estimate

estimate • analyze • plan • control

Why Can't People Estimate: Estimation Bias and Mitigation

Dan Galorath

galorath@galorath.com

310 414-3222 x614

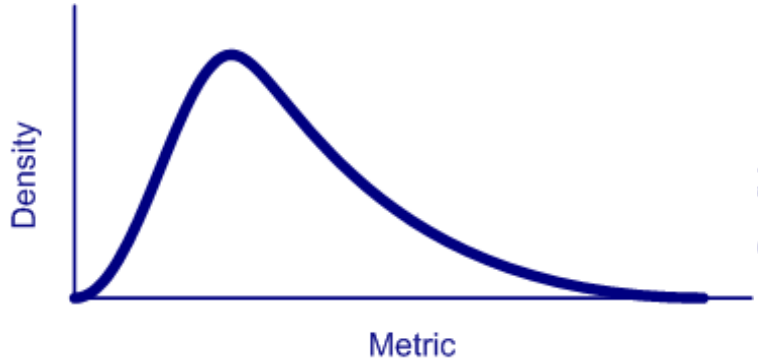
SEER by GALORATH

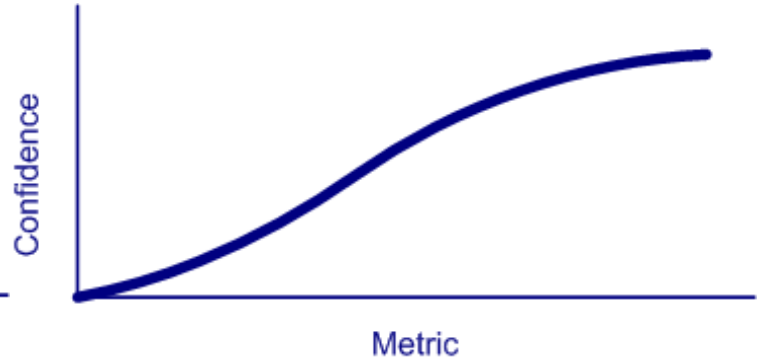
ESTIMATION & PLANNING:

An Estimate Defined

- An ***estimate*** is the most knowledgeable statement you can make ***at a particular point in time*** regarding:
  - Effort / Cost
  - Schedule
  - Staffing
  - Risk
  - Reliability

- Estimates more precise with progress
- ***A WELL FORMED ESTIMATE IS A DISTRIBUTION***





Estimation Methods Summarized

SEER by GALORATH

Category	Description	Advantages	Limitations
Guessing	Off the cuff estimates	Quick Can obtain any answer desired	No Basis or substantiation No Process Usually Wrong
Analogy	Compare project with past similar projects.	Estimates are based on actual experience.	Truly similar projects must exist Or analogy techniques used
Expert Judgment	Consult with one or more experts.	Little or no historical data is needed; good for new or unique projects.	Experts tend to be biased; knowledge level is sometimes questionable; may not be consistent.
Vendor Quotes	Vendor identification of scope & costs	Vendor has experience and (hopefully) data Vendor can commit to scope	Often assume best case.. Then exceed Customer costs not included
Agile Velocity		Helps root level management of Agile Projects	Doesn't estimate up-front well or provide answers for management decision making
Comprehensive Parametric Models	Perform overall estimate using design parameters and mathematical algorithms.	Models are usually fast and easy to use, and useful early in a program; they are also objective and repeatable.	Models can be inaccurate if not properly calibrated and validated; Bias in parameters may lead to underestimation.

Human Nature:
Humans Are Optimists

Harvard Business Review explains this Phenomenon:

- Humans seem hardwired to be optimists
- Routinely exaggerate benefits and discount costs

Delusions of Success: How Optimism Undermines Executives' Decisions (Source: HBR Articles | Dan Lovallo, Daniel Kahneman | Jul 01, 2003)

Solution - Temper with "outside view":
Past Measurement Results, traditional forecasting, risk analysis and statistical parametrics can help

Don't remove optimism, but balance optimism and realism

Cognitive Bias: How Fair Are We
(Source BeingHuman.org)

- Cognitive bias: Tendency to make systematic decisions based on cognitive factors rather than evidence
- Human beings exhibit inherent errors in thinking
- Researchers theorize in the past, biases helped survival
  - Our brains using shortcuts (heuristics) that sometimes provide irrational conclusions

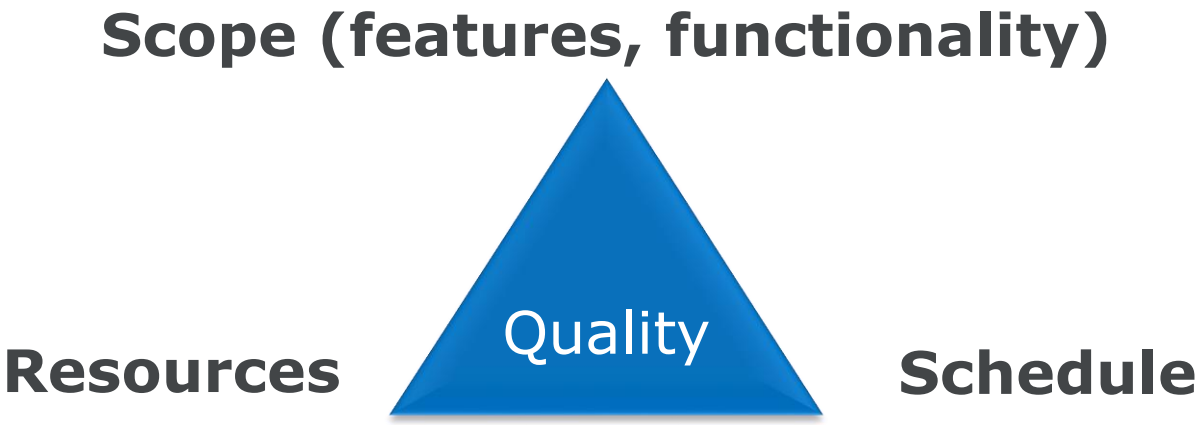
"We usually think of ourselves as sitting the driver's seat, with ultimate control over the decisions we made and the direction our life takes; but, alas, **this perception has more to do with our desires—with how we want to view ourselves—than with reality**." Behavioral economist Dan Ariely

- Bias affects everything:
  - from deciding how to handle our money
  - to relating to other people
  - to how we form memories

Essence of the problem: Memory is unreliable and we are hard wired to ignore risk & questioning

Trouble Starts By Bias or Strategic Mis-Estimation Ignoring Iron Triangle

- Typical Trouble: Mandated features needed within specific time by given resources



- At least one must vary otherwise quality suffers and system may enter impossible zone!

Sometimes strategic mis-estimation
is used to get projects started or to win
Some customers think price to win is strategic mis-estimation (it is not)

The Planning Fallacy (Kahneman & Tversky, 1979)

- Judgment errors **are systematic & predictable**, not random
  - Manifesting bias rather than confusion
  - Judgment errors made by experts and laypeople alike
  - Errors continue when estimators aware of their nature
- Optimistic due to overconfidence ignoring uncertainty
  - Underestimate costs, schedule, risks
  - Overestimate benefits of the same actions
- Root cause: Each new venture viewed as unique
  - "inside view" focusing on components rather than outcomes of similar completed actions
  - FACT: Typically past more similar assumed
    - even ventures may appear entirely different

Reference Class Forecasting (adapted from http://www.slideshare.net/assocpm/a-masterclass-in-risk)

- Best predictor of performance is actual performance of implemented comparable projects (Nobel Prize Economics 2002)
- Provide an "outside view" focus on outcomes of analogous projects
  - Attempts to force the outside view and eliminate optimism and misrepresentation
- Choose relevant "reference class" completed analogous projects
- Compute probability distribution
- Compare range of new projects to completed projects

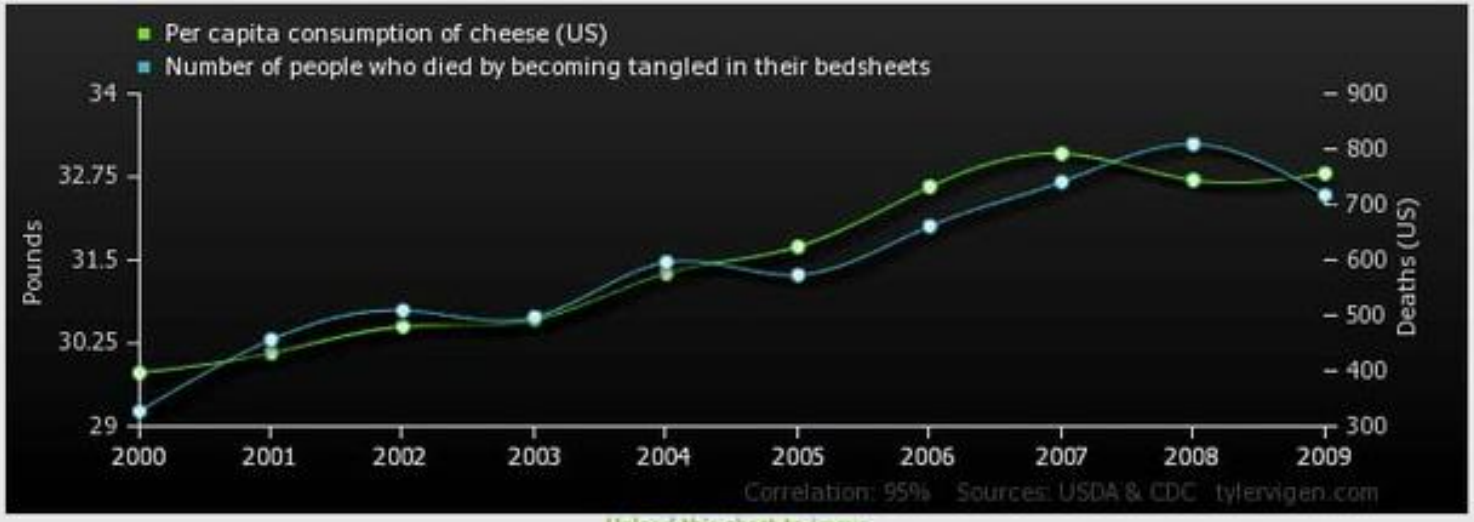
Correlation Doesn't Always Mean Causation (Source: www.memolition.com)

SEER by GALORATH

Per capita consumption of cheese (US)

correlates with

Number of people who died by becoming tangled in their bedsheets



	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Per capita consumption of cheese (US) Pounds (USDA)	29.8	30.1	30.5	30.6	31.3	31.7	32.6	33.1	32.7	32.8
Number of people who died by becoming tangled in their bedsheets Deaths (US) (CDC)	327	456	509	497	596	573	661	741	809	717

Correlation: 0.947091

Adding Reality to Estimates – Example – 2 (Source SEI)

Step	Expected
1	30
2	50
3	80
4	50
5	90
6	25
7	35
8	45
9	70
10	25
	500

What would you forecast the schedule duration to be now?

Example Bias Mitigation Using Multiple Sources

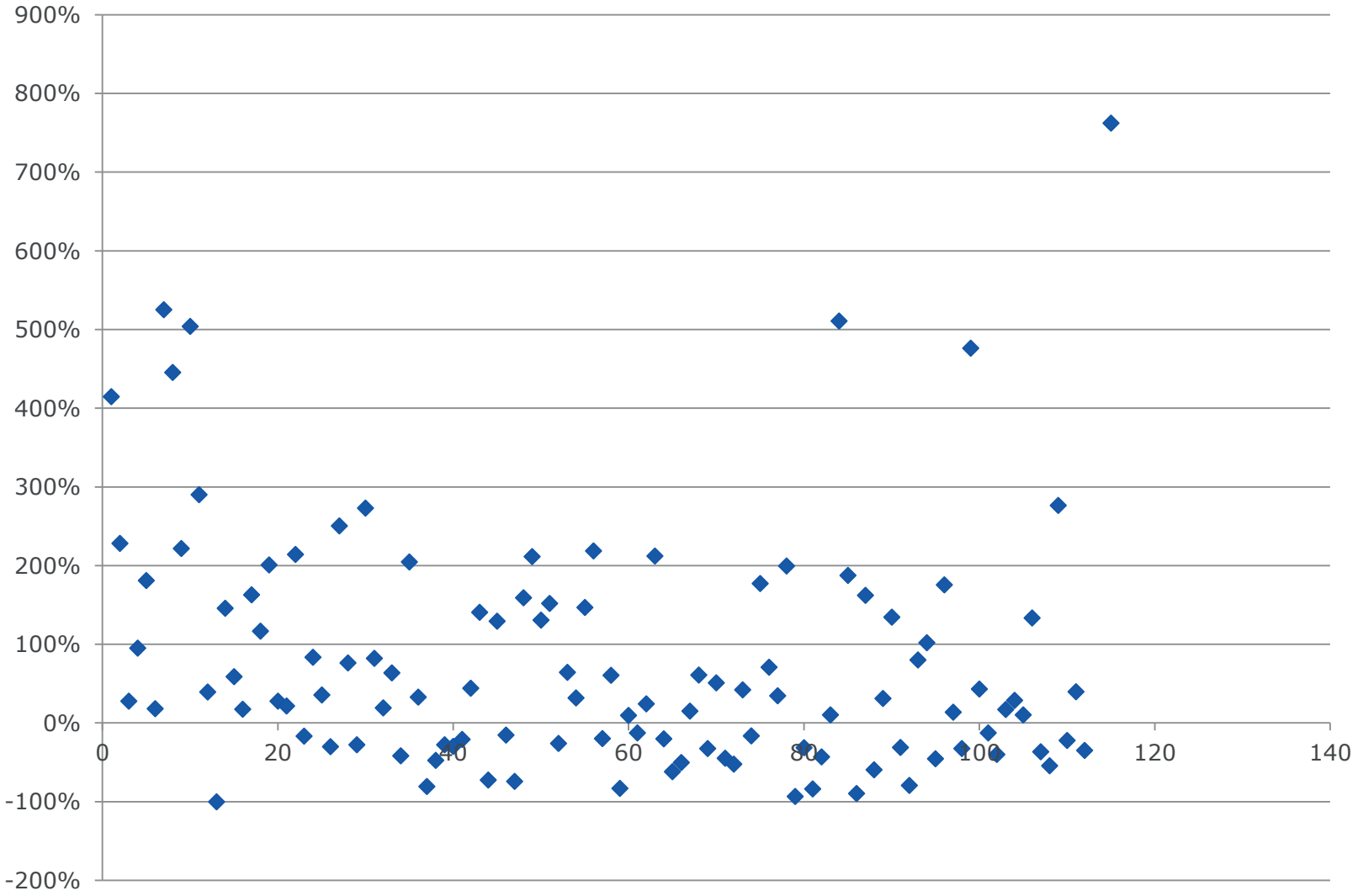
Evaluate All Sources of Software Size…

Total Size Estimates	Least	Likely	Most
Expert Judgement	12000	15500	17000
Relevant Range by Analogy	19850	24750	32540
Sizing Database	8000	32000	46000
Functional Analysis	19680	27540	35400
SEER-EstimateByCompare	15450	22650	29850
Delphi Analysis	16788	19750	22713
Estimate Range	**12000**	**22650**	**46000**

Estimate Independently then show table to minimize anchoring and other bias

SRDR v1 Estimate New SLOC vs Actual (Note:

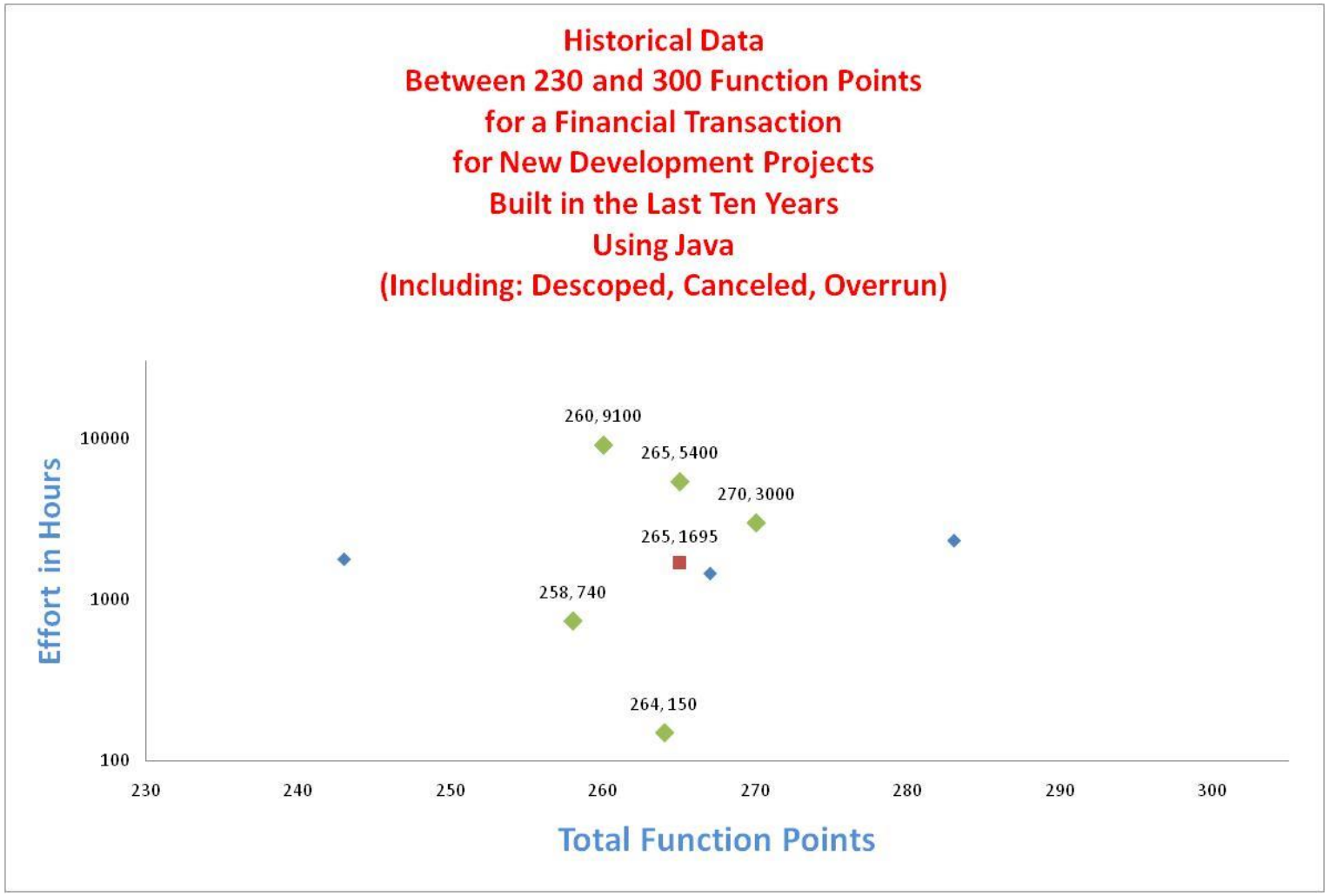
HUGE outliers removed to make the graph more readable)



Gross underestimation of software size versus actual

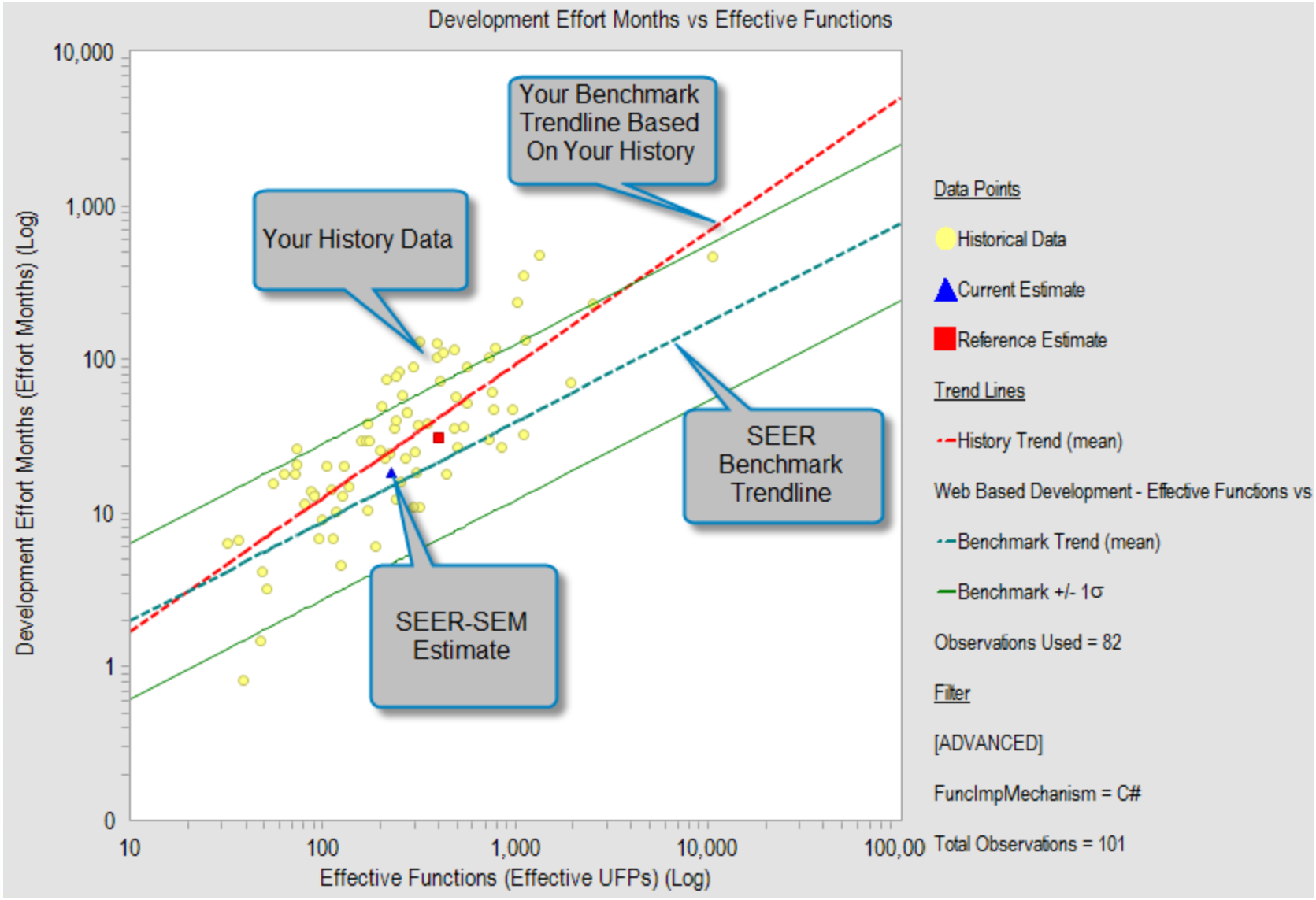
Fallacy of Silent Evidence

What about what we don't know?

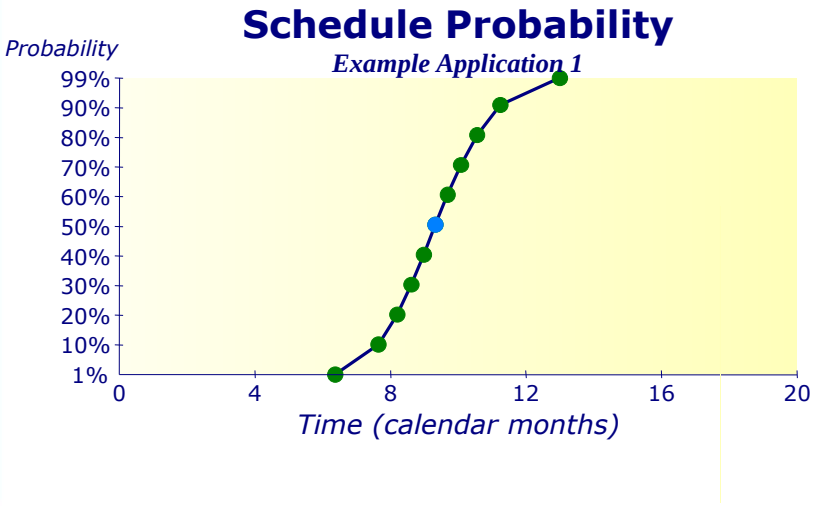


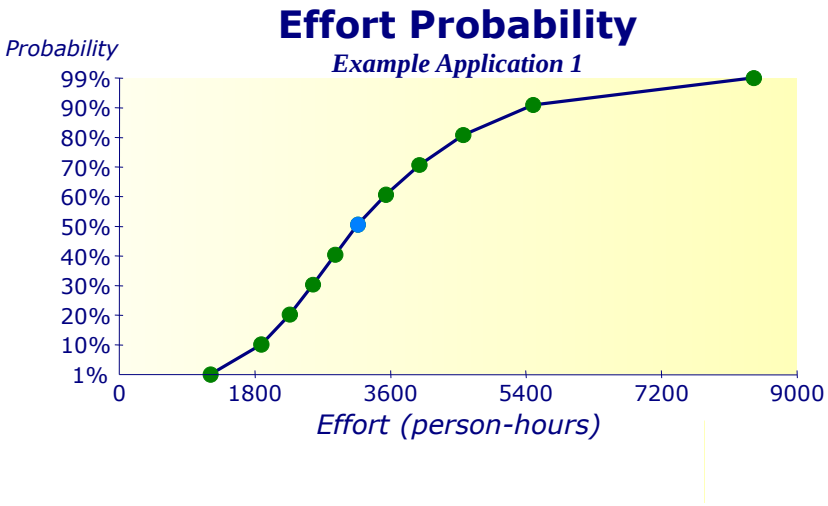
How confident would you feel if the Silent Evidence was visible?

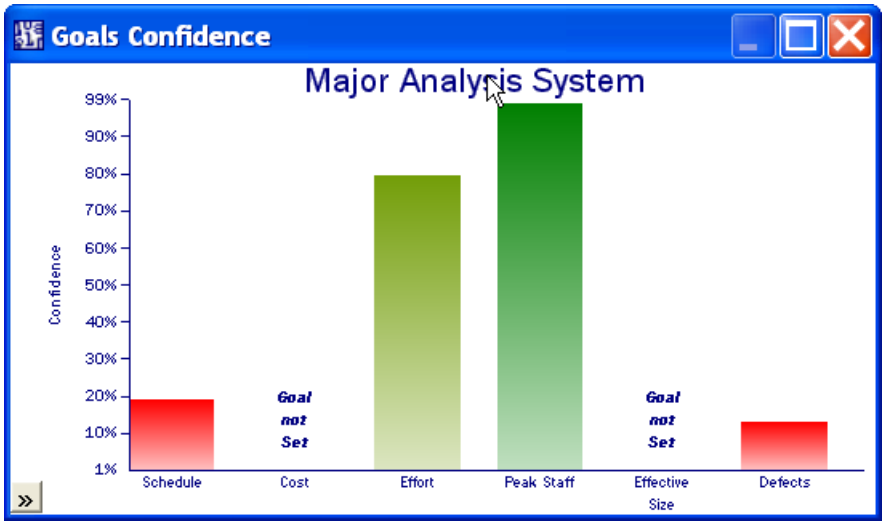
Example: Parametric Estimate Compared With History

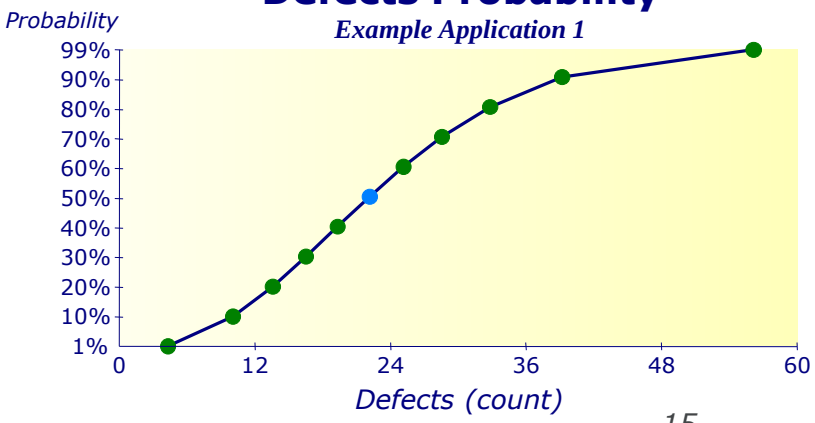


Understand Project Risks Include Them In Planning Decisions (Example SEER-SEM Outputs)

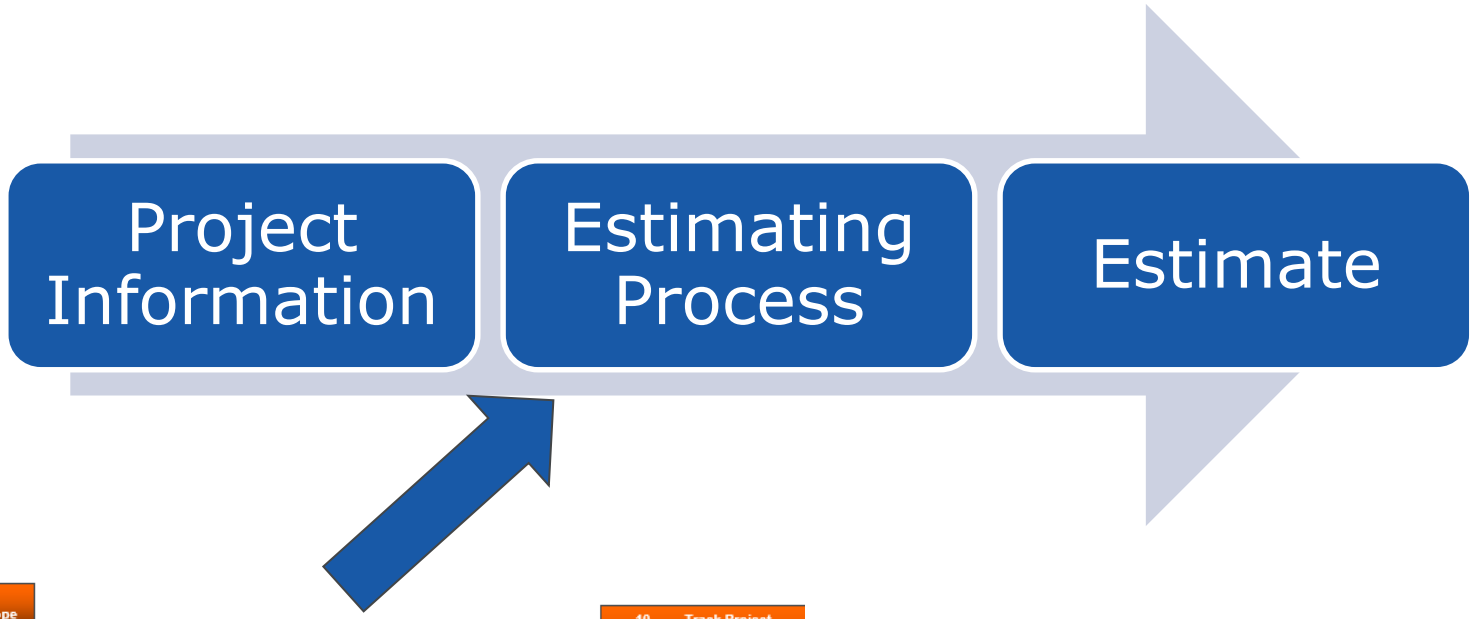


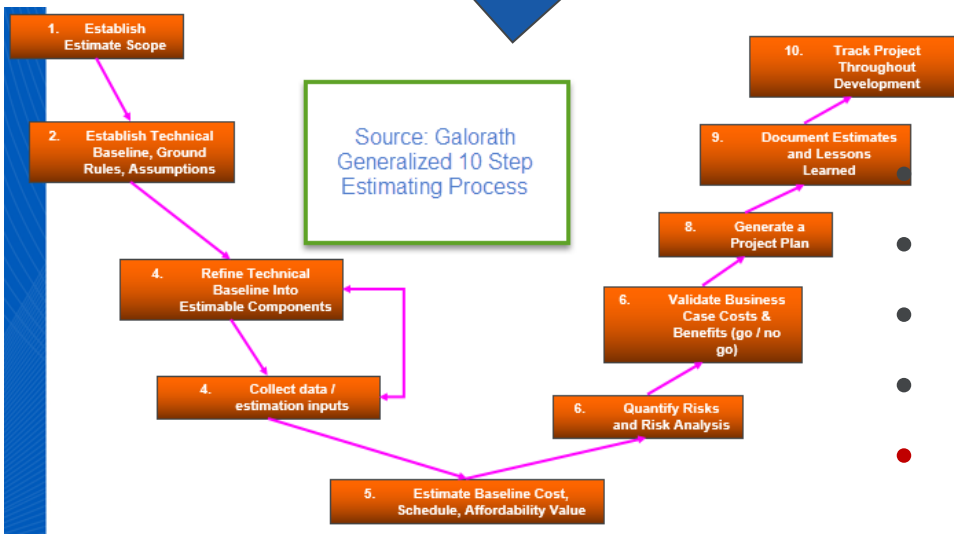






Estimating Process Should Help Mitigate Bias (Adapted from Andy Prince)





Process Provides

- Traceability
- Repeatability
- Best Practices
- Analytical Mindset
- STEPS TO MITIGATE BIAS

Anchoring Experiment: Anchoring Biases Estimates (Source: myweb.liu.edu/~uroy/eco23psy23/ppt/04-anchoring.pptx)

1. Subject witnesses the number that comes up when a wheel of fortune is spun
2. Is asked whether the number of African countries in the U.N. is greater than or less than the number on the wheel of fortune
3. Is asked to guess the number of African countries in the U.N.

Result: those who got higher numbers on the wheel of fortune guessed bigger numbers in Step 3

If given a number that biases estimates

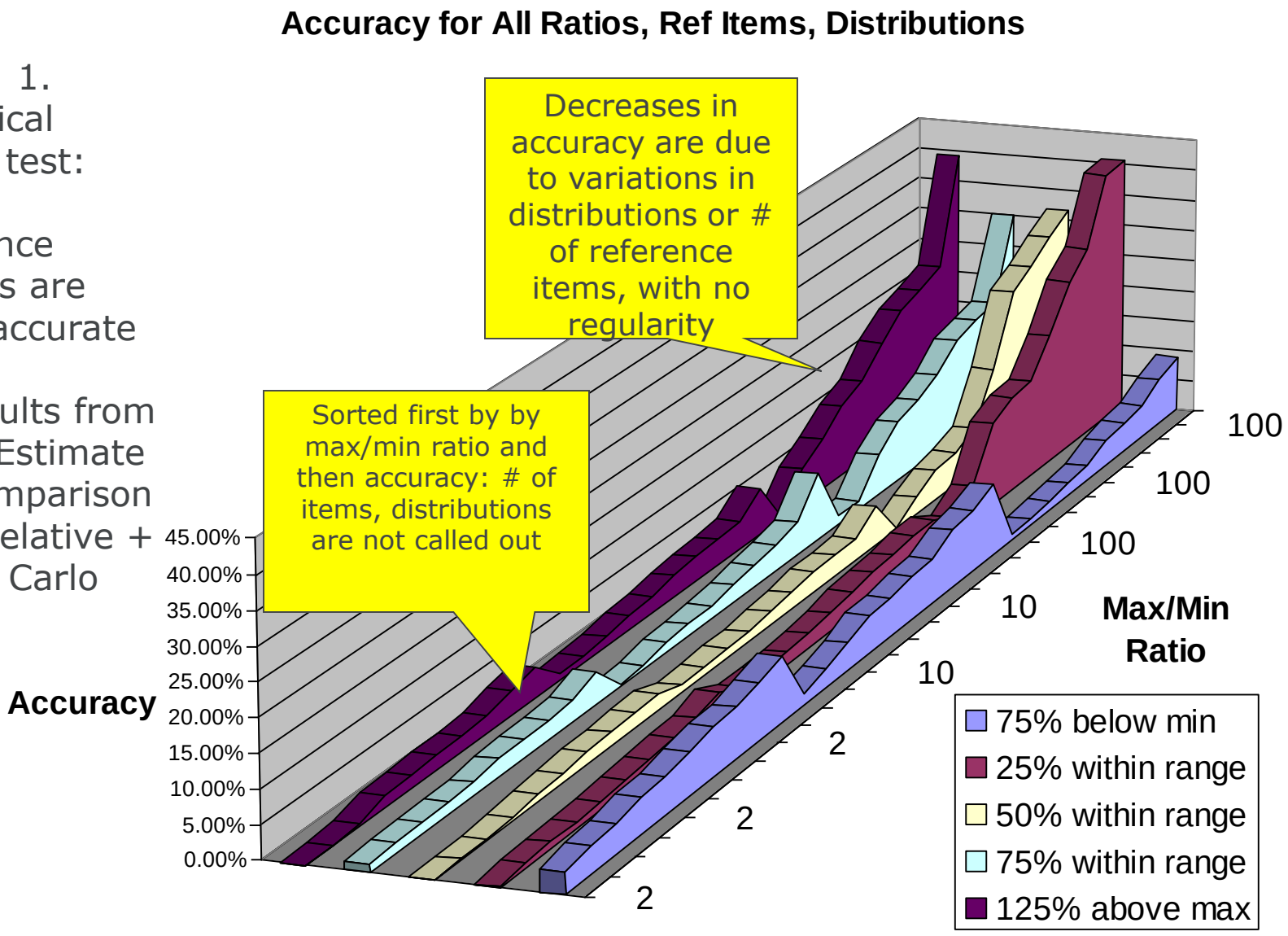
Anchoring

How we choose by comparing with a nearby reference point

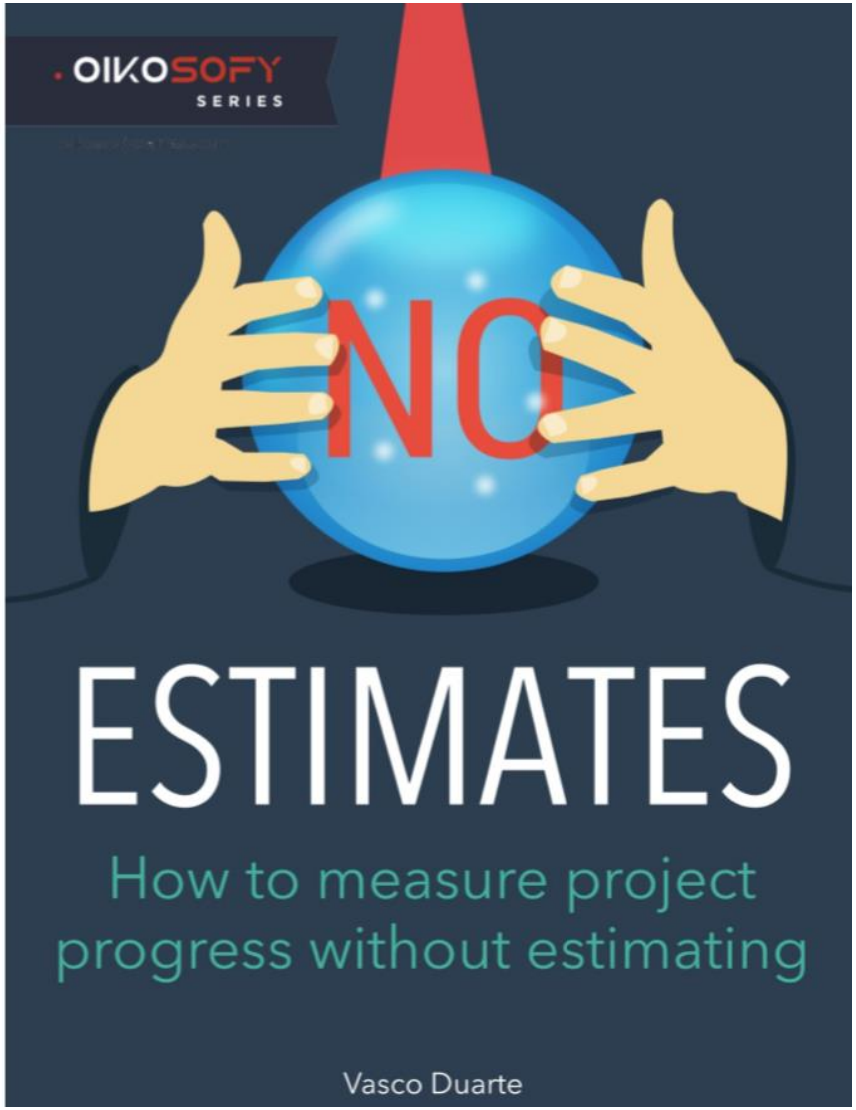
AHP Type Relative Analysis Can Be Within 10% of Actuals

Notes: 1. statistical stress test: Viable reference choices are most accurate

2. Results from SEER Estimate By Comparison Uses relative + Monte Carlo

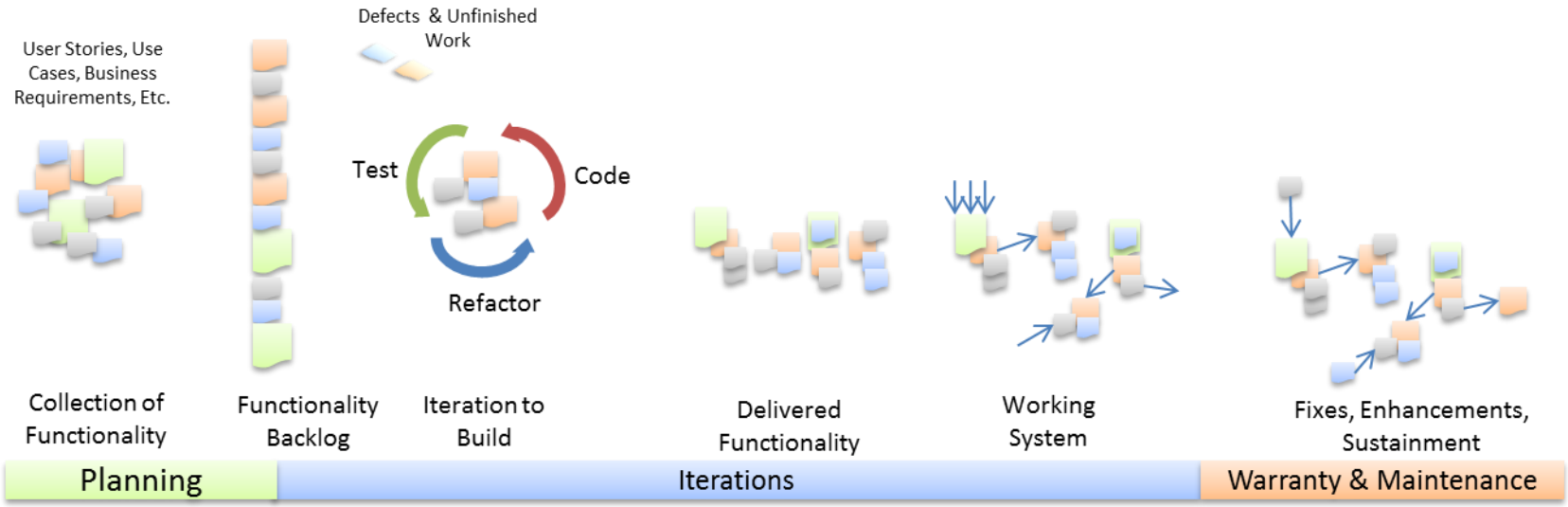


Add In The Agile Bashing of Estimating For a Full View



The Agile "Life Cycle"
(Scrum Example)

- Focus is on what features can be delivered per iteration
- Not fully defined what functionality will be delivered at the end?

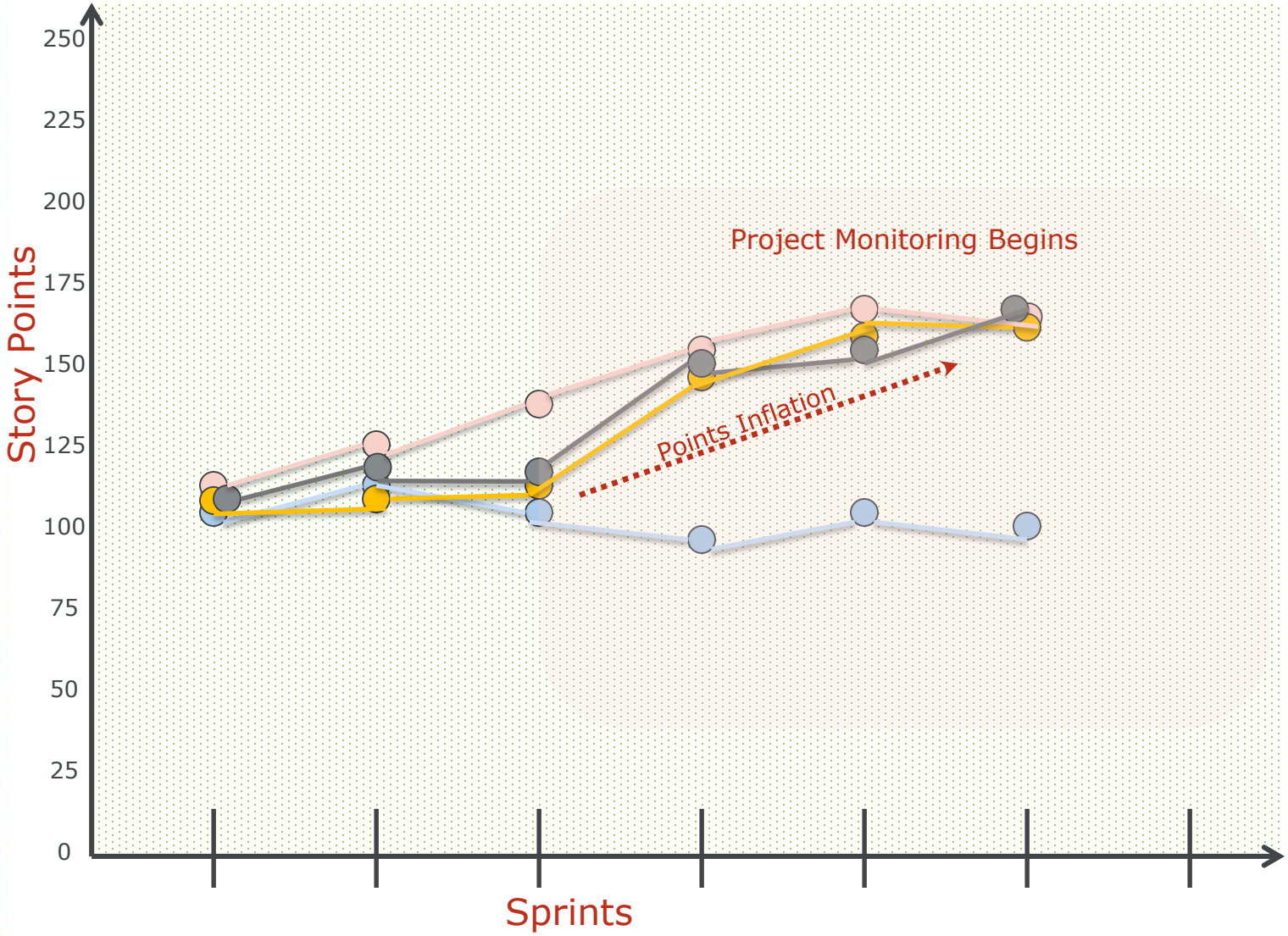


- Iterations are often called a "Sprint"

Root Causes Of Bad Estimates & Bias In Agile Projects As An Example

- Team not really doing Agile
  - Everyone seems to have their own "hybrid" which is code for management controls
- Immature process
  - No one with previous experience, i.e.: no Scrum Master
  - No training in the process being used
- Management gets in the way
  - Micromanage the burn down chart
  - Want to use velocity as productivity
  - Assume Ideal Days = Capacity Days
- Bad Story Counting
  - Trying to use counts across teams
  - Using historical story point counts for new work

Inflation in Story Point Productivity



Key Points

Without care estimates are usually biased (even with experts)

Tempering with an "outside view" can mitigate some bias

Estimates can be better, squelching bias & strategic mis-estimation… Parametrics help.

estimate

estimate • analyze • plan • control

Backup slides

Confirmation Bias (Source: Beinghuman.org)

- Give more weight to information that confirms what we already believe
  - Automatic unconscious way our brains process information
  - Selectively remember information that confirms what we already think
  - When we approach new information, we interpret it in a biased way
  - Spin news story so it vindicates their own beliefs?
- We subconsciously only pay attention to the information that confirms what is already known

You would think this would help ensure viable estimates but… Its what we believe, not necessarily what is reality

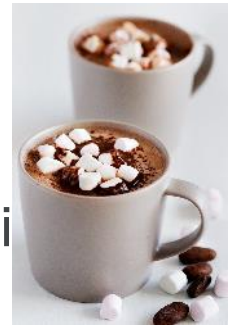
Negativity Bias (Being Human.org)

- Unconsciously pay give more weight to negative experiences than positive ones
- Brains react powerfully to negative information than they do to positive information
- Daniel Kahneman explained:
- "The brains of humans and other animals contain a mechanism that is designed to give priority to bad news. By shaving a few hundredths of a second from the time needed to detect a predator, this circuit improves the animal's odds of living"
- More important for our ancestors to be able to avoid a threat quickly than to gain a reward

Again, this should yield viable estimates but is usually overridden

Loss Aversion Bias (Source BeingHuman.org)

- Tendency to strongly prefer avoiding a loss to receiving a gain
  - Explains making same irrational decisions over and over
- Kahneman: Experiment giving one third of the participants mugs, one third chocolates, and one third neither
  - Option of trading
    - 86 percent who started with mugs chose mugs
    - 10% who started with chocolate chose mugs
    - 50% who started with nothing chose mugs



- Throwing good money after bad (sunk cost fallacy) i perfect example of loss aversion
- To avoid feeling the loss we stick with our plan, hoping for a gain, even when that just leads to a bigger loss

Explains why it is so hard to kill a failing program

Affect Heuristic Bias (Source: Beinghuman.org)

- Involuntary response to a stimulus that speeds up the time it takes to process information
  - If we have pleasant feelings, we see benefits high and risks low, and vice versa
  - affect heuristic behaves as a first and fast response mechanism in decision-making
  - Helpful in life or death situations where time was of the absolute essence.
- **System 2** The analytic, rational system of the brain is relatively slow and requires effort
- **System 1** The experiential system is different—speedy, relying on emotional images and narratives that help us to estimate risk and benefit.

Hopefully estimates elicit system 2... But often are off the cuff via system 1

Thinking Fast & Thinking Slow (Source: Kahneman)

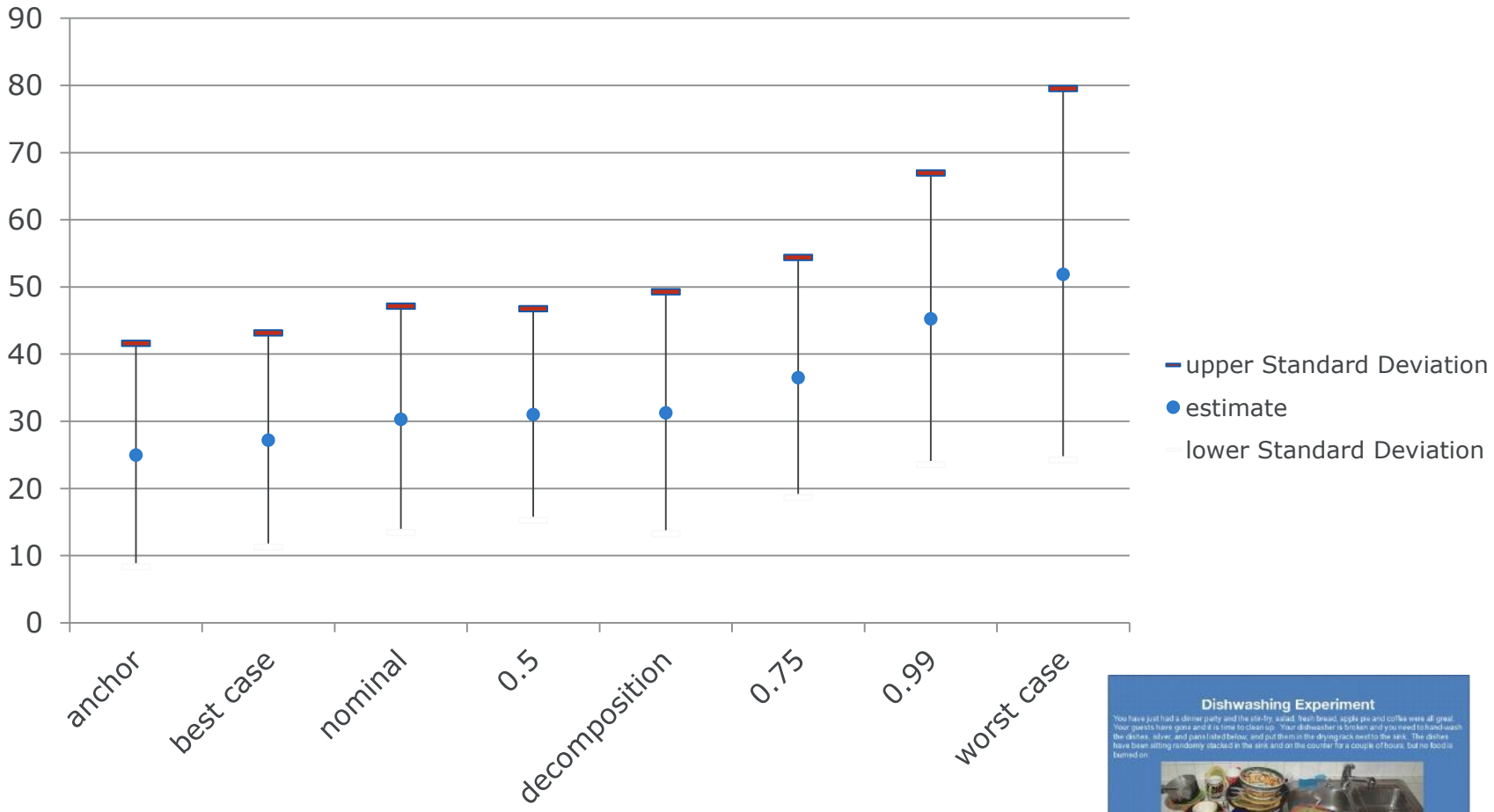
System 1: Thinking Fast	System 2: Thinking Slow
• Operates Automatically • No effort • Quick • No voluntary control	• Allocates attention to mental activities that demand it • Complex computations
• Coherent interpretation of what is going on	• Good at balancing probabilities but often indecisive
• Intuitive answers quickly	• Takes over when System 1 can't process the data • If the person is willing • Can correct or override System 1 if it determines intuition is wrong

Illusion of Control (Source: BeingHuman.org)

- Tendency to overestimate their influence over outcomes that they cannot affect
- Psychologist Ellen Langer Subjects given lottery tickets; either at random or allowed to choose their own
  - Had chance to trade tickets for others that had a higher chance of paying out.
  - Subjects who chose ticket were less likely to part with it than those who had a random ticket
  - Subjects felt their choice of ticket had some bearing on the outcome—demonstrating the illusion of control.
- Illusion of control especially strong in stressful and competitive situations, like gambling or financial trading or ESTIMATING

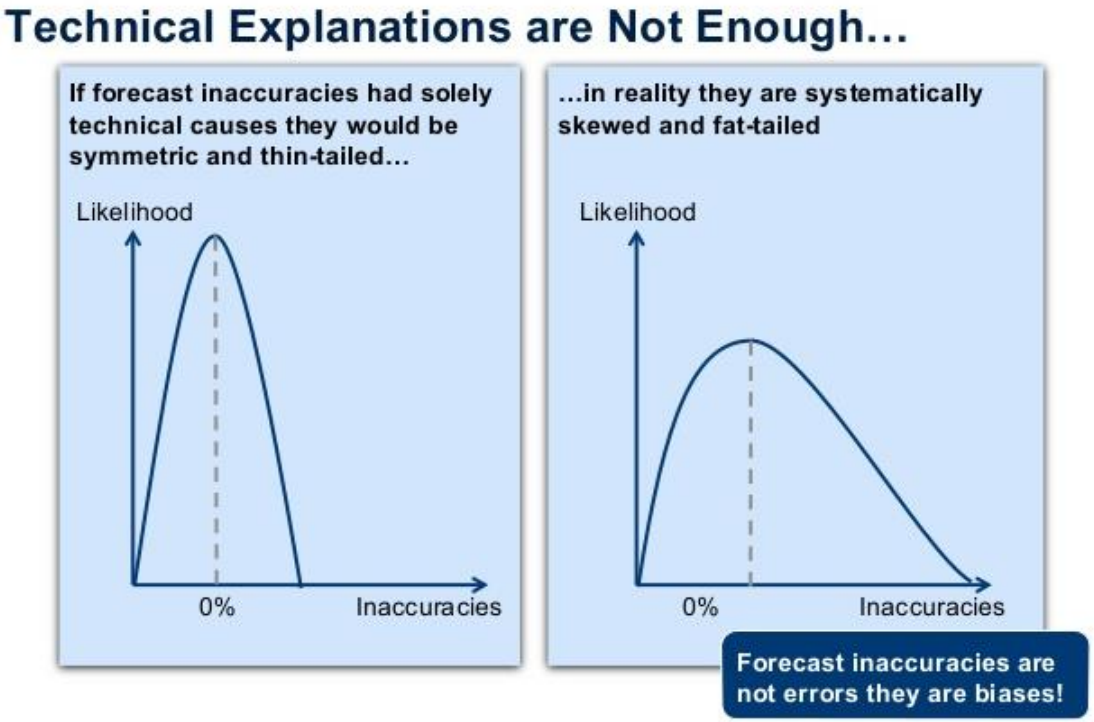
Illusion of control can lead bad decisions or irrational risks

Dishwashing Estimation Bias Study Summary (Source: JPL http://www.slideshare.net/NASAPMC/arthurchmielewski)



Explanations for Poor Estimating
(Adapted From Source Master Class on Risk, Flybjerg, 2013)

1. Technical: Inadequate data & Models (Vanston)
2. Psychological: Planning Fallacy, Optimism Bias - causes belief that they are less at risks of negative events
3. Political / Economic: Strategic misrepresentation - tendency to underestimate even when experienced with similar tasks overrunning (Flyvberg)



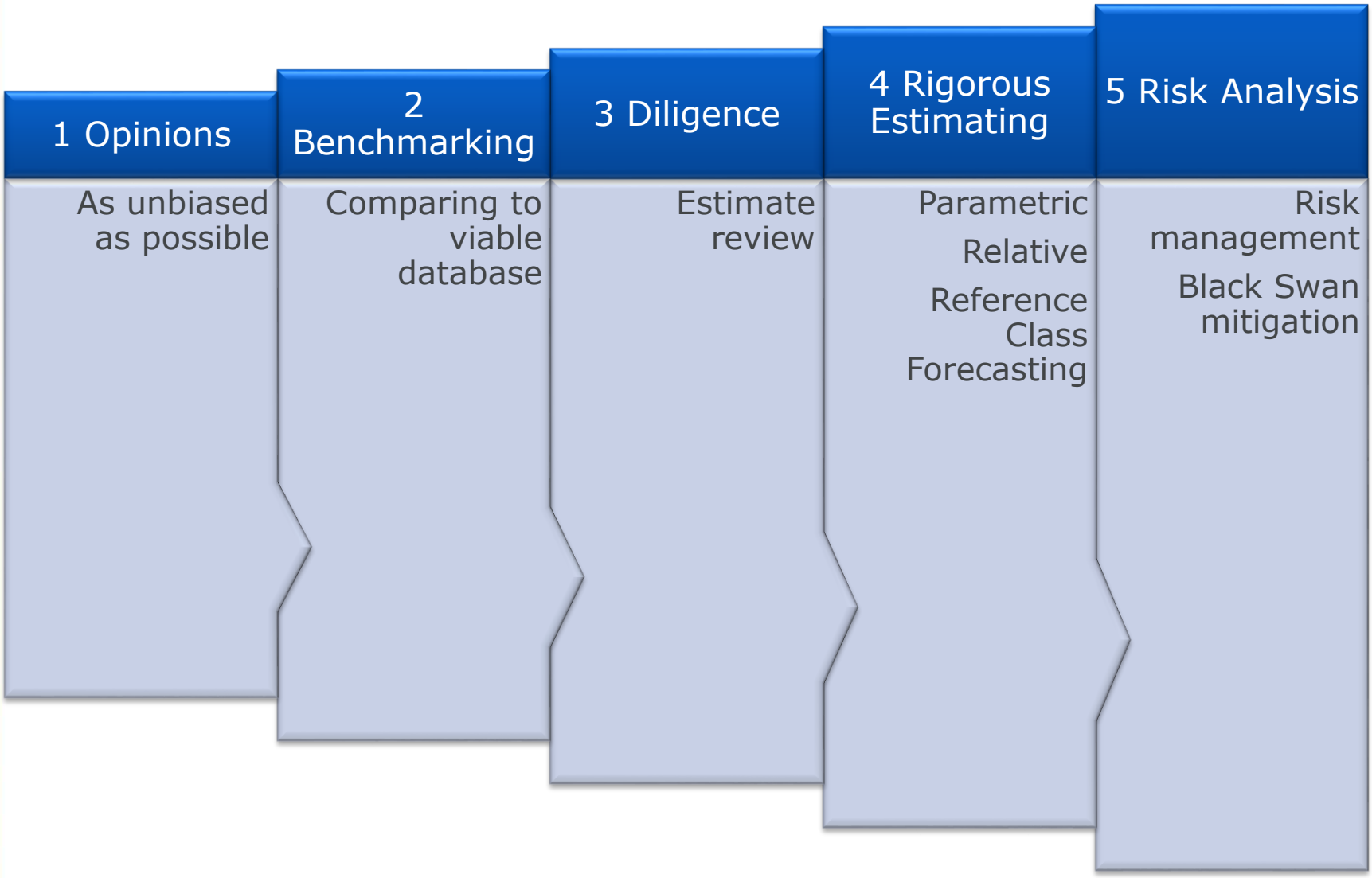
Draw Out Range By Obtaining 3 Estimates

- Optimistic value (s_{opt})
- Most likely value (s_m)
- Pessimistic value (s_{pess})

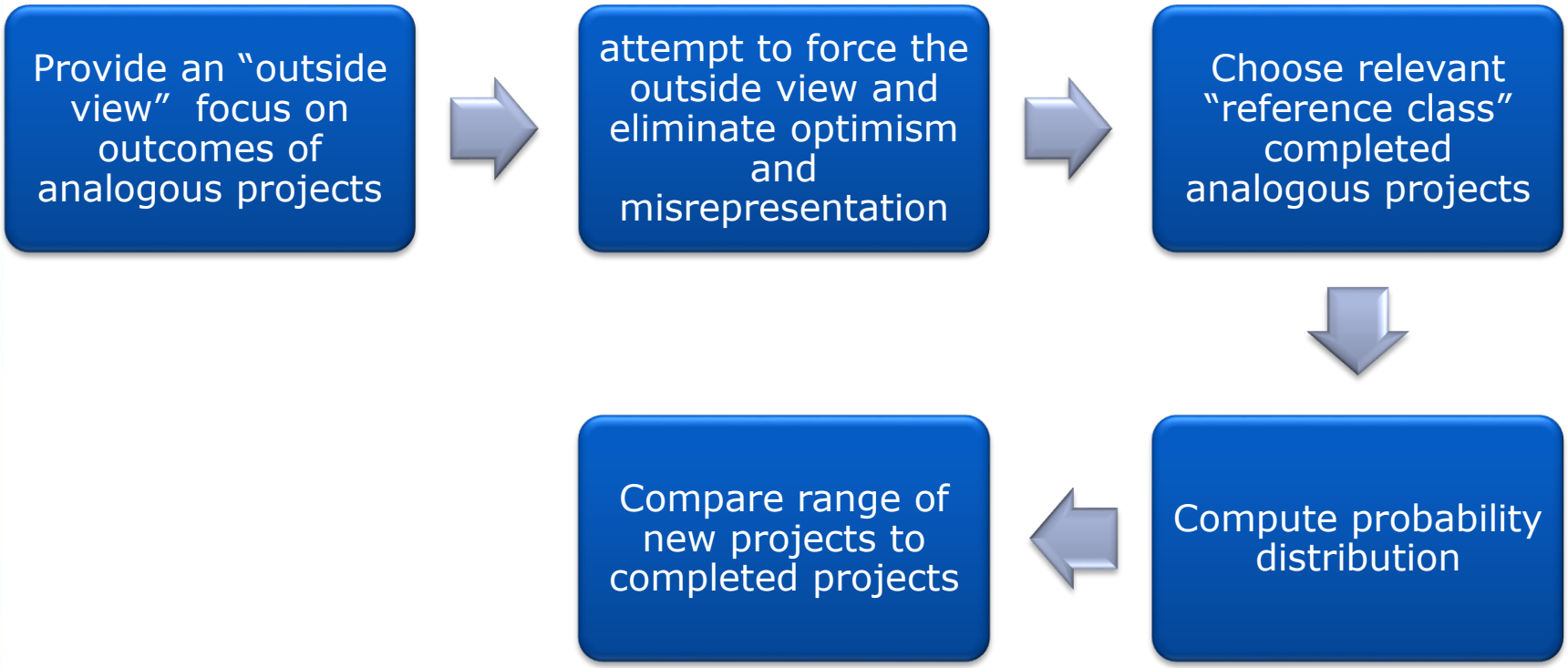
- Expected value (*EV*)

$$EV = \frac{(s_{opt} + 4s_m + s_{pess})}{6}$$

5 Levels of Risk Management (Adapted from Flyvbierg)



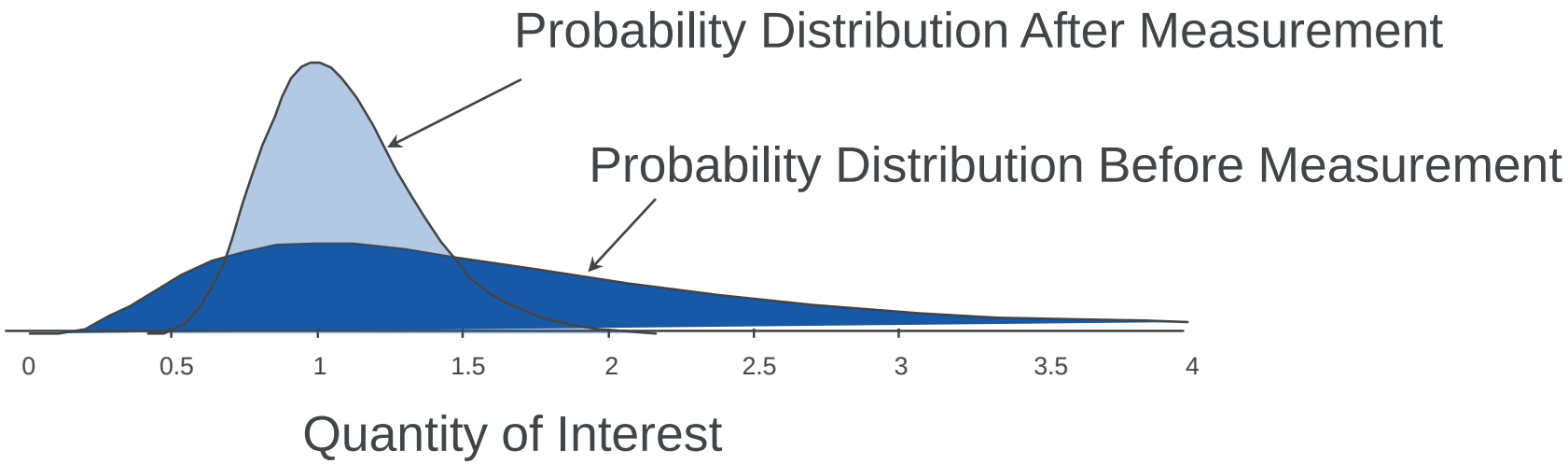
Reference Class Forecasting (adapted from http://www.slideshare.net/assocpm/a-masterclass-in-risk)



Best predictor of performance is actual performance of implemented comparable projects (Nobel Prize Economics 2002

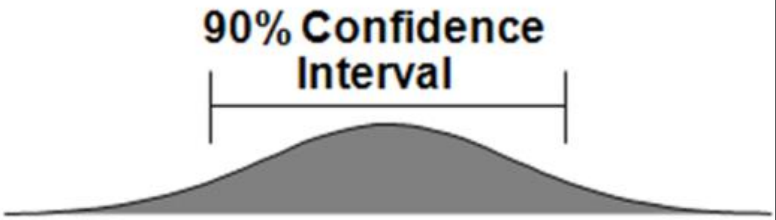
Hubbard: Measure To Reduce Uncertainty

- Perception that measurement is a point value is a key reason why many things are perceived as "immeasurable"
- Measurement: Quantitatively expressed reduction in uncertainty based on observation



Assumptions, Change Drivers & Expert Judgment Need Caution (Source: Hubbard)

- 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%

(AIE = Hubbard Generic Calibration Training)

Used with permission from Douglas Hubbard
Copyright HDR 2008 dwhubbard@hubbardresearch.com

Example - Pairwise Comparisons

- • Consider following criteria

Purchase Cost | Maintenance Cost | Gas Mileage

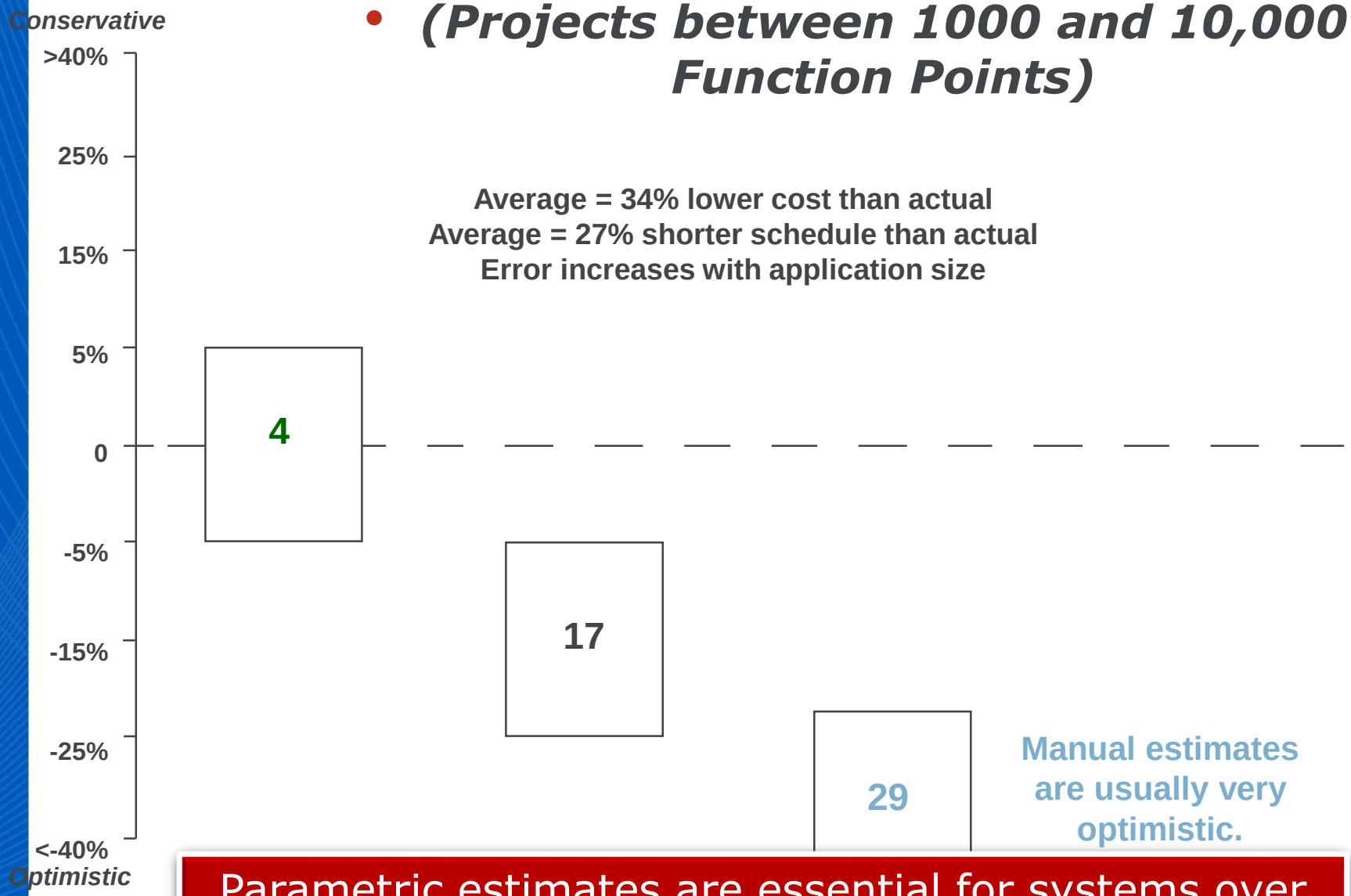
- **Want to find weights on these criteria**
- **AHP compares everything two at a time**

(1) Compare Purchase Cost **to** Maintenance Cost

– Which is more important?
Say purchase cost
– By how much? Say moderately → **3**

ACCURACY RANGES FOR 50 MANUAL ESTIMATES (Source: Capers Jones)

- ***(Projects between 1000 and 10,000 Function Points)***



Parametric estimates are essential for systems over 1,000 function points