-170.146649	0.0%
Load into the receiving system	3	2	125,000					
Extract from the sending system	$ 137,490	$ 125,000	$ 200,000					
Reformat data	$ 109,071	$ 200,000	$ 50,000			$0	-8.49691903	0.0%
Load into the receiving system	$ 82,970	$ 50,000	81%	0				
Productivity realization rate	86%	81%	$ 303,750					
Total savings from reduced interconnections	$ 291,417	$ 303,750						





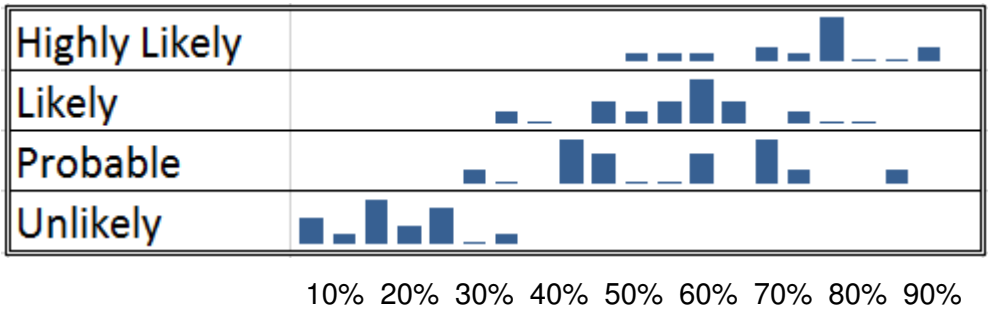
Hubbard Decision Research

Do “Scores” and “Scales” Work?

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

- Scales obscure (rather than alleviate) the lack of information (Budescu)
- Arbitrary partitions have unexpected effects on scoring behavior (Fox)
- The added error makes them “worse than useless” (Cox).

23 NATO officers estimates of probabilities for events described using common terms used in communicating likelihoods in intelligence reports (e.g. “War between X and Y is…”



Excerpt from: Richards Heuer, *The Psychology of Intelligence Analysis,* Center for the Study of Intelligence, CIA, 1999

Hubbard Decision Research

Are You Calibrated?

- Imagine processes in the industry where the use of an uncalibrated instrument could be very costly.
- Where could the use of an uncalibrated instrument do the most damage?
- Related Questions:
  - What is your most important decision?
  - What is your single biggest risk?

Hubbard Decision Research

Overconfidence

"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

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

Will Rogers

Hubbard Decision Research

What You Know: Calibrated Estimates

- Decades of studies show that most managers are highly uncalibrated when assessing their own uncertainty.
  - Studies showed that bookies were great at assessing odds subjectively, while doctors were terrible.
- Studies also show that measuring *your own* uncertainty about a quantity is a general skill that can be taught with a ***measurable*** improvement
- Training can "calibrate" people so that of all the times they say they are 90% confident, they will be right 90% of the time.

Hubbard Decision Research

Benchmarking Your Calibration

For the initial calibration test, you have two types of questions:

- For the questions that ask for a range, provide an upper and lower bound that you are 90% certain contains the correct answer.
- For the true/false questions, indicate “T” or “F” and then enter the percentage that best represents your confidence in your response

Hubbard Decision Research

Expected vs. Actual

- To determine your level of calibration, we need to compare actual outcomes to “expected” outcomes.
- In decision analysis, the word “expected” literally means probability weighted average.
- For the questions that ask for a 90% confidence interval, you expect to get 90% between your upper and lower bounds by definition.
- For the true/false questions, your expected number correct is equal to the total confidence on your answers. That is, if you were 50% confident on each, you expected to get half right, if you were 100% confident on each, you expect to get them all right, and so on.

Hubbard Decision Research

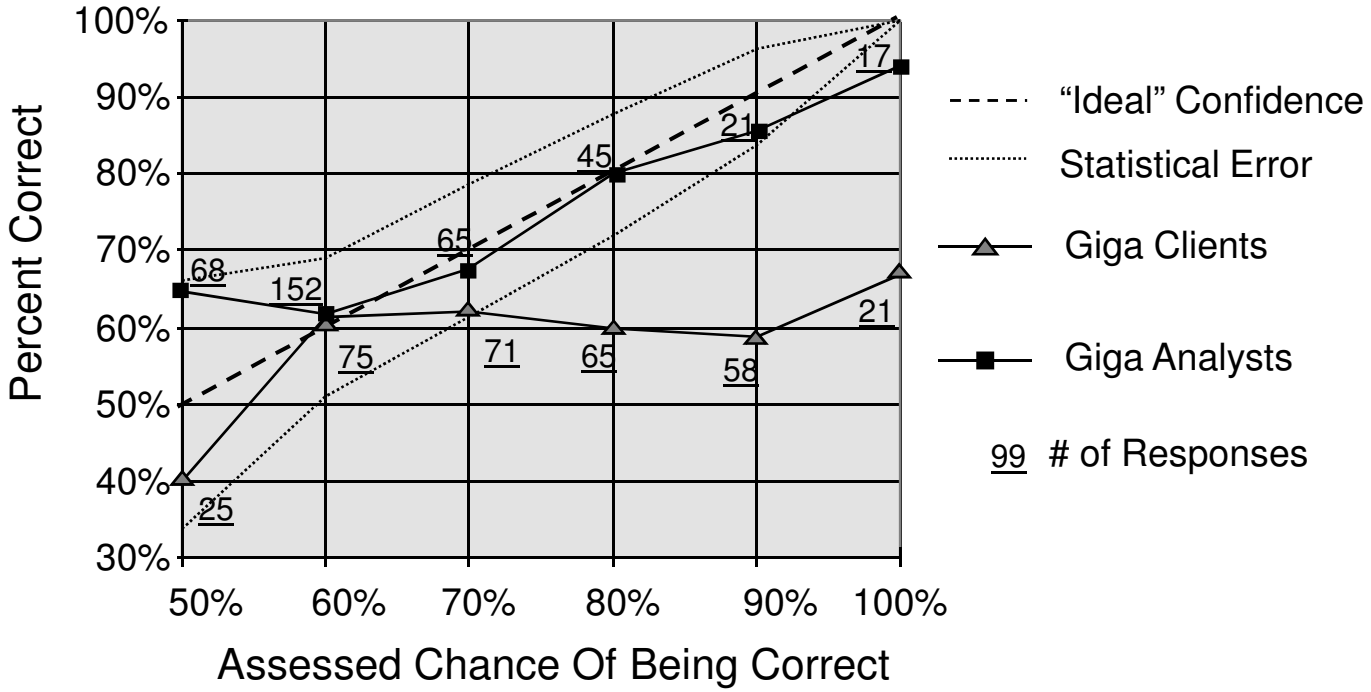
Common Confusions

- "How can I put a probability on it if I don't know the answer?"
- "We lack the data to put a probability on it."
- "I'm so uncertain about this, there is no way I can put a precise range or probability on it."
- "For the ranges, why don't I just put a ridiculously wide range on everything? Then everything will be right."
- "This is subjective. How can it be any good?"
- "I know some statisticians/I took a stats course once 20 years ago…this isn't mathematically valid."

Hubbard Decision Research

1997 Calibration Experiment

- In January 1997, I conducted a calibration training experiment with 16 IT Industry Analysts and 16 CIO's 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?

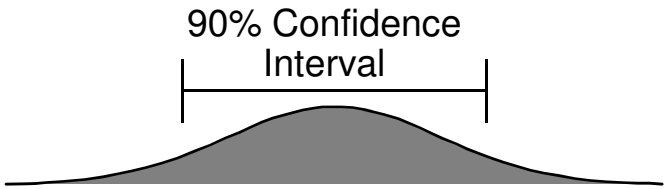


Source: Hubbard Decision Research

Hubbard Decision Research

Overconfidence in Ranges

- Most people are significantly ***overconfident*** about their estimates, especially educated professionals.



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%

Hubbard Decision Research

Calibration Aid: "The Equivalent Bet"

- For 90% Confidence Interval questions, which would you rather have?
  - **A**: Win $1,000 if your interval contains the correct answer
  - **B**: A 90% chance to win $1,000
- For the Binary Confidence questions, which would you rather have?
  - **A**: Win $1,000 if your answer is correct
  - **B**: A chance to win $1,000 equal to your stated confidence?

Hubbard Decision Research

The Odds: 10 Question Test

Even for a 10 question test many results will be conclusive.

Range Questions Correct out of 10	Probability of a calibrated person getting this number correct	Suggested range multiplier
10	34.87%	NA
9	38.74%	NA
8	19.37%	NA
7	5.74%	1.59
6	1.12%	1.95
5	1 in 612	2.44
4	1 in 6,807	3.14
3	1 in 109,630	4.27
2	1 in 2.7 million	6.49
1	1 109 million	13.09
0	1 in 10 billion	(Error)

93% chance a calibrated will be in this range

Possibly/slightly overconfident

Overconfident

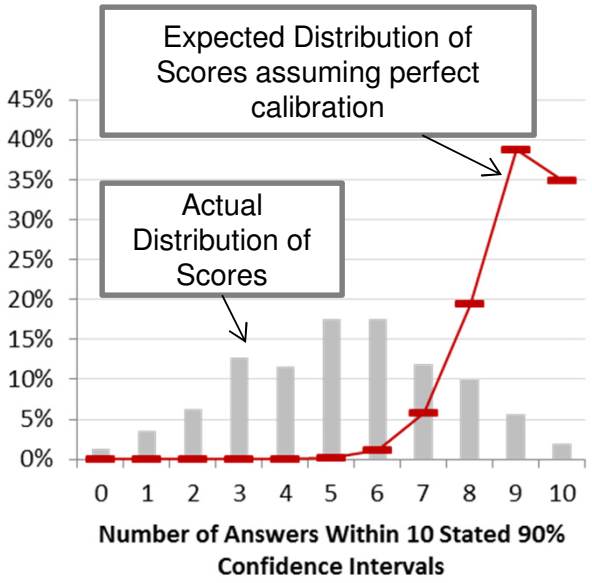
Indicates extreme overconfidence

Hubbard Decision Research

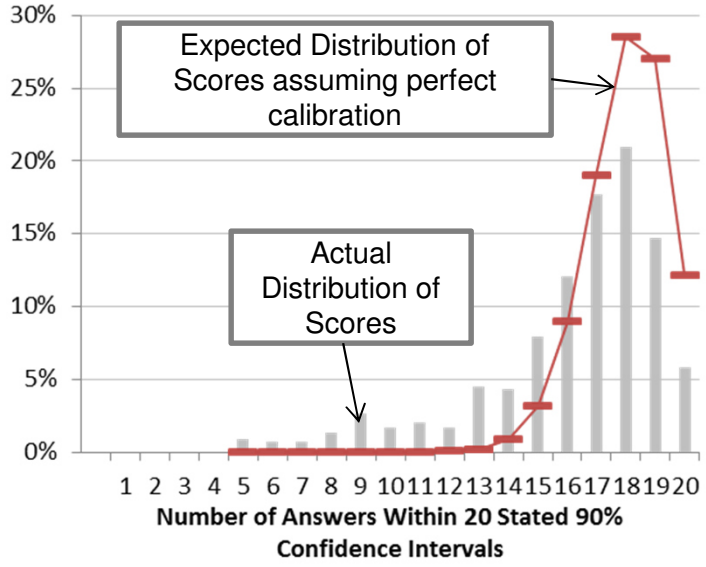
Newest Results on 90% CI Calibration

- With over 880 subjects who have taken the same calibration tests, and over 100,000 individual responses, a clear pattern emerges.
- Training has a major impact on 90% CI tests.
- About 15% don't quite reach calibration

Initial 10 Question 90% CI Test



Final 20 Question 90% CI Test



Hubbard Decision Research

A Few More Calibration Training Techniques

In addition to applying the equivalent bet:

1. Try something you almost never do - Challenge your answer
2. Anti-Anchoring strategies
   - Don't think of one number then add and subtract an error. Instead, treat each bound as a separate binary question (e.g. are you 95% certain the value is less than the upper bound?).
   - Think of absurdly wide ranges and then narrow them based on your knowledge instead of starting with narrow ranges and widening them.
3. Compressed Iteration and Feedback – more exercises, practice multiple strategies, strive to get better each time

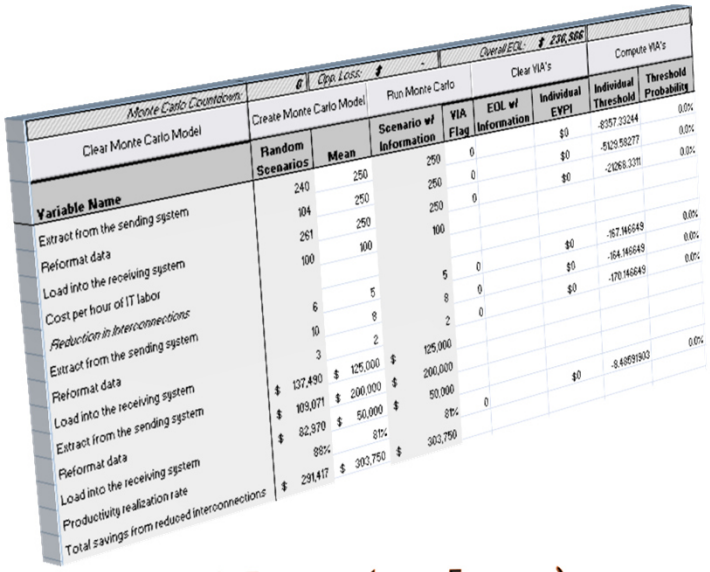
Hubbard Decision Research

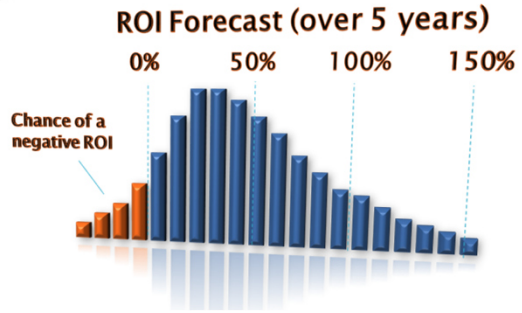
How to Think About Subjective Probabilities

- We use probabilities *because* we lack perfect data, not in spite of it.
- You already *imply* probabilities but you do it with ambiguous labels like “unlikely” or “medium”
- Ambiguous scales or labels do not alleviate the lack of data or alleviate complexity or change. They merely mask it.
- Excessively wide ranges don’t represent your knowledge any better than excessively narrow ranges. Overconfidence and Underconfidence are equally undesirable – but you are probably overconfident to start with.
- Training experts to use probabilities instead of ambiguous terms isn’t any less subjective and doesn’t require more data. It is just unambiguous.
- Skill with probabilities can also be measured against outcomes and outcomes can be used to update probabilities in mathematically-sound ways.

Hubbard Decision Research

Making the Best Bet





Define the Decision – Identify relevant variables and set up the "Business Case" for the decision using these variables.

Model The Current State of Uncertainty – Initially use calibrated estimates and then actual measurements.

Calibration Training

Compute the value of additional Information – Determine what to measure and how much effort to spend on measuring it.

Is there significant value to more information?

No

Yes

Measure where the information value is high – Reduce uncertainty using any of the methods.

Optimize Decision – Use the quantified Risk/Return boundary of the decision makers to determine which decision is preferred.

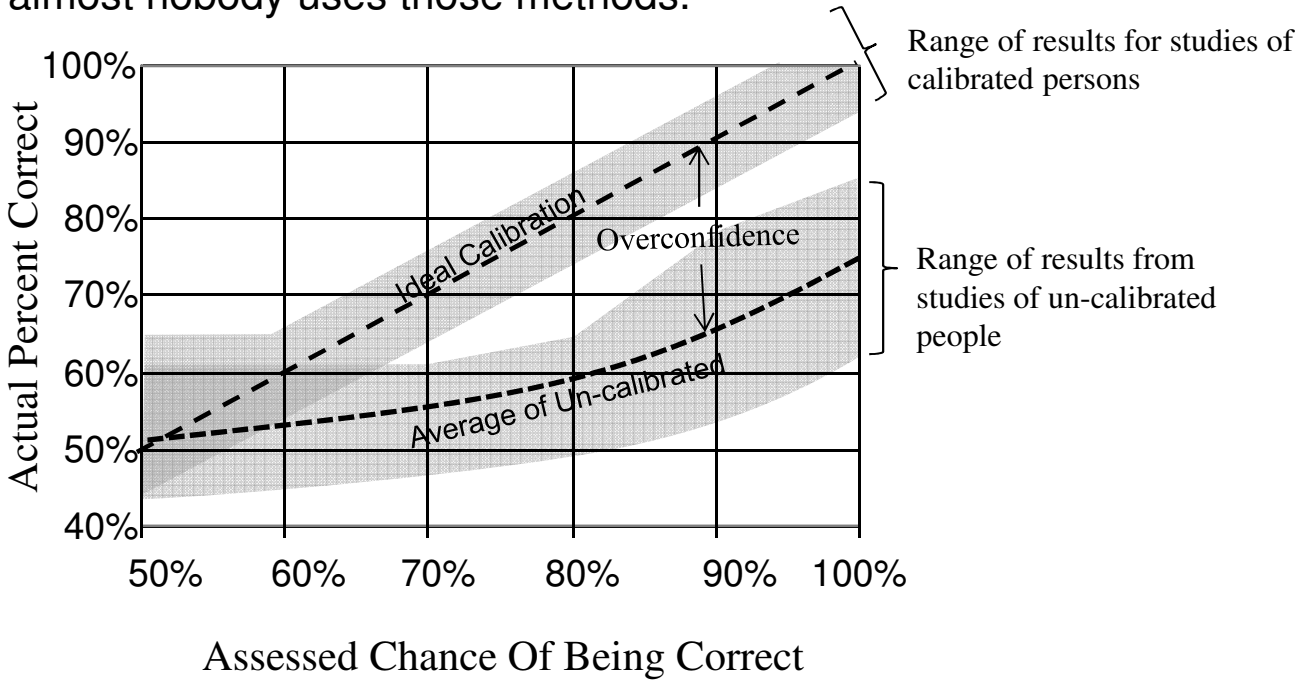
Hubbard Decision Research

Supplementary Calibration Material

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

Overconfidence

- This is the aggregate of 11 studies in how well people subjectively assess odds.
- The overwhelming evidence is that everyone is systematically “overconfident” when assessing probabilities.
- Fortunately, training and other techniques exist that adjust for this effect.
- Unfortunately, almost nobody uses those methods.



Hubbard Decision Research

The Odds: 20 Question Test

A 20 question test will have slightly better resolution – but still better at detecting overconfidence than under-confidence.

Range Questions Correct out of 20	Probability of a calibrated person getting this number correct	Suggested range multiplier
20	12.2%	NA
19	27.0%	NA
18	28.5%	NA
17	19.0%	NA
16	9%	1.28
15	3.2%	1.43
14	0.89%	1.59
13	0.20%	1.76
12	0.036%	1.95
11	0.005%	2.18
10	1 in 139,842	2.44
9	1 in 1.4 million	2.75
8	1 in 17 million	3.14
7	1 in 255 million	3.63

Possibly/slightly under-confident

75% chance a calibrated will be in this range

Possibly/slightly overconfident

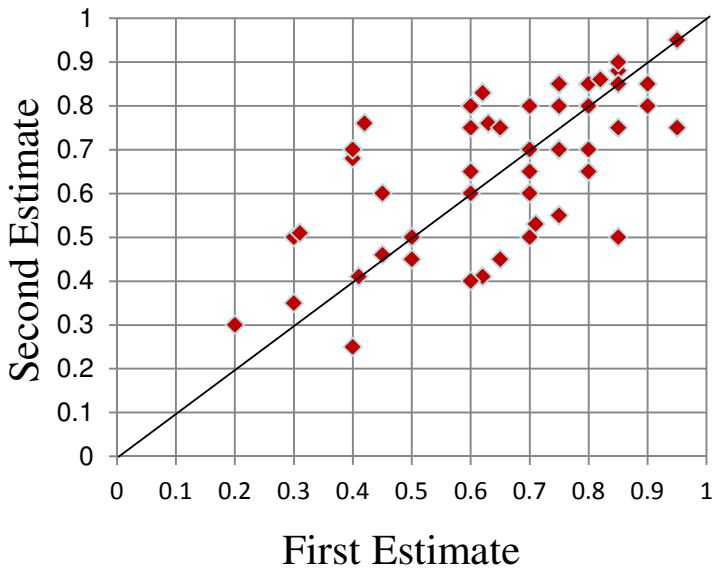
Overconfident

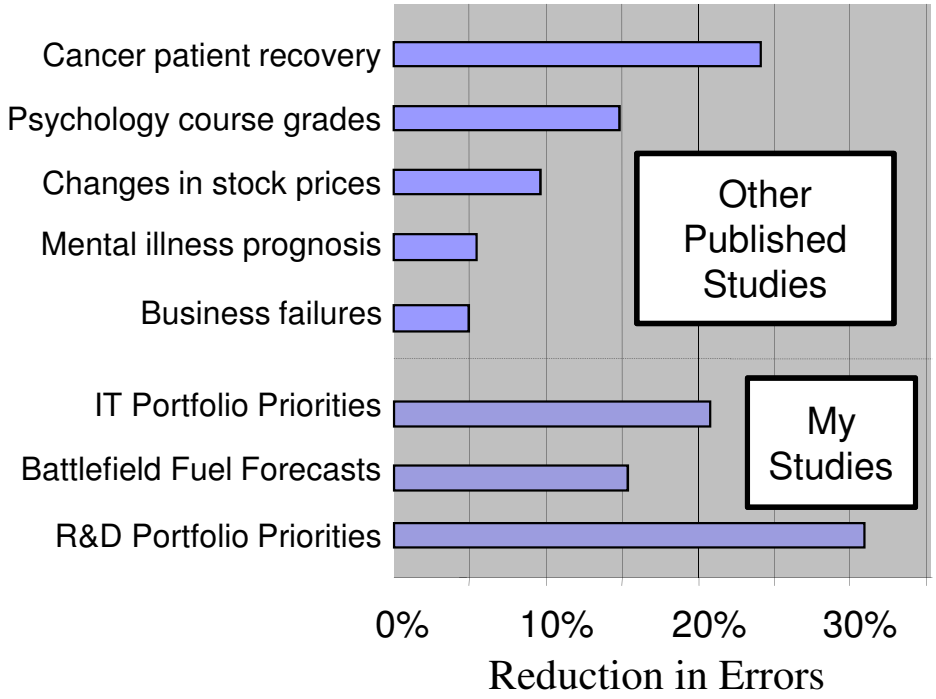
Indicates extreme overconfidence

Hubbard Decision Research

Measuring the Impact of Analysis Example

- No matter how much experience experts have, they appear to be unable to apply what they learned consistently
- Methods that statistically "smooth" their estimates show reduced error in several studies for many different kinds of problems





Hubbard Decision Research

Our Presenters

Richard Seiersen

Currently the General Manager of Cybersecurity and Privacy at GE Health Care. Data driven executive with ~20 years experience spanning subject matters in Cyber Security, Quantitative Risk Management, Predictive Analytics, Big Data and Data Science, Enterprise Integrations and Governance Risk and Compliance (GRC). Led large enterprise teams, provided leadership in multinational organizations and tier one venture capital backed start-ups.

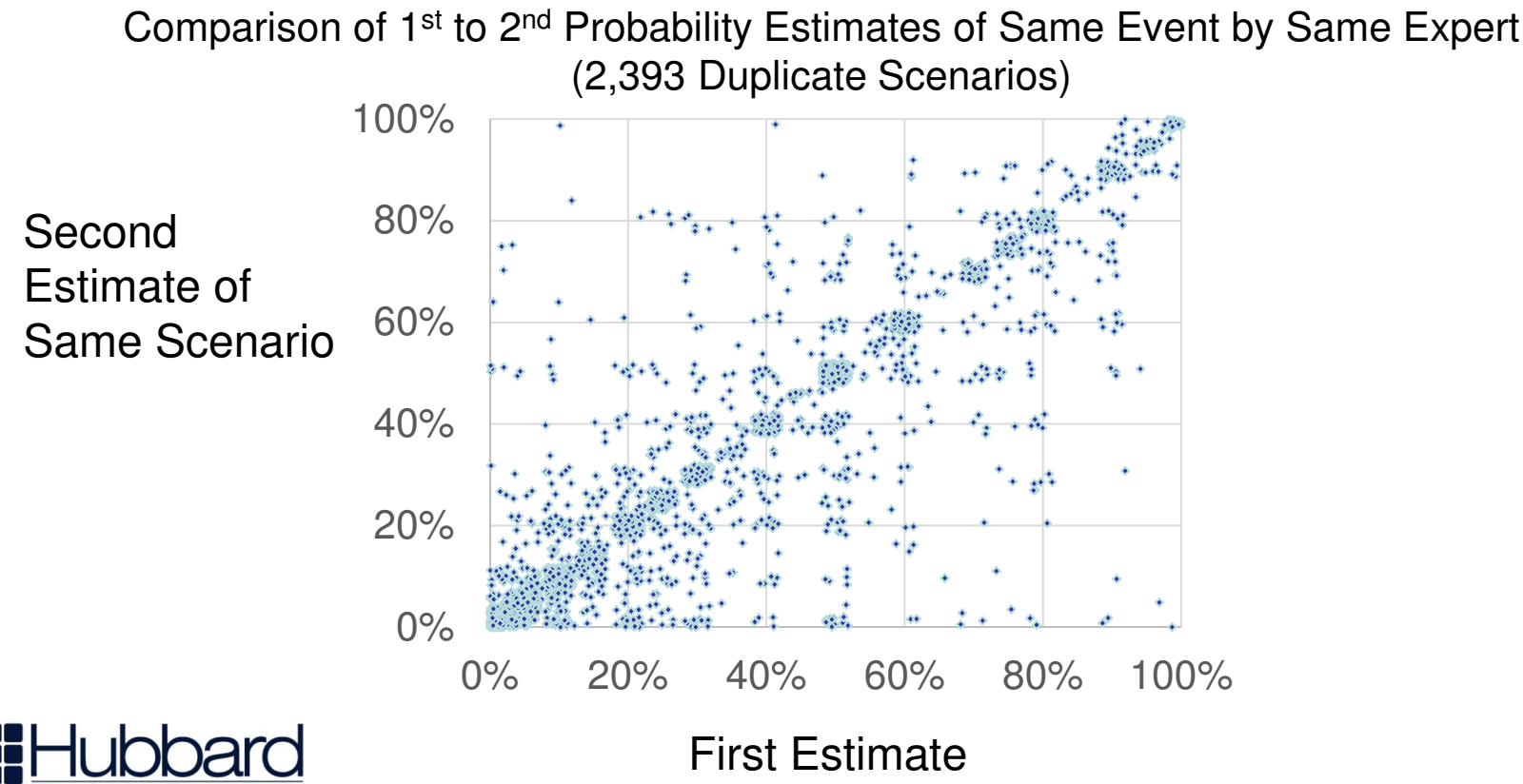
Douglas Hubbard

Mr. Hubbard is the inventor of the powerful Applied Information Economics (AIE) method. He is the author of the #1 bestseller in Amazon's math for business category for his book titled ***How to Measure Anything: Finding the Value of Intangibles in Business*** (Wiley, 2007; 3rd edition 2014). His other two books are titled ***The Failure of Risk Management: Why It's Broken and How to Fix It*** (Wiley, 2009) and ***Pulse: The New Science of Harnessing Internet Buzz to Track Threats and Opportunities*** (Wiley, 2011).

Hubbard Decision Research

Inconsistency Measurement Results

21% of variation in expert responses are explained by *inconsistency.*
(79% are explained by the actual information they were given)



Hubbard Decision Research