FORCE 10 Jun 15

Biases and their value-destroying impacts

Steve Begg
Emeritus Professor, University of Adelaide
& DecisionsDecisions

Outline

- Decision-making and the brain
- Common biases and traps
- Learning from visual judgements & mitigation tips

I am happy to provide a PDF, including extra biases & examples

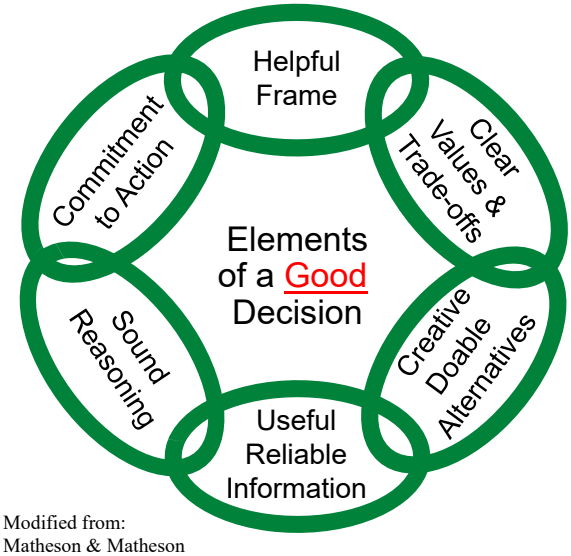
Decision-making context

The only control an organization has over its outcomes/future is the quality of its decisions – actual outcomes depend on uncertainties it cannot control: decisions of others, "nature", chance, ...

- "Natural talent" is overrated: expertise, whether in law, engineering, sport, or decision-making, is nearly always made not born
- Choosing wisely is a life *skill*. Cognitive Scientists have shown that good decision-making capabilities are not simply "wired-in" through evolution
  - but very few are taught how to make good decisions!

But what is a "good" decision – and how can we know that it is at the time we make it?

Decision Quality (DQ): An over-arching framework for making good decisions



Modified from:
Matheson & Matheson

A link is strong enough when doing more is a waste of time or money (the cost is more than the added value) - pragmatically, the best option is unlikely to change

The decision is only as good/strong as the weakest link!

Two major fields of study and application in decision-making

Prescriptive: Decision Analysis		Descriptive: Psychology
• How people *should* choose to maximize value, however they define it		• Observation of how people *actually* choose and make judgements
• A structured, explicit methodology supported by various "tools" - DA & DQ	≠	• Shows people often use simple rules or intuition, which lead to bias & error
• … best choice in terms of "delivering" what the decision-maker wants		• … choices that are often unlikely to result in what they really want!

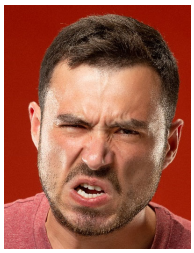
The descriptive side can protect you from
- biases (yours/other's) that unwittingly undermine getting more of what you want
- people who exploit psychological factors to "sell" you a project / product / idea that is in their interest – not necessarily yours or your organization's!

Type (System) 1 and Type (System) 2 "Thinking" (Stanovich)

Type 1	Type 2
– Fast, automatic, implicit	– Slow, conscious, explicit – but lazy
– Always on	– Needs to be "switched on"
– Effortless - based on associative memory	– Effortful – based on conscious reasoning
– Doesn't interfere with other ongoing cognitive tasks	– Interferes with other conscious thoughts (we can't multi-"task"!)
– Fallible – particularly when there is novelty, complexity or uncertainty	– Accurate – when in control it can correct System 1 errors
– The "gut-feel" / intuitive brain	– The "deliberative" brain

Type (System) 1 and Type (System) 2 "Thinking"

Type 1



Freepik.com

Type 2

$$\frac{23 \times 7}{9} = ?$$

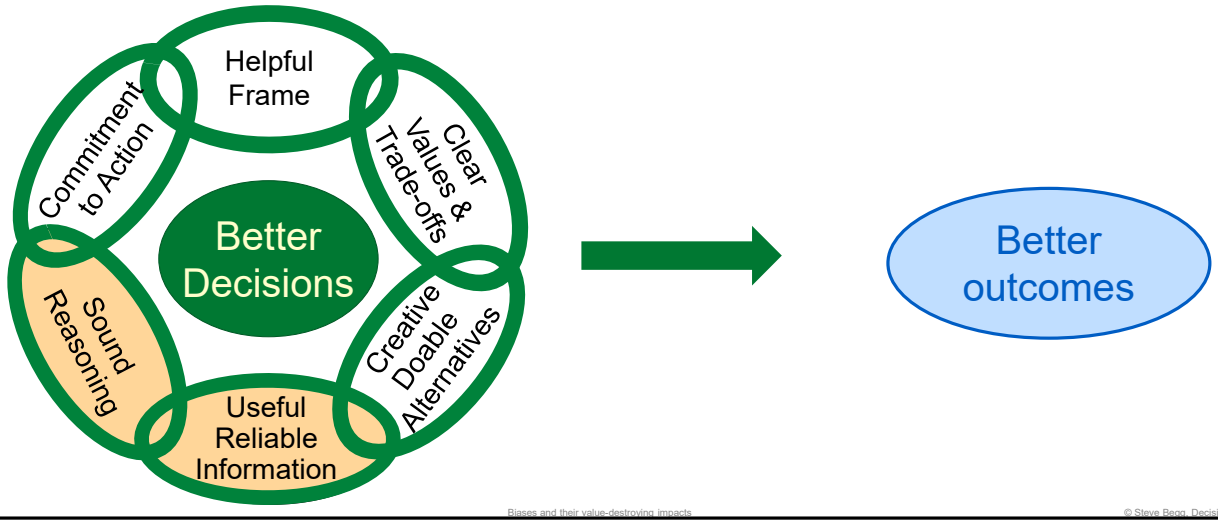
Q A bottle of wine and a glass cost $27 in total. The wine costs $22 more than the glass. How much does the glass cost?

- common intuitive answer: $5
- correct answer: $2.5 (bottle is $24.5, $22 more than the glass)

Evolutionarily, humans are cognitive misers and default to Type 1 – which can yield compelling, but wrong "answers"/decisions

Why does this matter ?

Our "caveman" brains are not (yet!) evolutionarily-endowed with unaided, sound thinking under uncertainty, complexity and novelty



Outline

- Decision-making and the brain
- Common biases and traps
  - simple judgements and chance assignments
  - "reasoning" under uncertainty
- Learning from visual judgements & mitigation tips

Avoiding confusion: Biases vs Preferences

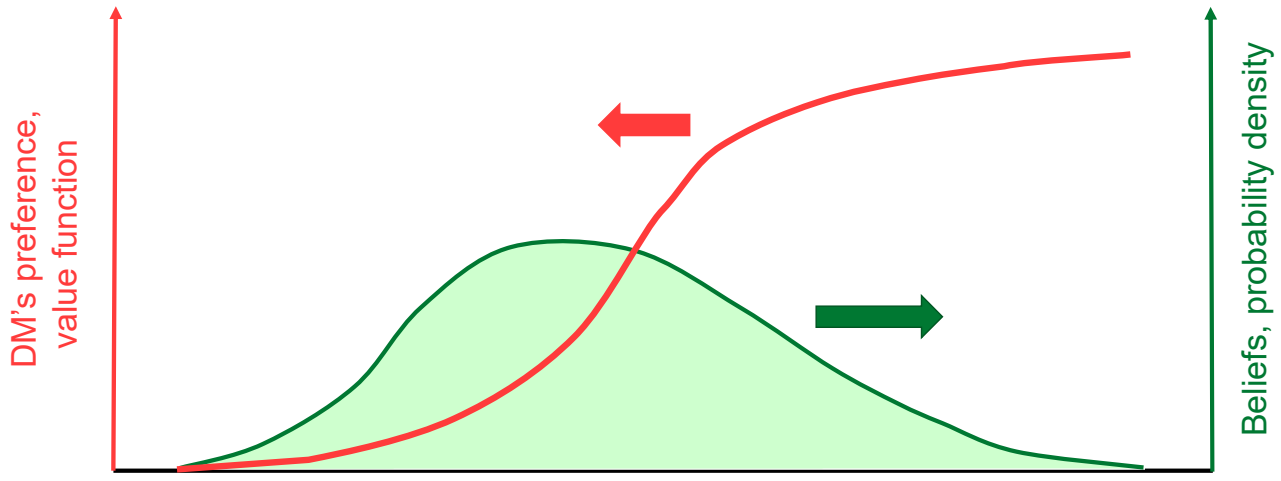
- Cognitive Biases: subconscious systematic "reasoning"/neural-processing errors when coming to judgements or beliefs -> inconsistent with data / evidence / reality
- Motivational Biases: we would like something to be true (or false) - so assign a higher (or lower) degree of belief to it occurring
- "Structural" Biases: built-in to calculations/models/algorithms

If our beliefs or reasoning are biased, we can mistakenly order the desirability of our options - and fail to choose the best one (for us)!

- Preferences: desirability of "states of the world", based on our values (and beliefs if we are uncertain)
  - "I prefer red wine to white"; "80% chance of oil is better than 20% chance of gas"

Preferring one thing to another is not being biased !
(high DQ requires us to know and use the DM's preferences)

* Values (preference) vs Beliefs (probability) - Graphically



Possible outcomes of a (continuous) uncertain quantity

Focus and Data

Wikipedia has 170+ "biases" – a range of judgement and decision-making behaviours / effects / traps / errors and genuine biases

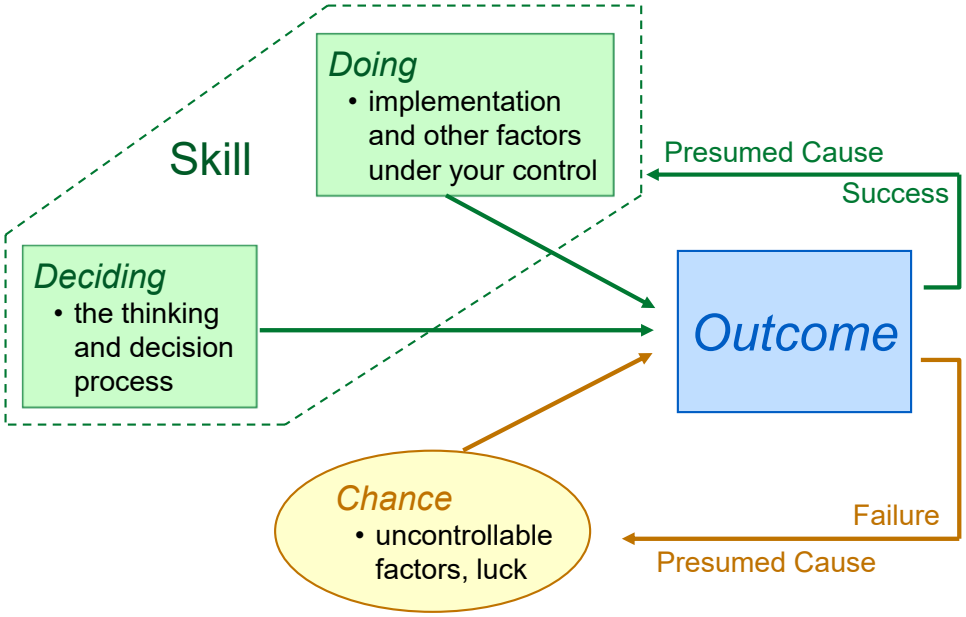
My focus:

- Biases mainly due to uncertainty, novelty & complexity
- Chosen based on personal view of most damaging & frequent

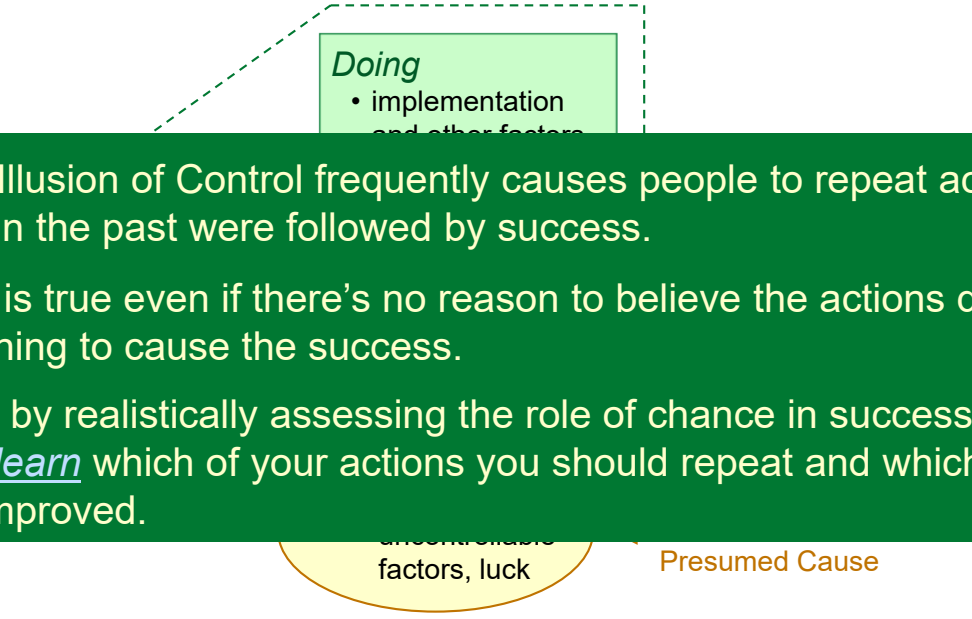
Data for examples

- Subjects: mainly O&G and mining professionals (many whose job is dealing with uncertainty)
- Sources: funded research projects and company training courses

* The Illusion of Control and its problems



The Illusion of Control and its problems



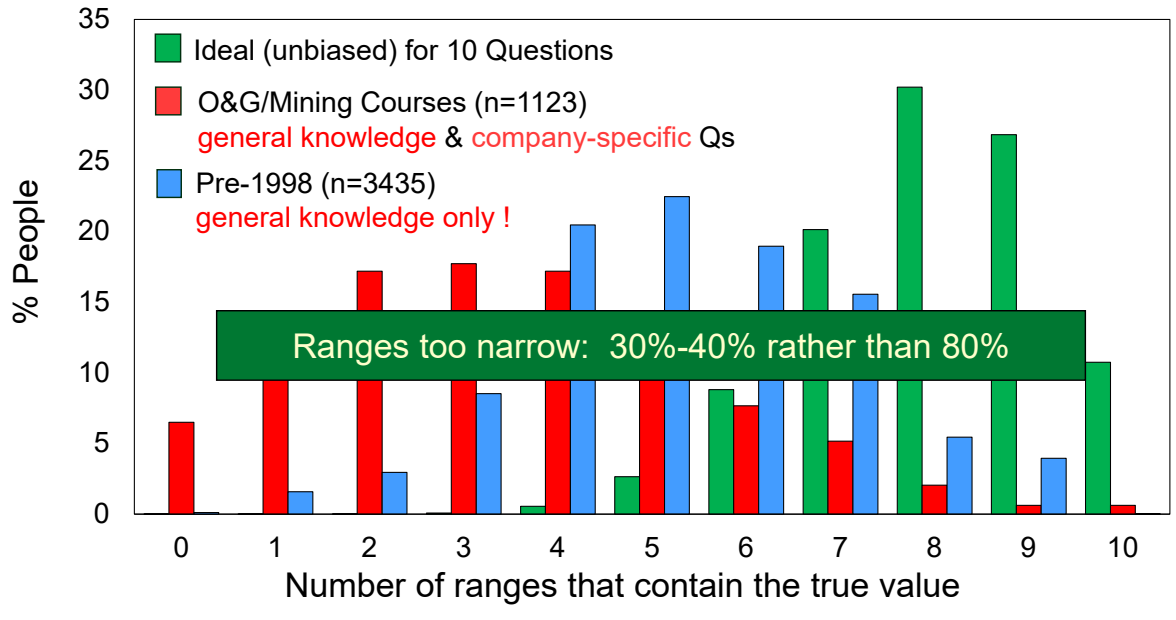
- The Illusion of Control frequently causes people to repeat actions that in the past were followed by success.
- This is true even if there's no reason to believe the actions did anything to cause the success.
- Only by realistically assessing the role of chance in successes can you *learn* which of your actions you should repeat and which could be improved.

Uncertainty range estimation exercise: 10 questions

	10%	Lower (P10)	80%	Upper (P90)	10%
Question 1		_____		_____	
Question 2		_____		_____	
Question 3		_____		_____	
...		...		...	
Question 10		_____		_____	
Eg What is the height of Chicago's tallest building, m?		250		450	

- If respondents are unbiased (their ranges are consistent with their state of information), on average, 8 out of 10 ranges should contain the true value

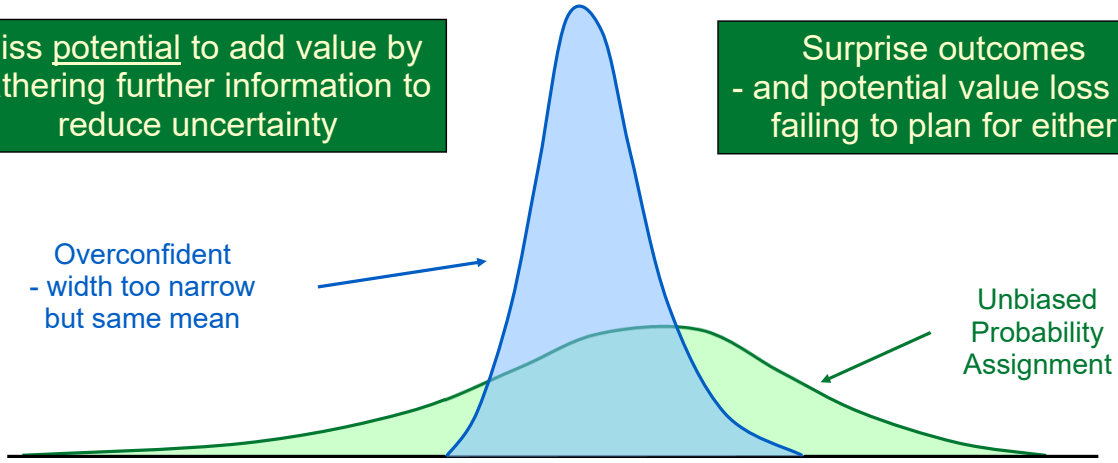
Overconfidence (over-precision): Industry professionals



Overconfidence and its impacts

Miss potential to add value by gathering further information to reduce uncertainty

Surprise outcomes - and potential value loss by failing to plan for either



Greater chance of surprise undesired low outcomes (risks)

Greater chance of surprise desirable high outcomes (opportunities)

Overconfidence: economic impact on project NPV

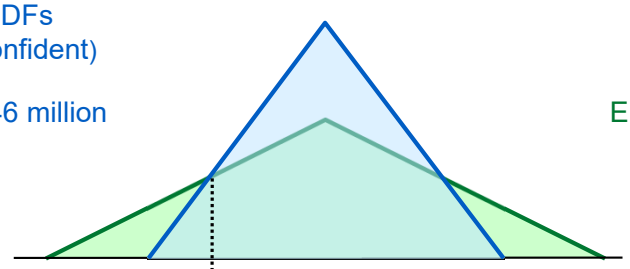
- Investment decision - development plan depends on the hydrocarbon volume estimates*
  - uncertainty in volume calculation inputs assumed to be overconfident (EV the same)

Biased PDFs
(30% overconfident)

E[NPV] = $246 million

Unbiased PDF
(0% overconfident)

E[NPV] = -$10 million!!

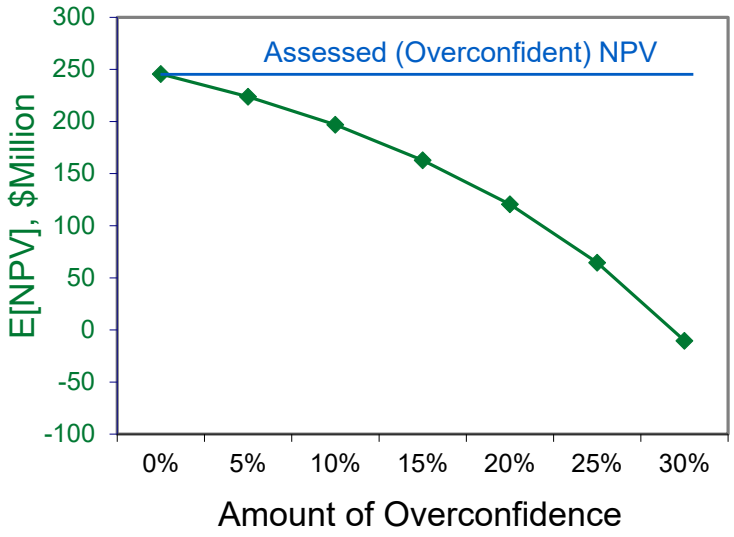


Input to volume calculation

P10 of biased (overconfident) PDF = P25 of unbiased ("true") PDF

*underlying model is non-linear w.r.t. inputs

Overconfidence: Economic impact on project* NPV



- If inputs weren't overconfident the E[NPV] would be $246m
- Inputs "corrected" for different levels of over-confidence and the "true" E[NPV] computed
- If the inputs were 30% overconfident (typical) the true E[NPV] is -$10m!

*underlying model is non-linear wrt inputs

* Over-placement: Positive Illusions

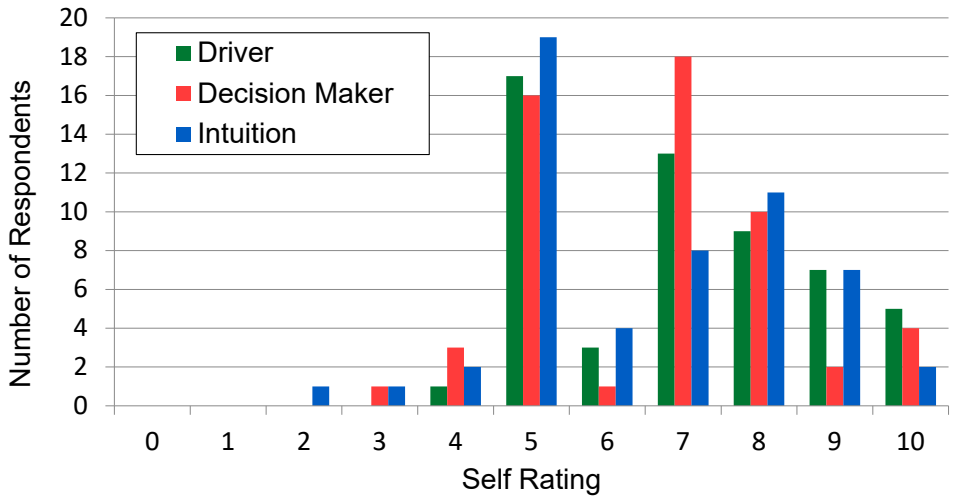
- 82% of people say they are in the top 30% of safe drivers;
- 86% of MBA students say they are better looking than their classmates;
- Doctors consistently overestimate their ability to detect certain diseases;
- 81% of new business owners think their business has at least 70% chance of success, but only 39% think that a business just like theirs would be likely to succeed.

When assessing their position in a distribution of peers on almost any positive trait, 90% of people say they are in the top half!

Source: Russo & Schoemaker

* Positive Illusions: Industry professionals

- Compared to your peers, on a scale of 1 to 10, where 5 is average, how good a "xxxxxxxx" are you?



* Positive Illusions

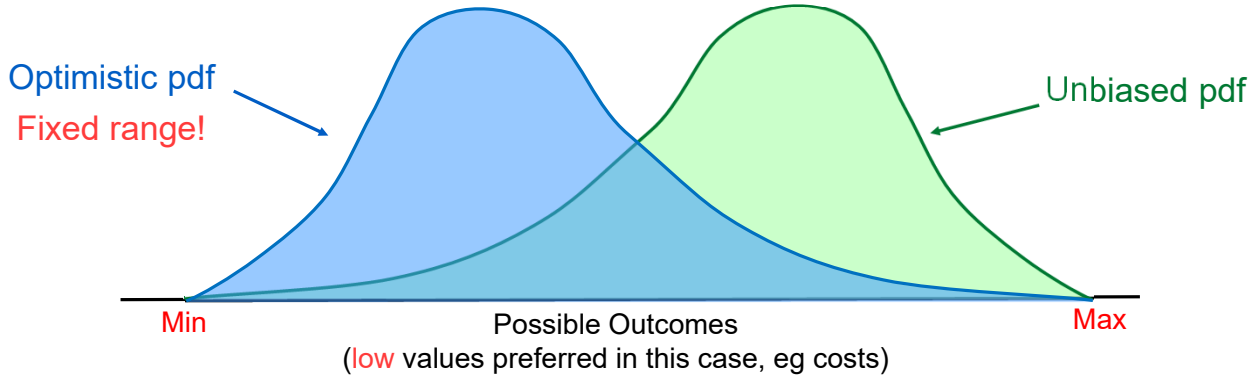
- Most behavioral scientists question whether positive illusions are really good for us.
  - it can lead them to believe that they are less at risk of experiencing a negative event compared to others in the same situation
- It may be difficult to disconfirm the positive belief.
  - eg: it is easier for individuals to maintain the view that they are more honest than others than to claim that they are faster runners.
- It is harder to have positive illusions when they are inconsistent with easily available objective data.
  - eg: people rate themselves more highly on the overall dimension of being environmentally friendly than on specific behaviours such as recycling, reusing paper, or turning off lights.

* Types of Overconfidence

- Over-precision – cognitive: what we commonly call Over-confidence. Placing too narrow confidence intervals (ie probabilities), eg a P10-P90 range, or PDF in general, too narrow
  - having a higher probability, than is warranted by our information, that a statement/event is true (and a too low probability that its complement is false)
- Over-placement – cognitive: what we commonly call Positive Illusions – in a group of people an individual ranks their ability, relative to others, higher than is warranted
- Over-estimation – cognitive: an individual thinks their ability is greater than it actually is (eg Dunning-Kruger effect; ability to pick stocks; academic ability; managerial ability; decision-making ability …….)

Optimism

- Assigning higher probabilities to more-preferred outcomes and lower probabilities to less-preferred outcomes
  - substituting (or confusing) preference with probability



- Usually a motivational bias - there is an incentive/pressure to be optimistic (at worst, lie) - but could be cognitive, or a mix

Optimism: $ impacts

- Sydney Opera House
  - construction began in 1959

	Predicted	Actual
Cost	$7 million	$102 million
Completion	1963	1973

- Norwegian Continental Shelf estimates of PV loss due to optimism

Reason	Lost Value, 2017 M NOK
Cost overrun	213,219
Schedule overrun	61,389
Under production	199,389
Total PV loss	474,313

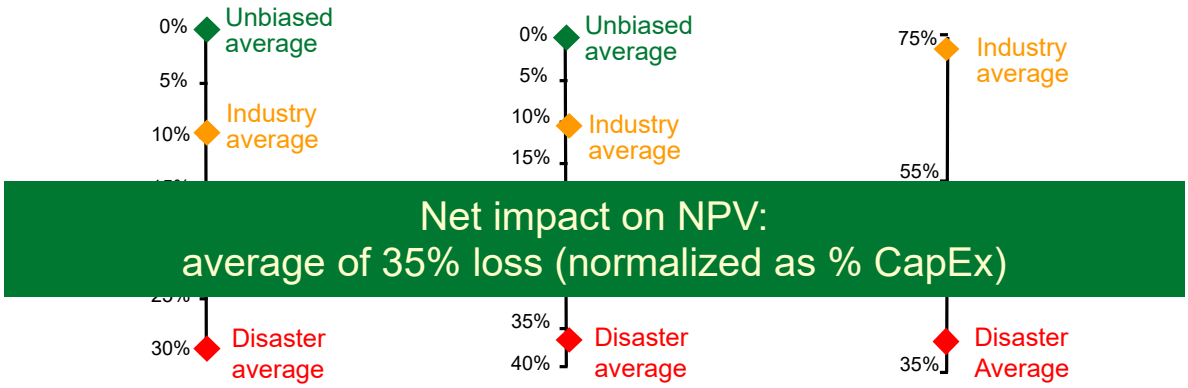
Mohus (2018)

~61.6 Billion US$

Optimism: $ impact on value (Merrow & Nandurdikar)

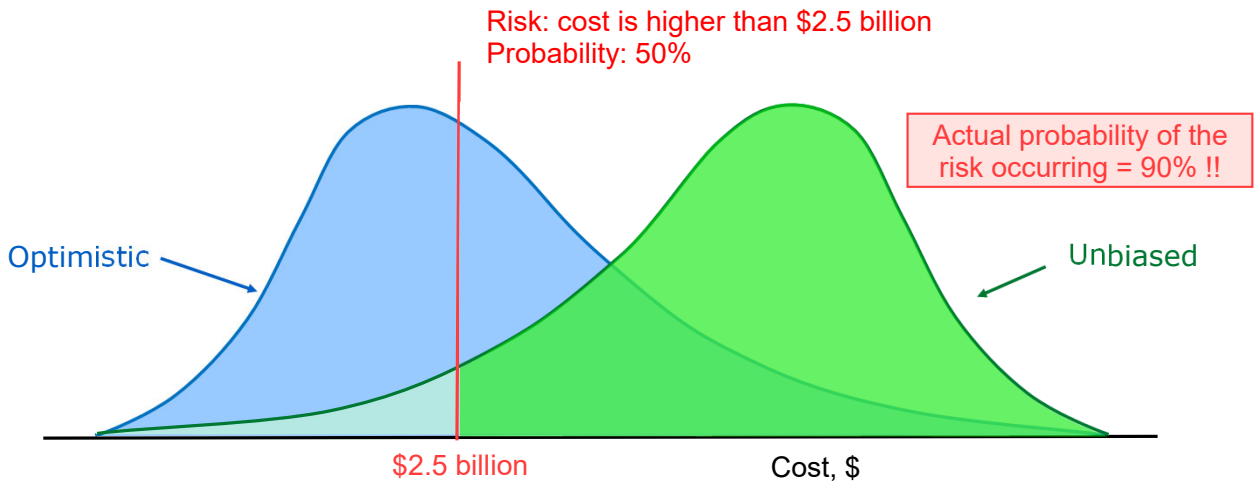
- IPA: Based on 1000 major offshore oil & gas developments
  - one in eight fell into the 'disaster' category: failed on two out of three "metrics"

>40% cost growth, >40% time slippage, < 50% 1st year production



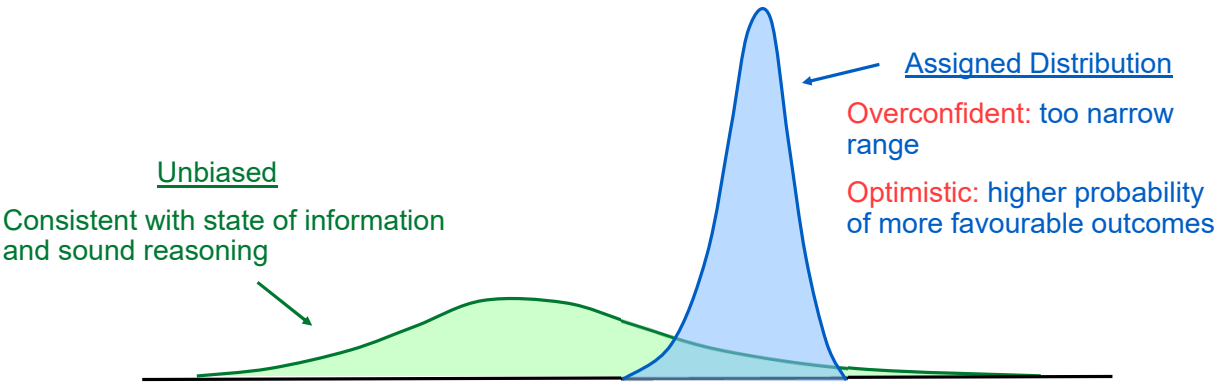
- Record even worse for mega-projects
- E&Y report 62% average cost overrun for mining mega-projects

Optimism: impact on risk



The more optimistic the forecasts, the more likely a project is to be chosen – but less likely to deliver! (and may displace a better option)

Overconfidence & Optimism – probably the most common situation!



Norwegian Continental Shelf probabilistic production forecasts (~500, 1995-2021)
Optimism: 15%-22%, under-production Overconfidence: 26-50% in 80% CI

Bratvold *et al*

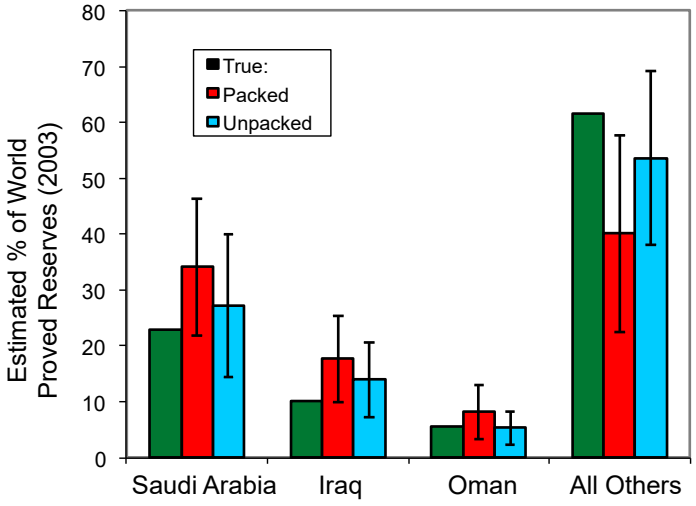
Overconfidence, optimism and organizational "dynamics"

- Every company has limited investment funds and time to devote to new projects.
- Fear that projects won't "get up" if the range of uncertainty is too wide because
  - decision-makers want the psychological comfort of narrower ranges, even when there is no value added
  - or don't know how to make decisions under uncertainty (choose option with highest EV, CE)
- The selection process often
  - encourages over-confidence in SMEs: "what are we paying you for?"; "you are supposed to be the expert, yet all you do is give these ranges"; "give me a number"
  - encourages optimism - uncertainty makes it plausible to forecast lower costs, shorter times, higher production
  - incentivises advocacy versus finding the best option – and there are rewards (implicit and explicit) for getting projects "across the line"
  - favours the most confident & articulate communicators – not necessarily the most knowledgeable

Tremendous "pressures" to be overconfident and optimistic !

Unpacking: Industry professionals

- Group A: packed Question
  - "What % of world proved oil reserves are in the following areas: Saudi Arabia, Iraq, Oman and All Others?"
- Group B: unpacked Question
  - "What % of world proved oil reserves are in the following areas: Saudi Arabia, Iraq, Oman Venezuela, USA, China, Russia, Nigeria and All Others?"

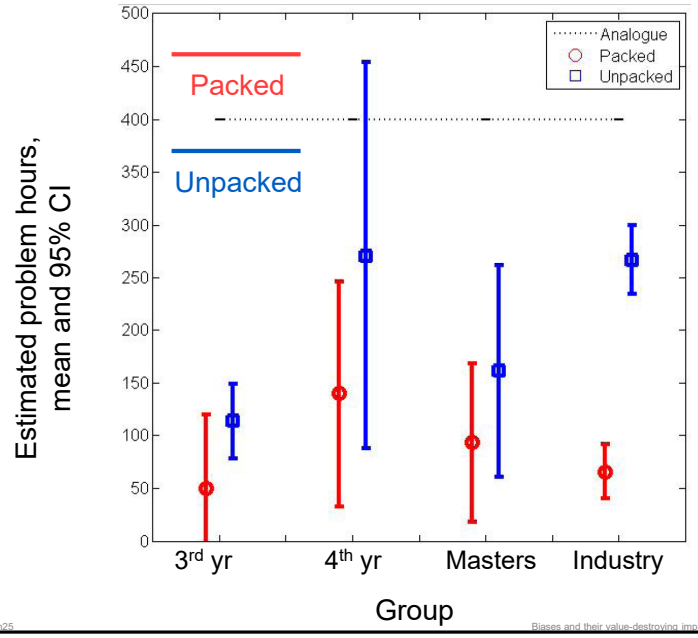


Unpacking example: estimating time to drill a well

- Four groups were given a description of a real well and asked to estimate drilling times, in hours
  - 3rd yr ASP Pet. Eng. undergraduates (no decision-making training)
  - 4th yr ASP Pet. Eng. undergraduates (some decision-making training)
  - "Conversion" Masters of Pet. Eng. (little Pet. Eng. Knowledge)
  - Industry petroleum engineers (with average 10 yrs experience)
- Approximately half were asked to estimate total time based on four Packed categories
  - Drilling, Tripping, Rigging and All associated problems

  and the rest given an Unpacked version where "All associated problems" was decomposed into 6 finer categories
  - Mud conditioning; Well-control operation; Fishing operations; Severe weather; Rig repair; Logistics delays

Unpacking: drilling "problem hours"



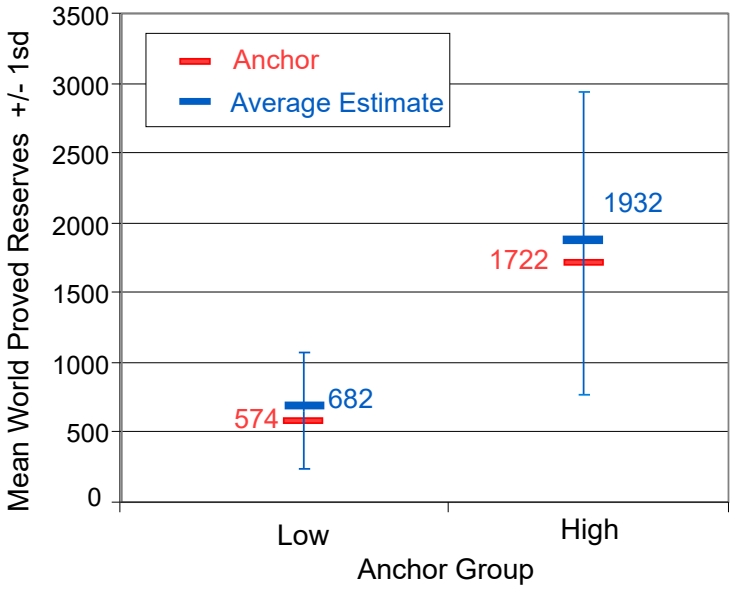
- All groups improved with unpacked question
- Industry ("expert") group had narrowest range – but had greatest improvement !
- 4th yr students had best estimate - and range included true value!

Unpacking Effect and The Planning Fallacy

- Planning Fallacy – a specific instance of the unpacking effect
  - tendency of people to underestimate completion times for complex tasks
- What causes the Planning Fallacy?
  - research suggests people construct a single mental scenario comprised of broad stages …. in which most things go according to plan ……
  - rather than with consideration of everything that could go wrong at each sub-stage
- Mitigation: be aware of situations when it may be present and decompose coarse categories into finer ones before attaching estimates to them
  - finer categories needs to be in some sense "equivalent" - don't use categories that are likely to have 1% and 30%
  - keep number of categories operational

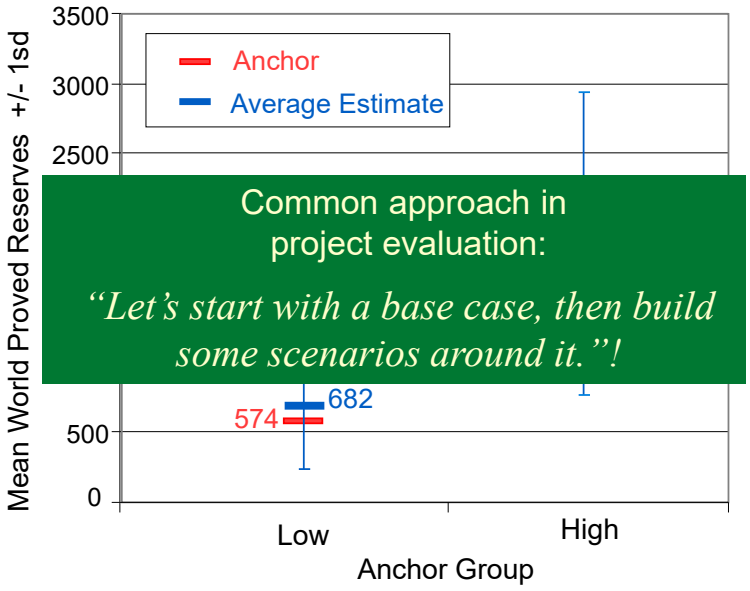
* Anchoring question: Industry professionals

- Alternate versions with high and low anchors
  - Group A: "Were world proved oil reserves in 2003 greater or less than 1722 Billion Barrels?"
  - Group B: "Were world proved oil reserves in 2003 greater or less than 574 Billion Barrels?"
- Both groups then asked the same question:
  - "What is your best estimate of the world proved oil reserves in 2003?"



Anchoring question: Industry professionals

- Alternate versions with high and low anchors
  - Group A: "Were world proved oil reserves in 2003 greater or less than 1722 Billion Barrels?"
  - Group B: "Were world proved oil reserves in 2003 greater or less than 574 Billion Barrels?"
- Both groups then asked the same question:
  - "What is your best estimate of the world proved oil reserves in 2003?"



Anchoring & Adjustment bias

- Describes a common heuristic when estimating values.
  - they use any given number/statistic/fact as a starting point (anchor) that sub-consciously dominates their judgement process
  - they adjust away from the anchor too little
- Examples
  - skilled negotiators start be setting an anchor
  - trial lawyers - "strike that comment"
  - resource industry mangers/leaders
- Random anchors can have just as large effects as credible anchors!
  - Attila the Hun;
  - San Francisco average daily temperature (558º)
- Subtle wording changes in a question can significantly impact responses
  - referenda and opinion polls

Availability Heuristic

- Are there more words in English that have the letter '*k*' as their first letter (e.g., kill) or as their third letter (e.g., awkward)?
  a) first letter
  b) third letter
- Experimental results (Tversky and Kahneman):
  - 2/3 of people thought words with *k* in the 1st position were more probable
  - reality: there are approximately twice as many words with *k* in the 3rd position as begin with a *k*
- The tendency of people to base judgements/beliefs on more easily recalled, recent, vivid or salient information.

> *"decision makers assess the frequency of a class, or probability of an event, by the ease with which instances or occurrences can be brought to mind"*

- Short-term memory is limited to ~7 "chunks" of information (5 for most people) - these fill up quickly with the most recent, salient and vivid cases

Availability: vividness

- Which of the following caused more deaths in Australia in 2003
  a) renal (kidney) disease?
  b) all transport-related accidents combined (ie land, air and sea)
- Renal failure - 15000 compared to 2100 for transport accidents (ABS, 2003)
- Typically, media do not report deaths by renal failure whereas they do report dramatic/vivid accidents
- Same applies to what is "reported" in casual conversations and social media

Beware: your memory (conscious & sub-conscious) is the source of information for intuitive probability judgements! Is its content, and use, reliable / unbiased?

Outline

- Decision-making and the brain
- Common biases and traps
  - simple judgements and chance assignments
  - "reasoning" under uncertainty
- Learning from visual judgements & mitigation tips

Judging likelihoods from information

- Linda is a 31 years old, single, outspoken and very bright. She majored in philosophy. As a student she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. (1970's context!)

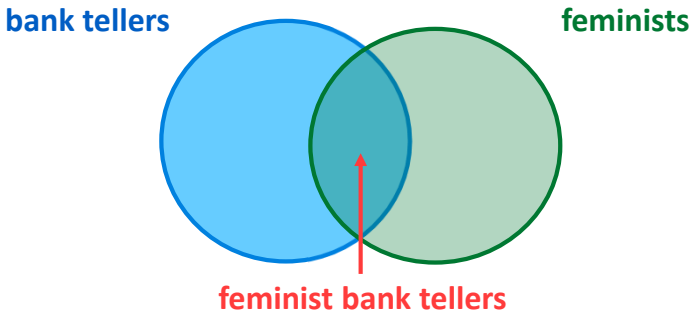
Which is the more likely alternative?

a) Linda is a bank teller
b) Linda is a bank teller and active in the feminist movement.

Answer:_____

Probability Intuition: Results and discussion of Linda question

- Nearly 90% of respondents choose the second alternative (bank teller *and* active in the feminist movement), even though this is *logically incorrect*)



P(Bank Teller) > P(Bank Teller AND Feminist)

Junctions ("ands") are always less likely than stand-alone statements.

Probability Intuition: Representativity <> Probability (likelihood)

- The description of Linda seems more representative of a feminist bank teller so people, wrongly, conclude it is more likely. Kahneman & Tversky (1982):
  - *"As the amount of detail in a scenario increases, its probability can only decrease steadily, but its representativeness and hence its apparent likelihood may increase. The reliance on representativeness, we believe, is a primary reason for the unwarranted appeal of detailed scenarios and the illusory sense of insight that such constructions often provide."*
- Related effect: the brain loves stories
  - we feel "satisfied", or relieved, when disparate bits of information "click into place" and develop too strong a degree of belief (=probability!) in the explanatory scenario (story)
  - most people prefer stories, and remember them better, than stand-alone data, facts, numbers – and stories often engage our emotions, reinforcing this effect
  - stories can be very persuasive – used to manipulate us to develop beliefs and make decisions that are in the story-teller's interest, not necessarily the organization's

Probability Intuition: Representativity <> Probability (likelihood)

- The description of Linda seems more representative of a feminist bank teller so people, wrongly, conclude it is more likely. Kahneman & Tversky (1982):
  - *"As the amount of detail in a scenario increases, its probability can only decrease steadily, but its representativeness and hence its apparent likelihood may increase. The reliance on representativeness, we believe, is a primary reason for the unwarranted appeal of detailed scenarios and the illusory sense of insight that such constructions often provide."*
- Related effect: the brain loves stories

Just as we need to beware of data/statistics that are manipulated to deceive / persuade us, we need to train ourselves to be sceptical of stories, especially when they weave in "facts" – they can be very compelling. Explore the motivation and ethics of the story-teller.

Representativity heuristic: short sequences of events

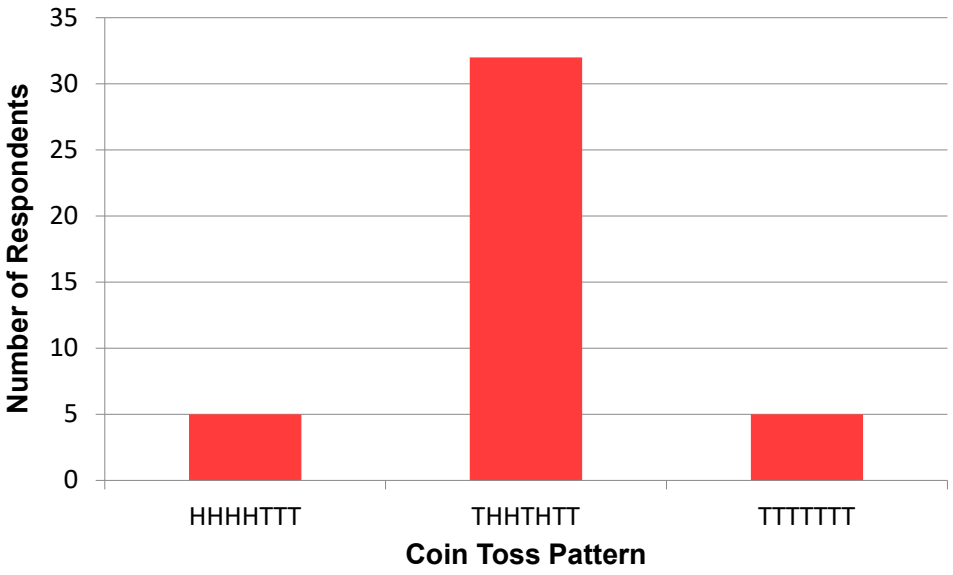
- I have just tossed a fair coin 7 times. You have not seen the result. You are invited to play a betting game to guess which of the three sequences below is the one I actually observed.
- Which sequence would you bet on?
- Using multiplicative rule for independent events

  $P(A\&B\&C\&D.) = P(A)*P(B)*P(C)*P(D)$ ….

  a) HHHHTTT $P = (1/2)^7 = 1/128$
  b) THHTHTT $P = (1/2)^7 = 1/128$
  c) TTTTTTT $P = (1/2)^7 = 1/128$

- Sequence b) seems more typical, but A sequences have the SAME probability and are thus EQUALLY likely (or rare!)
  - What if these are 7 projects, or products, with a probability of success of 0.5 – there is only a 1/128 chance of them all succeeding
- Don't confuse representivity (typicality) with chance (probability) !

Representativity heuristic: short sequences of events



The "Law of Small Numbers" (Tversky & Kahneman)

- We expect to see the same behaviour in small sequences that we would observe in large sequences
  - the mathematical "Law of Large Numbers" informs us of behaviours that are approximately true for large sequences, and rigorously true for sequences near to infinity
- Some sequences are seen as more balanced or more "typical" and are thus thought to be more probable. Typicality is mistaken for probability.
  - with the result that we over-estimate probability

Illusory Associations: Industry professionals

- You are testing a new technology, "High Spot", which might indicate the presence of commercial hydrocarbons (CHs). Observations from 25 unbiassed tests are:

When commercial hydrocarbons were

High Spot "said"	Present	Absent	Total
"Present"	16	4	20
"Absent"	4	1	5

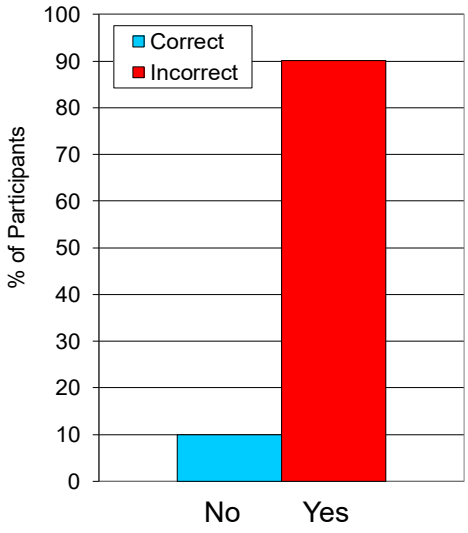
Do the data suggest that a "Present" High Spot signal is associated with the presence of commercial hydrocarbons? (yes/no)

- To check the claim, compare observed frequencies

CHs Present when High Spot said they were = 16/20 = 80%
CHs Present when High Spot said they were not = 4/5 = 80%

Frequency is the same irrespective of what High Spot says
- an Illusory Association

Illusory Associations: Industry professionals



* Some tips to avoid Illusory Associations

Don't come to conclusions based on just the number of positive occurrences you observe, or worse, remember!

- Collect sufficient observations on all combinations & calculate the probabilities

Real world outcome is

Test/Action "says"	True	False
"True"	Data	Data
"False"	Data	Data

- Check that your information collection itself is not biased
  - by data sampling, memory, vividness, recency, saliency, or confirmation bias
  - don't "mold" the evidence to fit your existing beliefs ☺ !

Beware of anecdotes, or a few occurrences, being used to create, prove or justify a belief or judgement (but they might be good to illustrate one)

Base-rate neglect: reliability of tests / predictors question

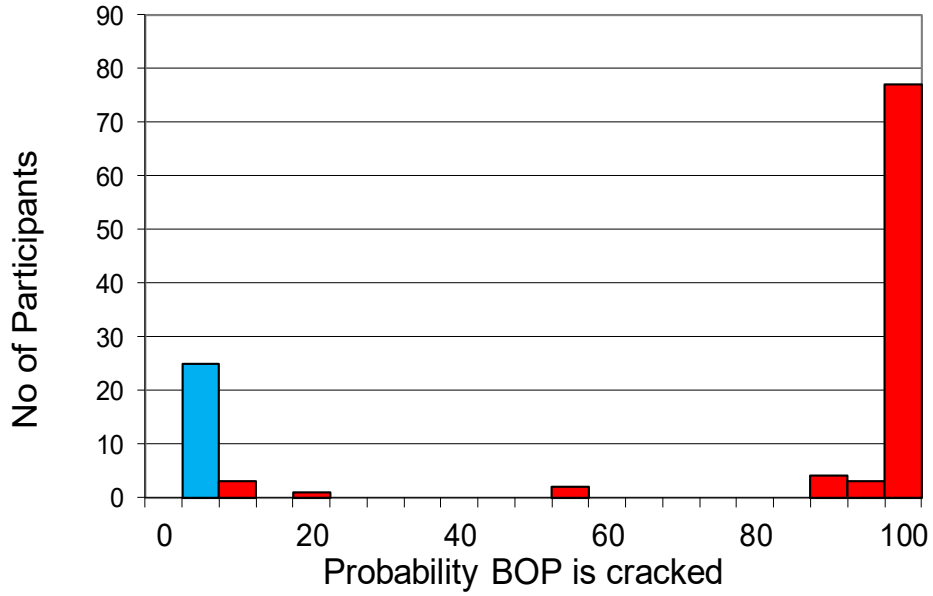
- Suppose historical estimates suggest 1 in every 1000 blow-out preventors (BOP) have serious cracks.
- Assume x-ray analysis is a very good, but not perfect, detector of these cracks.
  - if a BOP has cracks, x-rays will correctly say that it does 99% of the time
  - if a BOP does not have cracks, x-rays will correctly say that it doesn't 98% of the time
- A BOP has been x-rayed at random and the result was positive!
  - what is your intuitive assessment of the chances (%) that is cracked?

4.7%!

Probability (test positive given cracked) = 99%

Probability (cracked given test positive) = 4.7%

Base Rate Neglect: Industry professionals



OldSniff v NewSniff

- Your company's current remote sensing test (OldSniff) indicates (doesn't prove) the presence or absence of commercial hydrocarbons. You are approached by a start-up vendor who has developed a faster and cheaper test (NewSniff).
  - they claim that in any case when OldSniff indicates commercial hydrocarbons, their test will do likewise and will loan it to you to test this claim.
- You start a program to compare the two techniques on four prospects and have the following results to date:

	Prospect 1	Prospect 2	Prospect 3	Prospect 4
OldSniff says	present	absent		
NewSniff says			present	absent

- What is the minimum number of further tests needed to be to determine the truth of the vendor's claim?

1 with NewSniff: Y / N 2 with NewSniff: Y / N

3 with OldSniff: Y / N 4 with OldSniff: Y / N

OldSniff v NewSniff: Test Prosects 1 & 4

	Prospect 1	Prospect 2	Prospect 3	Prospect 4
OldSniff	positive	negative		Test
NewSniff	Test		positive	negative

The consequences of selecting each prospect for further testing are:

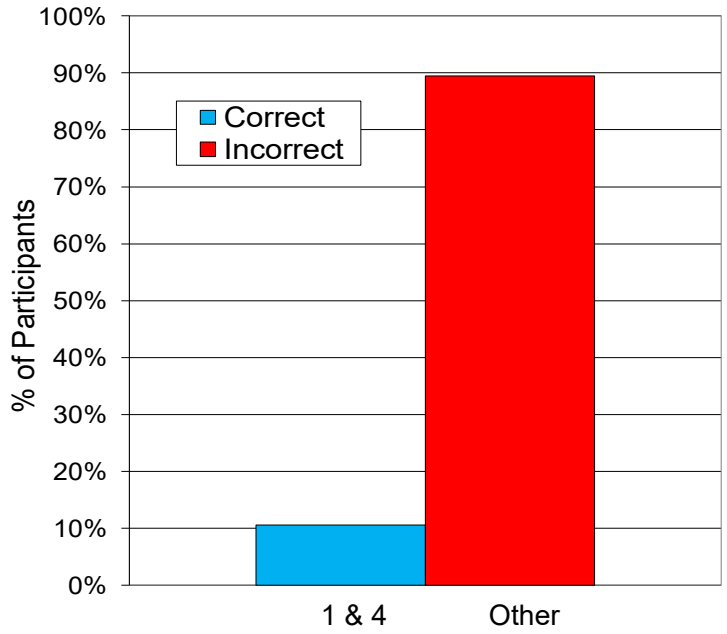
Prospect 1: If NewSniff returns a positive result, this is in accord with the claim but if it is negative this contradicts the claim. Therefore it should be tested.

Prospect 2: The vendor has made no claims regarding what result NewSniff will give when OldSniff is negative. Therefore further testing of this prospect will yield no additional information.

Prospect 3: If OldSniff is positive, this accords with the consultants' claim. If OldSniff is negative, then the results still accord with their claim. Further testing this prospect yields no additional information.

Prospect 4: If OldSniff is positive, then this contradicts the vendor's claim. If negative, it is in accord. Given that results from this prospect will provide a test of the vendor's claim, it should be tested.

Confirmation Bias: Industry professionals



Intuition

- Human beings are imperfect information processors (particularly sub-consciously)
- Many best judgements, beliefs and choices are non-intuitive, especially when uncertainty, novelty and complexity are present
  - we fall prey to systematic biases that result from mental short-cuts (heuristics), "gut-feel" and emotion (cognitive hands over to limbic)
- We need to use appropriate tools and "instruments" to keep our judgements and reasoning/thinking correct, to aid making good decisions
  - pilots are trained to ignore their intuitions and trust their instruments

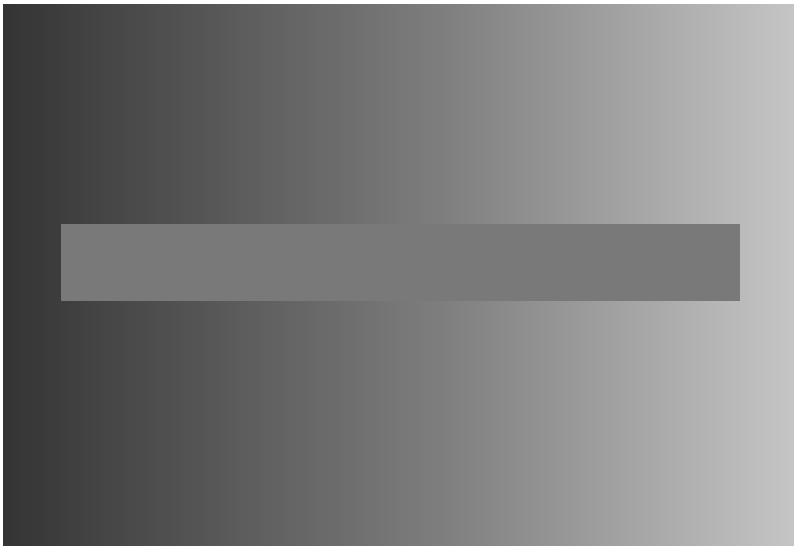
When to trust intuition when making a decision or judgement

- General consensus amongst cognitive scientists (eg Kahneman, Klein, Pearson, …) on when intuition (not = instinct!) can be used
- 5 questions to ask:
  - Am I in an emotional state? ("no"= pass)
    - could be subtle or strong - often initiated by "emotional tags" from previous similar situations
  - Is the "environment" predictable/regular? ("yes" = pass)
    - a stable relationship exists between identifiable cues/signals and subsequent events/outcomes
  - Has your intuition been trained/learnt in this environment? ("yes" = pass)
    - you have had prolonged practice and received rapid, unequivocable feedback on outcomes
  - Does the situation require probabilistic "thinking"? ("no"= pass!)
  - Do you have any personal motivation for a particular choice/outcome? ("no"= pass)
    - motivations that are inconsistent with the objectives of the decision or judgement
- If you "pass" all, then your intuition may be reliable. If not, use "System 2" supported by relevant computational and reasoning "tools"

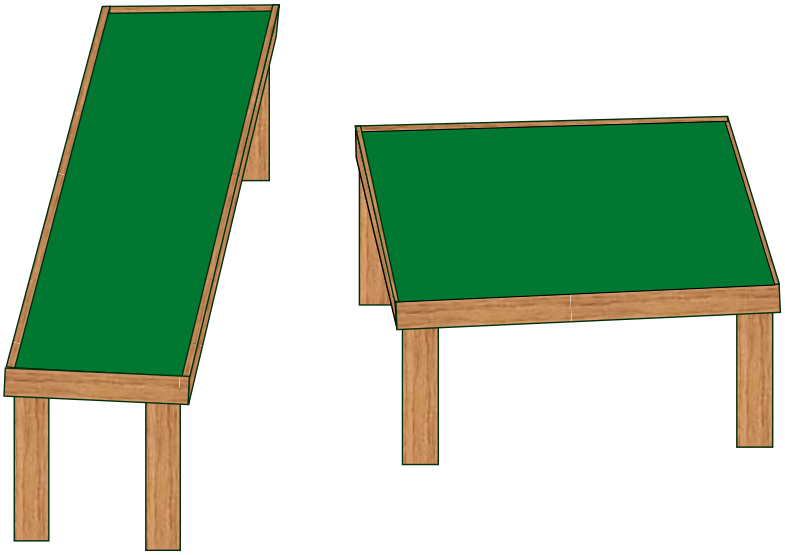
Outline

- Decision-making and the brain
- Common biases and traps
- Learning from visual judgements & mitigation tips

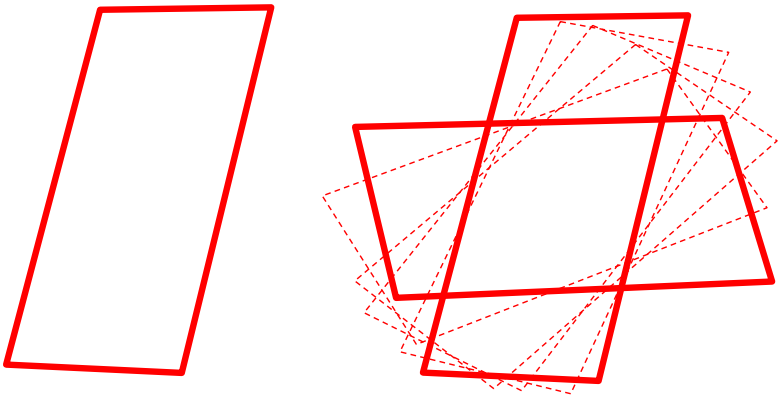
Grey gradation



Shepard's Table



Shepard's Table: both tables are same shape and size !



Learning from visual illusions

- Folk Wisdom states "Seeing is Believing"
  - people are convinced of the validity of their own sensory experience (and intuitions, "thought" processes and memories)
- Psychological research disagrees - people are subject to many cognitive illusions - contrary to the popular saying, perception is not always reality!

Visual illusions are a metaphor for cognitive illusions

- Awareness of illusions, by itself, does not necessarily produce a more accurate perception.
- Like visual illusions, erroneous beliefs and/or reasoning can remain compelling/convincing even when shown the truth - and thus be extremely difficult to overcome.

Kida's "6 basic mistakes we make in thinking":

1. We prefer stories to data, evidence, statistics
2. We seek to confirm, not question our ideas, beliefs
3. We underestimate the role of chance and coincidence in shaping events
4. We misperceive the world around us
5. We oversimplify our "thinking", and at worst, use intuition and gut-feel
6. We have faulty memories

And I think Kahneman (*"Thinking, Fast and Slow"*) would have added

Believing the above observations apply to other people, but not to me!

So what do we do? Some mitigation tips

Educate: create awareness of bias types, nature and situations where they are likely to be present - and an environment where their discussion is "normal"

- necessary, but not sufficient – a starting point!

- Seek unbiased information: become sceptical of small samples, anecdotes/stories, vividness/recency or confirmatory evidence
- Record & track probability assignments: compare to actual outcome frequencies (calibration) and make SMEs accountable
  - practice P10 – P90 ranges on upcoming events
- Pre-mortem: jump to the future and imagine there was an outcome outside your range – think of reasons why that might have happened – re-assess your range
- Assess (plausible) extremes first - use multiple 'anchors' (make use of "availability")

So what do we do? Some mitigation tips

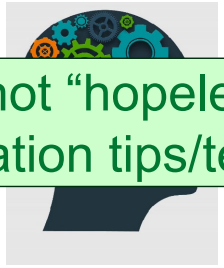
- Reference Class Forecasting: use the frequencies of known outcomes of similar projects ("outside" view) to scale project "inside view" uncertainty assessments
  - establish "base rates"
- Develop trained elicitors, using a formal protocol, for interviewing SMEs
  - eg Stanford Research Institute 5-step protocol
  - elicitation tools based on what the brain does well (relative comparisons) eg MOLE
- Motivational biases
  - leaders: identify & remove incentives (motivations) that drive biases
  - SME's: seek "external" opinions from people who don't have your motivations ("outside view")
- Think! Drop "system 1" and engage "system 2": reason using the rules & "tools" of logic and probability
  - trees, Bayesian belief networks, Monte Carlo simulation, analytical equations

Create a "culture" where bias is not tolerated, or worse, encouraged!

Key Takeaways

Our "caveman" brains are not (yet!) evolutionarily-endowed with unaided, sound thinking under uncertainty, complexity and novelty

Unwarranted / flawed beliefs, judgem predictions, a intuitions



But its not "hopeless" ☺
– use mitigation tips/techniques

in bias/error:
=> poor decisions
=> greater chance of undesired outcomes

… and this applies to Subject Matter Experts as much as lay-people!

Acknowledgements & Further Info

Dr. Matthew Welsh, University of Adelaide
Prof. Reidar Bratvold, University of Stavanger
Prof. Michael Lee, University of California, Irvine

"Bias in Science and Communication: A Field Guide"
MB Welsh, Institute of Physics Publishing, Bristol, 2018

"Thinking, Fast and Slow",
D Kahneman, Farrar, Strauss and Giroux, New York, 2011

"Don't Believe Everything you Think: The 6 basic mistakes we make in thinking"
TE Kida, Prometheus, New York, 2006 (more "accessible" than Kahneman)

References / Sources

Stanovich, K. E. (2011): "*Rationality and the Reflective Mind",* New York, Oxford University Press

E Merrow (2003): "*Taking on a Cult of Mediocrity"*, Upstream, May & Nandurdikur from JPT Oct14 p16

RB Bratvold, E Mohus, D Petutschnig & E Bickel (2020) *"Production forecasting: Optimistic and overconfident—over and over again",* SPE Reservoir Evaluation & Engineering 23.03

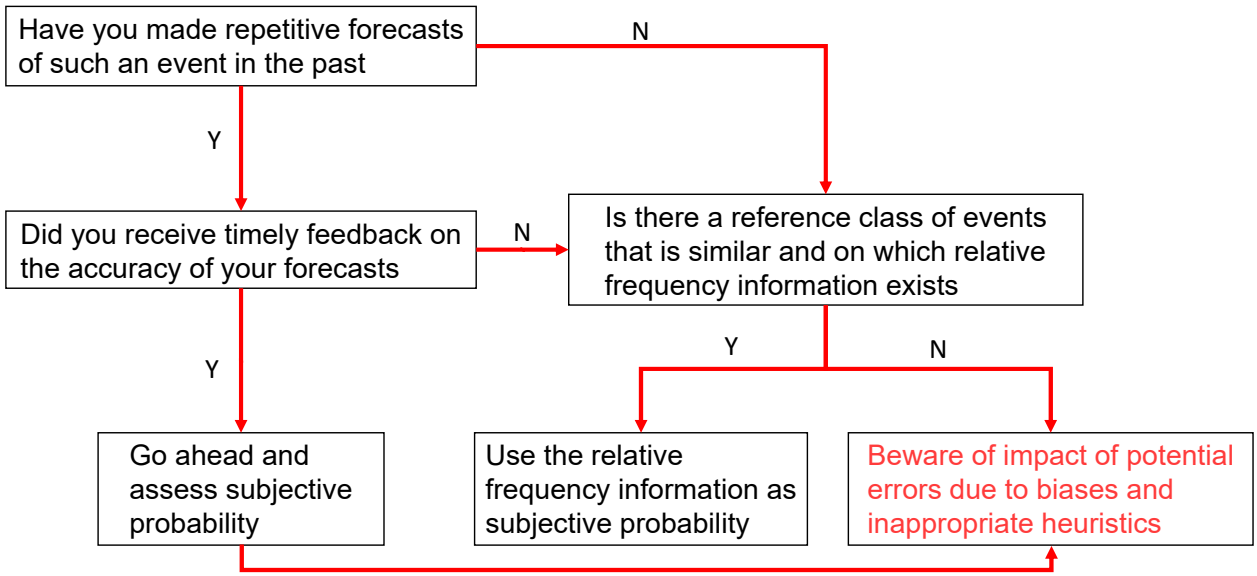
McKinsey Quarterly (May 2010) based on Kahneman & Cline (2009): *"Conditions for intuitive expertise: A failure to disagree"*, American Psychologist, Sept

BR Conway *et.al.* (2010): *" Advances in color science: from retina to behavior",* J Neurosci. 2010 November 10; 30(45): 14955–14963. (original image from Beau Lotto)

.

Additional information on avoiding biases associated with probability assessment & reasoning under uncertainty

Developing subjective probabilities



MOLE background

- There is a small benefit in just repeating the question and averaging the two results, Vul & Pashler (2009)
  - the benefit increases if a significant time elapses between the repetitions thus increasing independence of error in each estimate
  - problem if the person remembers their answer or is anchored by it
- People are better at relative judgements than absolute ones
- Welsh, Lee & Begg (2008, 2009) developed a More-Or-Less Elicitation (MOLE) technique for repeatedly asking a person about the same stimulus
  - maintaining the independence of errors without the need for time-lapse
  - using multiple anchors to mitigate effect of anchoring

The MOLE

- Industry partnership led to the development and testing of the MOLE tool
- Requires no specialised elicitation knowledge nor direct probability estimates by SMEs
- Subjects are repeatedly offered pairs of values selected at random - starting with very wide differences and gradually narrowing
  - possible values are only excluded if the subject is certain that they are not possible
- For each pair they are asked:
  - which is closer to the truth (a relative judgement)
  - to indicate how confident they were in their selection
- Algorithm calculates full PDF from a series of 10 such judgements
- Shown to improve calibration (including reduction in anchoring)
  - not restricted to elicitation from SME's

Eliciting subjective probabilities: The de Finetti Game

- Most people "lie" about probability without even being aware of it – they even lie to themselves.
- The de Finetti game: a device to "objectively" measure subjective probability
  - Bruno de Finetti(1906–1985), an Italian statistician. Worked in the middle ground between mathematics and psychology
- Example: suppose your colleague has been assessing the potential for a trap to exist at a prospect location
- She might tell you:
  - *"Given all the data I've looked at, I am very sure there is a trap present."*
- The de Finetti game is a way to measure how sure she *really* is about a trap being present.
- We need to ask her a series of questions to assess her true subjective probability

Eliciting subjective probabilities

- Tell you colleague the following: *"Let's play a game. You have a choice*
  - *you can either draw a ball from a bag that has 98 red balls and 2 green balls. If you happen to draw a red ball, I will give you one million dollars,*
  - *or you can decide to wait untill we drill and find out if there really is a trap – if there is one I will give you one million dollars.*

  what's your choice: draw or wait?"
- If your colleague says *"Draw from the bag"* she must be less then 98% certain that there really is a trap
- Now you ask the next question:
  - *"Now there are 70 red balls and 30 green balls. Do you want to draw and get a million $ if it is red, or wait to drill and if there really is a trap get a million $?"*
- If your colleague says *"Wait till we drill and find out"* then you know that she is more than 70% sure that there really is a trap.

Eliciting subjective probabilities

- Now choose a value in between, such as 85 red balls and ask:
  - *"Now there are 85 red balls in the bag and 15 green ones. Do you want to draw, and if you obtain a red ball get a million dollars, or wait to see the well result?"*
- If your friend says *"Draw from the bag."* then you know that she less than 90% sure that she will get an HD
- Continue asking such intermediate questions. At some point, say at 83 red balls (83%), your colleague may say something like:
  - *"I'm indifferent (or I can't decide) between drawing a ball and waiting for the well result."*
- This (83%) is her degree of belief (subjective probability) that a trap really exists

A formal probability elicitation protocol for trained elictors

- Based on that developed at Stanford University and Stanford Research Institute in ’70s*
- 5-phase process
  1. Motivate
  2. Structure
  3. Condition
  4. Encode
  5. Verify

* This version adapted from summary in “Uncertainty” by Morgan & Henrion

1. Motivate

- Elicitor develops rapport with Subject Matter Expert (SME)
  - explains why quantifying uncertainty through probabilities is important for decisions
  - clarify that this is not about making a prediction that SME (or anyone else) will be held to
- Elicitor searches for any motivational biases of SME
  - wants to influence decision
  - reward/penalty based on outcome
  - maintain status as an authority
  - has a stated “position” to defend
- Elicitor distinguishes between
  - probability (likelihood) of outcomes, and
  - importance/severity/desirability of outcomes

2. Structure

- Unambiguous definition of the quantity to be assessed
  - find the form that the SME is able to assess most reliably – eg appropriate units
  - might need to decompose (unpack) the quantity into others that are more easily assessed – the use a model to aggregate back to the quantity of interest
- Clairvoyant test
  - could someone who knows everything about the past, present and future tell you the value of the quantity without have to ask clarifying questions?
- Identify any unstated assumptions
  - are the probabilities conditional on the assumptions – eg probability of time, or cost, dependent on assumption of no change in design?
  - need to get "unconditional" probabilities

3. Condition

- Elicitor makes sure the SME understands the subjective nature of uncertainty
- Elicitor helps SME to
  - think about what data, science, logic, models or other information they have available and how these will be used
  - understand cognitive biases and how to avoid them (see previous tips) – particularly overconfidence, anchoring, optimism, availability, representativeness, positive illusions – use some P10-P90 calibration questions

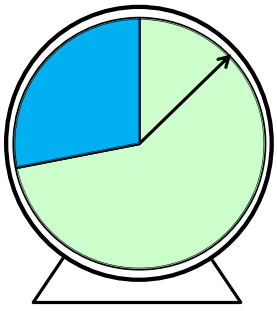
4. Encode

- Continuous quantities
  - ask for extreme values first (i.e. minimum, and maximum)
  - challenge for even more extremes e.g using pre-mortem
  - take a set of values between min and max of the uncertain quantity, x, and ask for P(X <= x), or P(X > x) if more natural – could use probability wheel if SME not familiar with probabilities
  - don't ask in order of values of x, or start with a central value
  - use odds, "1 in 500 chance", for very low probabilities
- Discrete quantities
  - ask for probability of least likely event first
- In both cases
  - Might need to cycle back to assumptions, bias education, etc.
  - do not show CDF to the SME at this stage

Use the de Finetti game aided by a probability wheel

Unambiguous uncertain quantity: average price of barrel of WTI in US$ for 2025

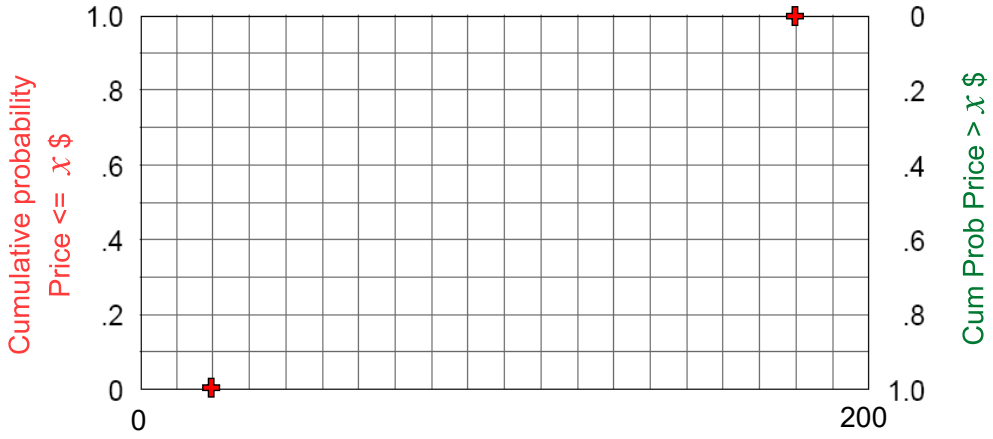
- Ask SME: would you bet on
  - the price being below \$ x dollars? or
  - spinning the arrow and it landing in the blue sector
- Adjust blue area until SME is indifferent between the two bets
  - start with green:blue area big enough to make the choice easy, then reverse, to make the choice easy again
  - hone in on the point of indifference



Continuous Quantity Probability Assessment:

(don't show chart to SME yet)

Start with absolute Min and Max
(test/push for wider range, eg using a pre-mortem)

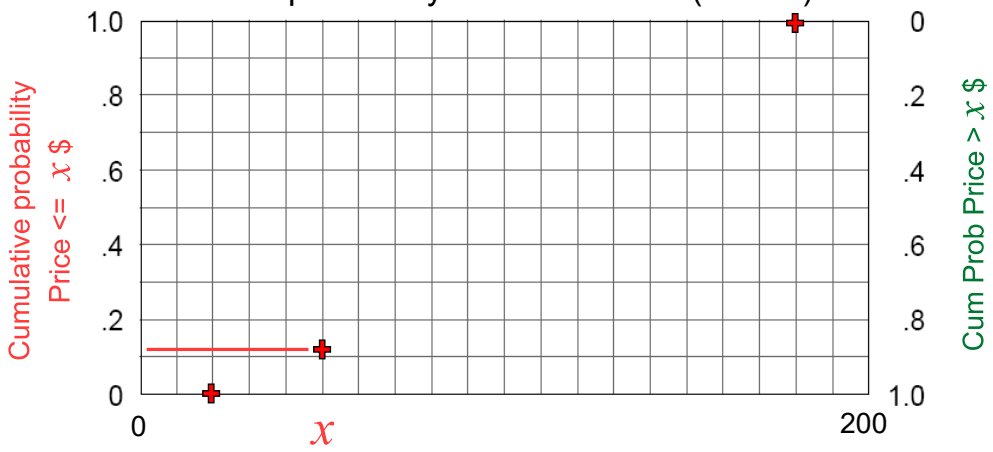


Unambiguous uncertain quantity: average price of WTI for 2025, US$/bbl

Continuous Quantity Probability Assessment:

(don't show chart to SME yet)

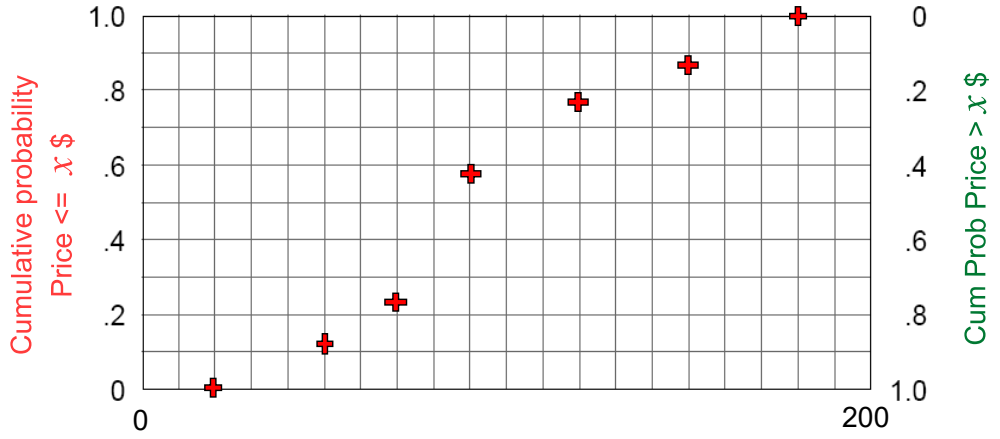
Choose an x and define the event as $X <= x$ or $X > x$
- use probability wheel to elicit P($X <= x$)



Unambiguous uncertain quantity: average price of WTI for 2025, US$/bbl

Continuous Quantity Probability Assessment:
(don't show chart to SME yet)

Continue define more events as $X <= x$ (or $X > x$) and elicit their probabilities



Unambiguous uncertain quantity: average price of WTI for 2025, US$/bbl

5. Verify

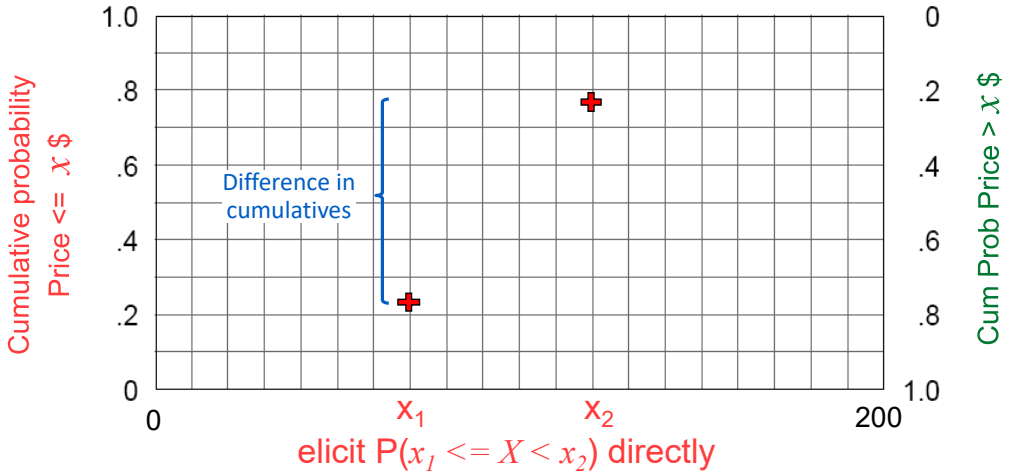
- Test if the resultant CDF and PDF really do reflect the SME's beliefs
  - directly elicit probabilities for events $x_1 <= X < x_2$, and check if they match the difference in the cumulatives: $P(X < x_2) - P(X <= x_1)$
  - can also apply this to the complements: eg for an x that was previously elicited, directly elicit $P(X > x)$ and compare to 1 minus the $P(X <= x)$
  - now show PDF and CDF to the SME
- Ask indifference questions. Would the SME rather bet on:
  - whether the actual value will be below the 10th percentile or above the 90th percentile?
  - whether the actual value will be between the 10th percentile and 50th percentile or between the 50th percentile and 90th percentile?
  - whether the actual value will be above or below the 50th percentile?

If the SME is not indifferent to these bets, revise the assessment

Continuous Quantity Probability Assessment:

(show chart to SME after testing consistency)

Test consistency by directly eliciting probabilities of events $x_1 <= X < x_2$
– do they match the differences in cumulatives?



Unambiguous uncertain quantity: average price of WTI for 2025, US$/bbl

* Avoiding biases when reasoning under uncertainty: Probability training

- Learn the basics of probability (and statistics - remembering they are different: uncertainty vs variability!)
- In terms of practical applicability, probability theory is comparable with geometry;
  - both are branches of applied mathematics that are directly linked with the problems of daily life.
- While most people have a natural feel for geometry (at least to some extent), many people clearly have trouble developing a good intuition for probability.
- Arguably, in no other branch of mathematics is it so easy to make mistakes as in probability theory.
  - Conditional probabilities, and Bayes theorem in particular, can be especially difficult

* Avoiding biases when reasoning under uncertainty: Probability training

- Learn the basics of probability (and statistics - remembering they are different: uncertainty vs variability!)
- In terms of practical applicability, probability theory is comparable with geometry;
  - both are branches of applied mathematics, directly linked with the problems of daily life.
-
-

"The theory of probabilities is at bottom nothing but common sense reduced to calculus; ...

it teaches us to avoid the illusions which often mislead us;

... there is no science more worthy of our contemplations nor a more useful one for admission to our system of public education."

Laplace – Theorie Analytique des Probabilites