

e-Cartoon: The 12 Traps of Uncertain Reasoning:

Uncertainty Aptitude versus Weapons of Mass Deception

Roger M. Cooke

Sept 20, 2025

Gratis e-Cartoon. Donations to [Med Share](#) could express appreciation

Contents

Prologue

Nudge Theory

Aristotle versus Ramsey

References

1. Certainty trap

Uncertainty Denial

2. Confidence trap

Thinking high confidence in each equals high confidence in all

3. Propagation trap

Confusing joint and conditional probability

4. Chaining trap

Thinking that a chain of probabilistic reasons is no weaker than its weakest link

5. Innumeracy trap

Inability to do elementary math

6. Procrastination trap

Thinking that not deciding is cost free

7. Onus Probandi trap

Letting an adversary allocate your burden of proof

8. Disagreement trap

Thinking that disagreement reduces certainty

9. Agreement trap

Thinking that agreement raises certainty

10. Negative Learning trap

Think that new information always reduces uncertainty

11. Bogsat trap

Trusting Bunch of Guys/Gals Sitting Around a Table

12. Communication trap

Trying to communicate what you don't understand

Prologue

The U.S. Office of Government Ethics (OGE) was established by the Ethics in Government Act of 1978. New employees are required to complete initial ethics training. Senior staff may have additional training requirements. Training covers topics like conflicts of interest, financial disclosure, misuse of position, and other relevant laws and regulations. Ethics training aims to help employees understand their ethical obligations, identify potential ethical issues, and make sound decisions in the public interest. <https://www.oge.gov/>.

Decision makers in both the private and public sectors are often tasked with reasoning and communicating under uncertainty. There are correct and incorrect ways of doing this. All parties need to recognize the strategies for gaming uncertainty to advance particular interests. Absent that, the market place of uncertain opinions becomes a battleground of opposing interests wielding weapons of mass deception.

Historical background: Nudge Theory

Tversky and Kahneman's 1974, 1982 seminal works on Judgment under Uncertainty: Heuristics and Biases, studied biases in estimating subjective probabilities. Among the most well-known are representativeness, availability, anchoring and overconfidence. "Heuristics" denotes mental shortcuts that could bring one close to the right answers quickly. A "bias" on the other hand is a heuristic gone awry, it is a systemic deviation from the truth. Thaler and Sunstein (2009) picked up the thread in 2009, identifying two types of human thinking, the "Reflective System" and the "Automatic System". The Automatic System, tethered to the reptilian brain, is the province of heuristics. "Libertarian paternalism" provides "nudges" to aid the automatic system in taking better decisions. "Libertarian" in this context means "liberty preserving", and "paternalism" is engaged because people are fallible, not because they are irrational or dumb (Thaler and Sunstein, 2021, p.25). A neuro-physiological substrate for the automatic/reflective distinction is found in the dorsal / ventral vagal complex (Porges and Porges 2023).

Nobel-prize winning psychologist Daniel Kahneman took this idea further in his best-seller "Thinking, Fast and Slow" (Kahneman 2011). He promoted several "social priming" studies, showing how behavior is influenced by subtle psycho-social cues. By carefully 'nudging' people with such cues, one hopes to promote better decision making without coercion. Kahneman wrote: "Disbelief is not an option, ...The results are not made up, nor are they statistical flukes. You have no choice but to accept that the major conclusions of these studies are true". Subsequent research showed that most of these studies could not be replicated (Richie, 2020), with Kahneman writing to students of social priming that "a train wreck" was approaching (Chivers 2019).

The train wreck has arrived yet our thought leaders are still TEACHING people to reason incorrectly under uncertainty, year after year. The problem is not fast versus slow thinking, but the very slow process involved in deliberative decision making. Nor is the problem communicating uncertainty to the fast-thinking man on the street. The problem is that our thought leaders and science communicators do not understand uncertainty – you cannot communicate what you do not understand (Cooke 2014). The result is that people in general, as well as media, are unable to call out and disarm these defective narratives.

If there were no such thing as good/bad quantification of uncertainty there could be no such thing as good/bad reasoning under uncertainty. Validating combinations of experts' judgments underpins this work. A dedicated [website](#) provides a portal to this work with links to blogs, wiki's, videos, publications and software. Lite highlights are [here](#), [here](#), [here](#) and [here](#). Validation is not the subject of this cartoon. Some references are listed at the end of his prologue and a [video](#) explains in-, out-of sample validation and the persistence of performance. Suffice to say that reasoning and deciding under uncertainty is not a question of providing the

right psycho-social cues to “fast thinkers”, it is a question of developing uncertainty aptitude among scientists, decision makers, general public and especially uncertainty communicators. To that end, this humble Cartoon hopes raise interest in Uncertainty Aptitude Training with a syllabus of errors: the twelve traps and deceptions which arise when reasoning under uncertainty. Memes hope to promote rapid recall.

First however, we must establish the distinction between reasoning under certainty - deterministic reasoning - and reasoning under uncertainty, - probabilistic reasoning. It will seem like small affair to all but the philosophers, but probabilistic and deterministic reasoning are fundamentally different. Most of the traps come down to failing to appreciate this.

Aristotle vs Ramsey

Aristotle (384 BC – 322 BC) is the father of deterministic logic. Writing $A(x)$ for "x is in set A" or "x is an A", his BARBARA syllogism may be written as

For all x if $A(x)$ then $B(x)$
 For all x if $B(x)$ then $C(x)$
then
 For all x if $A(x)$ then $C(x)$



Aristotle

This inference preserves the truth of the premises in the conclusion; if the premises (above the line) are true then necessarily the conclusion (below the line) is true.

Frank Ramsey (1903 – 1930) may be regarded as the father of reasoning under uncertainty. Suppose we add a wee bit of uncertainty to BARBARA. For some very very small $\epsilon > 0$ and $\delta > 0$ consider

For all x if $A(x)$ then with **certainty $1-\epsilon$** , $B(x)$
 For all x if $B(x)$ then with **certainty $1-\epsilon$** , $C(x)$
then
 For all x if $A(x)$ then with **certainty $1-\delta$** , $C(x)$



Frank Ramsey

Is this truth preserving? If the premises are true, is there some $\delta < 1$ for which the conclusion is always true? NO. The only value of δ for which the conclusion is always true is $\delta = 1$. This syllogism is truth preserving only if the uncertainty is zero. Here's an example:

For all x, if person x is more than 2.2 m tall, then **with certainty $1-\epsilon$** , x weighs less than 500 kg.
 For all x, if person x weighs less than 500 kg then **with certainty $1-\epsilon$** x is less than 2.2 m tall.
then ????
 For all x, if x is more than 2.2 m tall, then **with certainty $1-\delta$** , x is less than 2.2 m tall.

This inference is NOT truth preserving with a vengeance. We can surely find a small $\epsilon > 0$ for which the premises are true, but the conclusion is a contradiction. The only value of δ for which it is true is $\delta = 1$. *Reasoning is not continuous in certainty.* A pattern of reasoning which is valid under certainty is not "almost

valid" if a wee bit of uncertainty leaks in. Admitting uncertainty requires a total overhaul. Contemplating how much uncertainty has already leaked into our daily discourse underscores the urgency.

References

- Bamber, J.,L., Oppenheimer, M. , Kopp, R. E., Aspinall, W. P. , and Cooke, R. M. (2019), "Ice sheet contributions to future sea-level rise from structured expert judgment," *Proceedings of the National Academy of Sciences*, vol. 116, no. 23, pp. 11195–11200.
- Burgman, M., (2005) *Risks and decisions for conservation and environmental management*. Cambridge University Press.
- Chivers, T. (2019) "What's next for psychology's embattled field of social priming," *Nature*, vol. 576, no. 7786, pp. 200–20.
- Cooke, R. M. (2014), "Messaging climate change uncertainty," *Nature Climate Change*, vol. 5, no. 1, p. 8, 2014.
- Cooke, R.M. Goossens, L.J.H., (2000) *Procedures guide for structured expert judgment Project report EUR 18820EN*, Nuclear science and technology, specific programme Nuclear fission safety 1994-98, Report to: European Commission. Luxembourg, Euratom.
- Cooke, R. M. and Goossens, L. L. (2008) "TU Delft expert judgment data base," *Reliability Engineering & System Safety*, special issue on expert judgment, vol. 93, no. 5, pp. 657–674, 2008.
- Cooke, R.M. (1991) *Experts in uncertainty: opinion and subjective probability in science*. Oxford University Press on Demand, 1991.
- Cooke and Goossens, L., L., (1994) "Procedures guide for structured expert judgment Project report EUR 18820EN," Nuclear science and technology, specific programme Nuclear fission safety, vol. 98, pp. 303–311.
- Colson A.R. and Cooke, R. M. (2-17) "Cross validation for the classical model of structured expert judgment," *Reliability Engineering & System Safety*, vol. 163, pp. 109–120.
- Colson A., R., and Cooke, R. M. (2018), "Expert elicitation: using the classical model to validate experts' judgments," *Review of Environmental Economics and Policy*, vol. 12, no. 1, pp. 113–132.
- Cooke, R.M., Marti, D. and Mazzuchi, T. (2021) "Expert forecasting with and without uncertainty quantification and weighting: What do the data say?," *International Journal of Forecasting*, vol. 37, no. 1, pp. 378–387.
- Cooke, Roger M., (2025) *Wisdom / Madness of Crowds and Perils of Point Forecasts*, ESREL 2025, Stavanger, Norway.
- Eggstaff, J. W., Mazzuchi, T. A. and Sarkani, S. (2014), "The effect of the number of seed variables on the performance of Cooke's classical model," *Reliability Engineering & System Safety*, vol. 121, pp. 72–82.
- Kahneman, D. , (2011) *Thinking, fast and slow*. Macmillan.
- Kahneman, S. P. Slovic, P. Slovic, and A. Tversky, (1982) *Judgment under uncertainty: Heuristics and Biases*. Cambridge U.Press.
- Porges, Sephan and Porges, Seth, (2023) *Our Polyvagal World: How Safety and Trauma Change Us* W. W. Norton & Company. ISBN: 9781324030256.
- Ritchie, S. (2020) *Science fictions: Exposing fraud, bias, negligence and hype in science*. Random House.
- Thaler, Richard & Sunstein, C., (2009). *NUDGE: Improving Decisions About Health, Wealth, and Happiness*, New York Penguin Books
- Thaler, R. H., & Sunstein, C. R. (2021). *Nudge: The Final Edition*. New York: Penguin Books
- Tversky, A. and Kahnemann D. (1974) *Judgment under Uncertainty: Heuristics and Biases*, *Science* vol. 185 pp1124 – 1131.

Trap 1

Certainty trap: Uncertainty Denial



"We should have no concern for the environment because after the great flood with Noah, god promised that he would never ruin the earth again"

Sen. James Inhofe [2003](#)



"As a Christian, I believe that there is a creator in God who is much bigger than us... if there's a real problem, He can take care of it." **US Rep Tim Walberg** [2017](#)



"It is obvious that the spread of this virus is an act of Allah.Allah sent a disease upon them and this disease laid siege to 40 million [Chinese people] **Hadi Al Modarresi** [Feb .28 2020](#)



"We are all created in the image and likeness of God. That image is seen the most by our face. I will not wear a mask." **Ohio Rep Nino Vitali** [2020](#)



"I won't get into the debate about climate change but I'll simply point out that I think in academia we all agree that the temperature on Mars is exactly as it his here. Nobody will dispute that. Yet there are no coal mines on Mars. There's no factories on Mars that I'm aware of" **Republican state senator of Kentucky Brandon Smith**, [2014](#)



"Never again will I curse the ground because of man, even though all inclinations of his heart are evil from childhood and never again will I destroy all living creatures as I have done." Genesis 8:21-22 quoted by **US Rep John Shimkus (R, Ill)** [2009 to deny climate change.](#)



"The concept of global warming was created by and for the Chinese in order to make U.S. manufacturing non-competitive." 2012. ["I know much about climate change. I've received environmental awards"](#). (Jan. 18, 2016)

The certainty trap involves finding reasons for denying that there is any uncertainty. This is a very abridged list of examples. For context US Senator James Inhofe was elected Chair of the *Senate Committee on Environment and Public Works* in 2003. He compared the Environmental Protection Agency to the Gestapo. John Shimkus was chairman of the House Energy and Commerce Committee, serving as Chairman and then Republican Leader of its Environment and Climate Change Subcommittee.

Trap 2

Confidence trap:

*Thinking confidence in each equals confidence in all:
Aristotle instead of Ramsey*

This is a big one and expands the gap between deterministic and probabilistic reasoning. We start with a fictionalized conversation between Bill and Jane. You can also watch a [video](#) for details.

It'll be a Huge Day

I'm highly confident my car won't be ticketed,
I'm highly confident the sun will shine,
I'm highly confident my team will win and
I'm highly confident my portfolio will go up

So, you're highly confident in
No ticket, sunshine, win and
portfolio up?

That's what I just said

No it isn't!

I'm highly confident that 6 won't come up on
the first throw with this die
And on the second
And on the third and fourth BUT NOT ON ALL FOUR
The probability of no 6 in four throws is about $\frac{1}{2}$



The confidence trap has been institutionalized by the IPCC with its levels of confidence:

**Guidance Notes for Lead Authors of the
IPCC Fourth Assessment Report on Addressing Uncertainties**

Table 3. Quantitatively calibrated levels of confidence.

Terminology	Degree of confidence in being correct
<i>Very High confidence</i>	At least 9 out of 10 chance of being correct
<i>High confidence</i>	About 8 out of 10 chance
<i>Medium confidence</i>	About 5 out of 10 chance
<i>Low confidence</i>	About 2 out of 10 chance
<i>Very low confidence</i>	Less than 1 out of 10 chance

Table 4. Likelihood Scale.

Terminology	Likelihood of the occurrence/ outcome
<i>Virtually certain</i>	> 99% probability of occurrence
<i>Very likely</i>	> 90% probability
<i>Likely</i>	> 66% probability
<i>About as likely as not</i>	33 to 66% probability
<i>Unlikely</i>	< 33% probability
<i>Very unlikely</i>	< 10% probability
<i>Exceptionally unlikely</i>	< 1% probability

This reasoning gets applied in the US National Research Council 2010 *Advancing the Science of Climate Change* (http://www.nap.edu/catalog.php?record_id=12782). The NRC report bases its first summary conclusion on “high confidence” (at least 8 out of 10) or “very high confidence” (at least 9 out of 10) in six (paraphrased) statements (we go with "high confidence"):

Baiting the trap
How confident are you in all of them?

1. Earth is warming
2. Most of the warming over the last several decades can be attributed to human activities
3. Natural climate variability ... cannot explain or offset the long-term warming trend.
4. Global warming is closely associated with a broad spectrum of other changes,
5. Human-induced climate change and its impacts will continue for many decades,
6. The ultimate magnitude of climate change and the severity of its impacts depend strongly on the actions that human societies take to respond to these risks.

High confidence $\geq 80\%$ chance

High confidence
High confidence

High confidence

High confidence

High confidence



Trap 3

Propagation trap *confusing conditional and joint probability*

High Confidence: *Earth is warming*

High Confidence: *Most of **the** warming over the last several decades can be attributed to human activities.*

“**THE**” warming?

Does the second sentence say, “High confidence Earth is warming” **AND** humans caused it” (joint probability)

or does it say

“High confidence that **GIVEN** Earth is warming, Humans caused it” (conditional probability)?

If the authors of the highly word-smithed document knew, we wouldn’t be asking.

If its a conditional, we need to multiply high confidences

High Confidence: *Earth is warming*

High Confidence: “**Given** Earth is warming, humans caused it”

Baiting the trap



High Confidence (condition) × High Confidence(consequence | condition) = 64%?

Medium confidence

If statements 2– 6 are conditional on precedents, then confidence in **ALL** is $0.8^6 = 0.26$: *Low confidence*

Trap 4

Chaining trap not realizing that a probabilistic chain is weaker than its weakest link

This gained prominence during the COVID pandemic. To set the stage, recall the discussions in the UK:



Boris Johnson, 3 Feb 2020 "...there is a risk that new diseases such as coronavirus will trigger a panic and a desire for market segregation that go beyond what is medically rational to the point of doing real and unnecessary economic damage.....humanity needs

....some country ready to take off its Clark Kent spectacles and leap into the phone booth and emerge with its cloak flowing as the supercharged champion, of the right of the populations of the earth to buy and sell freely among each other."

<https://www.gov.uk/government/speeches/pm-speech-in-greenwich-3-february-2020>

March 12: Patrick Vallance told BBC that a "key" aim would be to "build up some kind of herd immunity so more people are immune to this disease, and we reduce the transmission. **March 13** The stands were packed during day tour of the Cheltenham Festival

<https://www.bbc.co.uk/mediacentre/statements/panorama-mon-27-apr>

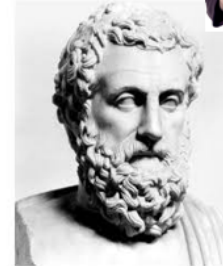
A photograph of a horse racing event. In the foreground, several horses and jockeys are running on a green track. In the background, there are large, multi-tiered stands filled with a large crowd of spectators. The sky is overcast.

"... steps to remove Covid-19 from the High Consequence Infectious Diseases (HCID) list were taken by the DHSC [Dept Health & Social Care] on 13 March, the same day PPE guidance was downgraded"

UK Reasoning for **no lockdown, Herd Immunity**; ‘We have good reason to believe:’ (GUARDIAN 30 April);

- Reproduction Nr < 2
- Case Fatality Rate < 1%
- Most covid-19 hospital cases do not need IC
- Social Distancing will fail

**Baiting
the trap**



Therefore: Herd Immunity, No Lockdown

Assign representative probabilities

- Prob (Reproduction Nr < 2) = 80%
- Prob (Case Fatality Rate < 1%) = 90%
- Prob (covid-19 hospital case not need IC) = 85%
- Prob (Social Distancing failure) = 80%

If independent: $P(\text{All four}) = 0.49$



Trap 5

Innumeracy trap inability to do simple uncertainty math

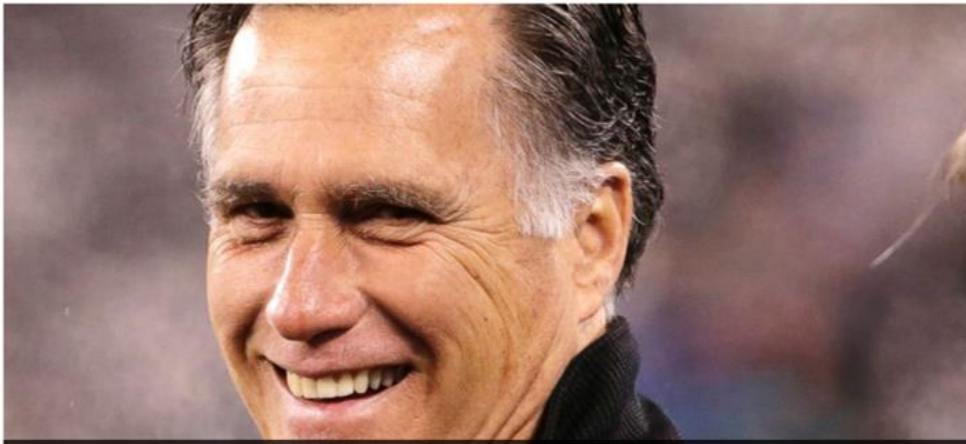
- IPCC: “virtually certain” = 99% ~ 100% probability
- Suppose US Nuclear Regulatory Commission licensed reactors each year as “virtually certain” not to fail
- With 100 reactors, each with Prob Failure = 0.01 /yr; Expected failure frequency = 1/yr.

**Baiting
the trap**



Trap 6

Procrastination trap: waiting till the facts are in when decisions can't wait



**Baiting
the trap**



“My view is that we don’t know what’s causing climate change on this planet. And the idea of spending trillions and trillions of dollars to try to reduce CO2 emissions is not the right course for us.” (2011, Pittsburgh)

<https://abcnews.go.com/Politics/mitt-romneys-climate-change-views-evolving/story?id=28405142>

Trap 7

Onus Probandi trap letting others gerrymander the proof burden

This one is not confined to uncertainty reasoning, but it flourishes under uncertainty. The laurels for baiting this trap go to Steve Koonin who proposed a trial by combat, a Red Team Blue Team war game for settling the climate debate.



Steve Koonin April 20, 2017 WSJ
Director of the Center for Urban
Science and Progress, New York
University

“Here’s how it might work: The focus would be a published scientific report meant to inform policy such as the U.N.’s Summary for Policymakers A Red Team of scientists would write a critique of that document, and a Blue Team would rebut that critique.”

<https://www.wsj.com/articles/a-red-team-exercise-would-strengthen-climate-science-1492728579>

The climate scientists get the proof burden, they have to defend the climate policy. The Red Team only has to attack. The climate deniers don't have to write a scientific report explaining why increasing CO2 emissions won't cause harm. Image a soccer game played with these rules: Koonin soccer.

Koonin Soccer

Baiting
the trap



The Blue Team has a goal to defend, the Red Team has no goal to defend. Blue can't win. The idea of trial by combat goes back to the middle ages. In the absence of evidence, a fair fight lets god decide. Koonin soccer is not a fair fight.

Let god decide... Trial by combat
settling disputes in absence of evidence



FOX news was receptive to Trump's proof burden shifting with regard to the 2020 presidential election:



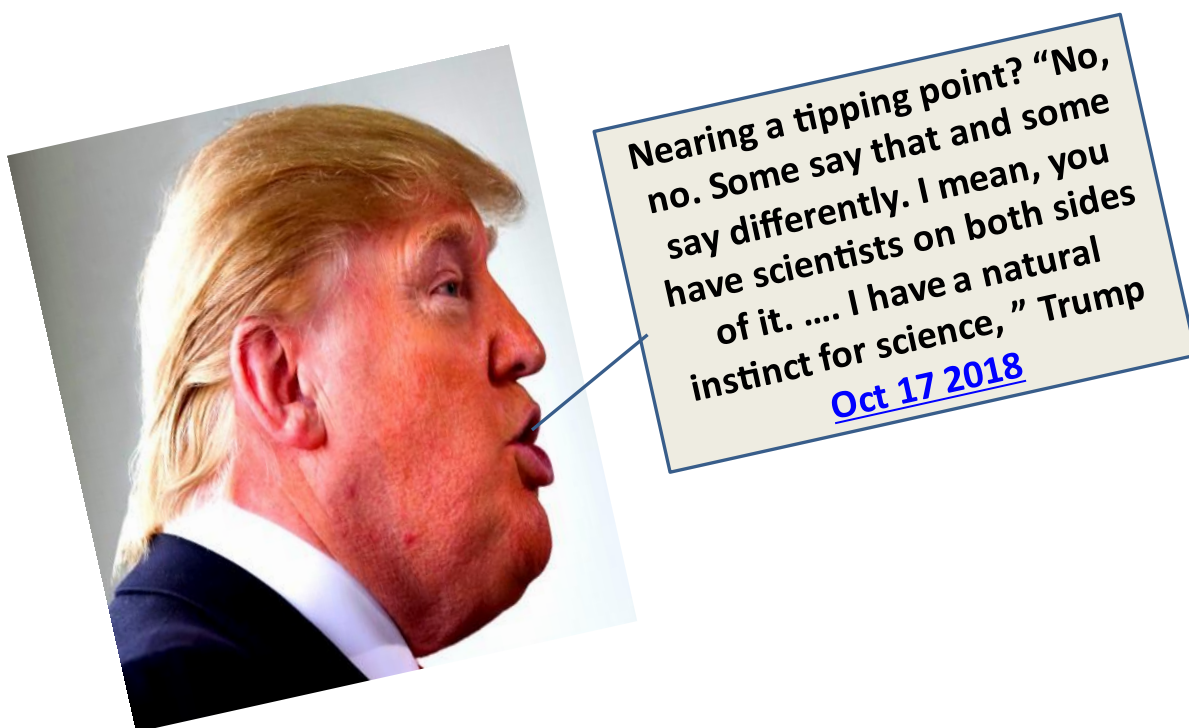
DONALD TRUMP · Published 3 days ago

Trump claims Biden must prove votes weren't 'illegally obtained' in order to enter White House

Trap 8

Disagreement trap thinking disagreement warrants dismissal

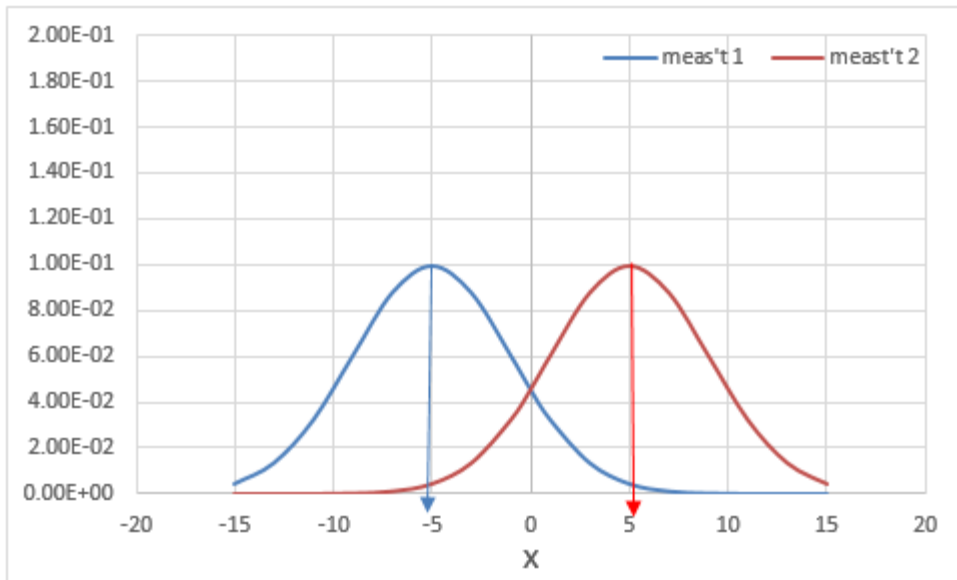
This one is very popular among climate deniers. Credit goes to Naomi Oreskes and Erik Conway (2010) for spotlighting this one¹. The treatment here focuses not on malign motives but on maladroit statistical treatment of discordant measurements..



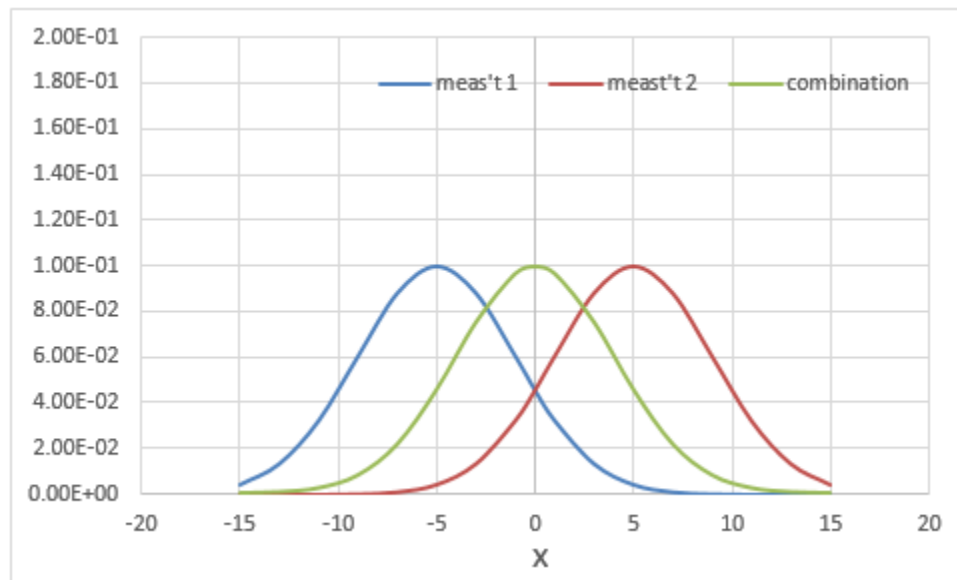
Oct. 17, 2018

The idea is simple. Measurement uncertainty often comes from noise in the measuring procedure. This can be studied by calibrating the instruments against things we know. Two different instruments for measuring the same thing can have different independent noise distributions. Their results can be combined even if they disagree. Suppose a Blue measurement returns the value -5 and a Red measurement returns the value 5. The noise of each distribution is shown below as bell curves. We see that the Blue and Red measurement have similar noise profiles centered at -5 and 5 respectively.

¹ [*Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. Bloomsbury Press. p. 6. ISBN 978-1-59691-610-4. merchantsofdoubt.org](#)



Donald Trump would say the scientists don't know what they're talking about so we can ignore both. A very simple analysis would say that, given both measurements, the most likely value is the average $(-5 + 5)/2 = 0$. The spread of the combined uncertainty is shown as the green bell curve below:



The combination bell curve shows uncertainty equal to that of the Blue and Red measurements, in spite of their disagreement. Why doesn't the Green uncertainty run from -15 to 15? This result is obvious for anyone with a "natural instinct for science". With only the Blue measurement the true value could with equal chance be either above or below the measured value -5. Similarly, with only the Red measurement the true value could be either above or below 5 with equal chance. However, given BOTH independent measurements, it is much more likely that the true value is above -5 and below 5. Doing the math, we get the green bell curve².

This simple analysis is in every elementary statistics text book, but it is often NOT correct. It assumes that the measurement errors are normally distributed and independent. We often measure proxies for the real

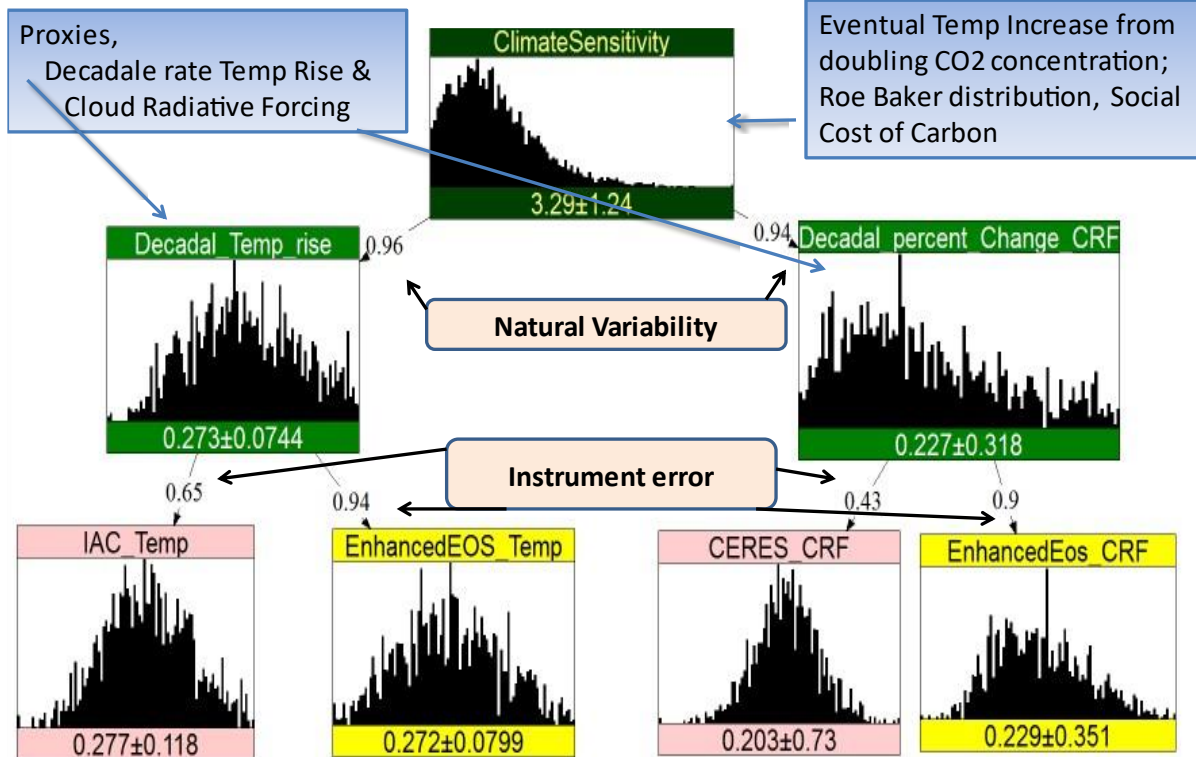
² The variance is the square of the standard deviation. The variance of the combination is the average of the variances of the red and blue measurements, which are both 4. The average variance is $(16+16)/2 = 16$, the standard deviation of the combination is 4.

variables of interest and these proxies are correlated by the fact that they are aimed at the same uncertain quantity (see Trap 10). For small positive uncertain values, the distributions are often right-skewed and thus not normal. The "prior uncertainty" of the unknown quantity is often quantified and must be taken into account.

The subtleties of merging discordant measurements were studied in research at Resources for the Future and NASA [on Probabilistic reasoning about measurements of equilibrium climate sensitivity: combining disparate lines of evidence](#). There is also a 15 min [video](#) on this. It concerns two proxies for Equilibrium Climate Sensitivity, being the amount by which the Earth's surface temperature would eventually rise as a result of doubling atmospheric CO2 concentrations. The proxies are Decadal Temperature Rise (DTR) and Percentage Change in Cloud Radiative Forcing (CRF). These proxies are measured by different independent satellites. According to the U.S. government approved models for [computing the Social Cost of Carbon](#), if we could measure either without error, we could then compute ECS. Unfortunately, we cannot measure these without error. Natural variability in the Earth's climate system causes the actual measured values of DTR and CRF to fluctuate around their long term trends. Natural variability includes things like the [ENSO cycles](#). In addition our satellite instrument are subject to measurement error caused by orbital decay, calibration drift and the like. All of these sources of error have been measured. These, together with a prior uncertainty on ECS approved for the computation of the [Social Cost of Carbon](#), enable us to compute the joint distribution of all variables. When we observe DTR and CRF for a number of years, we can estimate the long term trend with computable error, which decreases as the measuring time increases. We can then conditionalize, or update, our distribution on the observed values. NASA did these calculations in order to compute the expected dollar value of newly designed measuring systems for DTR and CRF (yellow, below) relative to that of the current measurement systems (pink, below). The goal is to demonstrate the value of information obtained by adopting the new systems. [Uncertainty always costs money](#). By reducing uncertainty in ECS we can take more effective adaptive measures and avoid wasting resources unnecessarily.

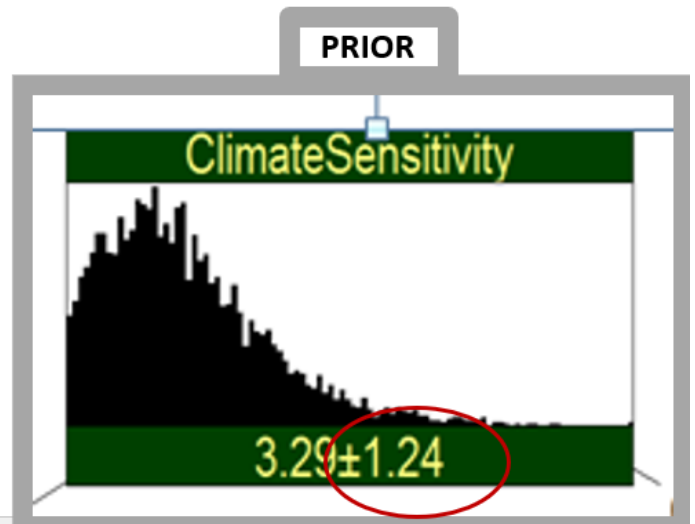
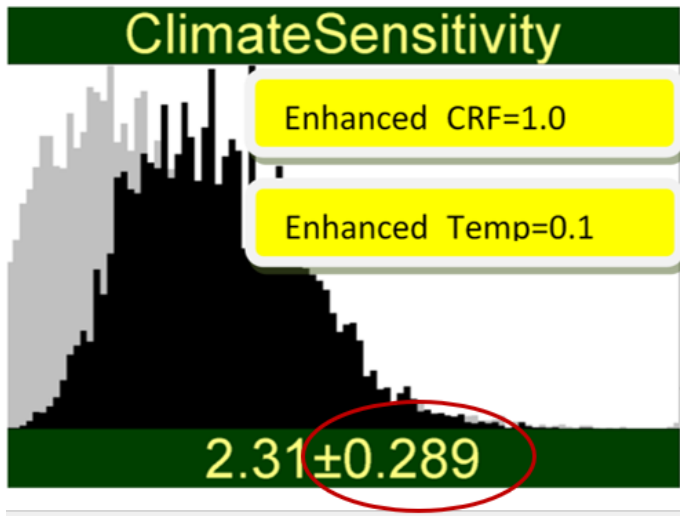
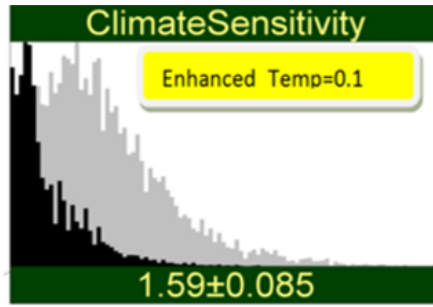
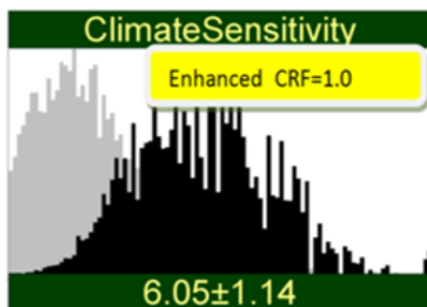
The picture below shows the joint distribution of all variables after launching in 2020 and observing out to 2050. Correlations induced by natural variability and instrument error are shown on corresponding arcs in the graph. If we measure any of the values on the bottom row, we can conditionalize the joint distribution on these observed values. By selecting possible measured values anno 2050 we visualize how discordant or concordant measured values affect our prediction of ECS.

Observations through 2050 (Business as Usual)



Below we see the results of observing only CRF in the enhanced system ($=1.0$) and only DTR in the enhanced system ($=0.1$). These values disagree strongly. The CRF measurement shifts the mean value of ECS from 3.29 to 6.05 with standard deviation 1.14. The DTR measurement shifts ECS down to 1.59 with standard deviation 0.085. The later value is low because ECS cannot be negative.

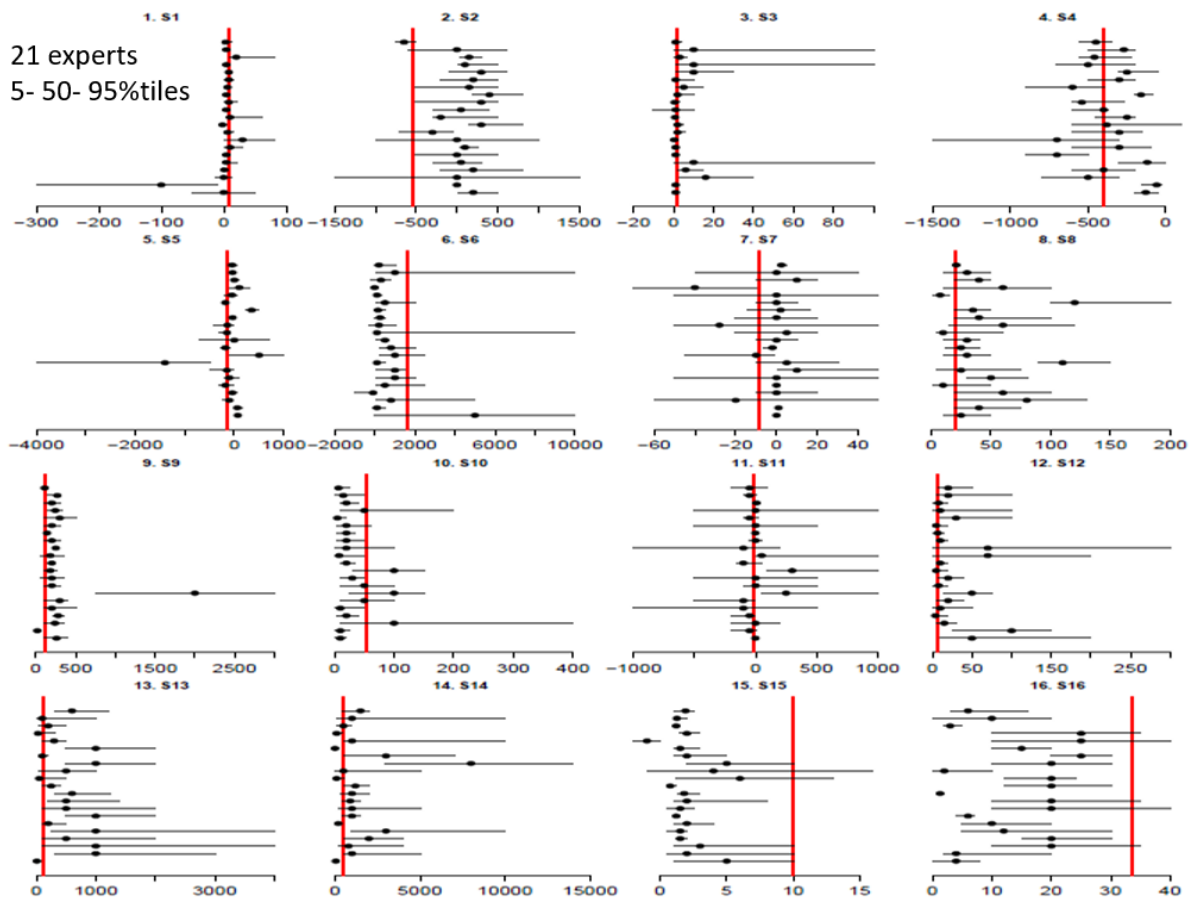
If we combine these two measurements then we predict a mean value for ECS of 2.31 and a standard deviation of 0.289. This is about $\frac{1}{4}$ of the prior standard deviation of 1.24. Combining discordant measurements leads to significant uncertainty reduction. Throwing away information from conflicting measurements would be profoundly maladroit.



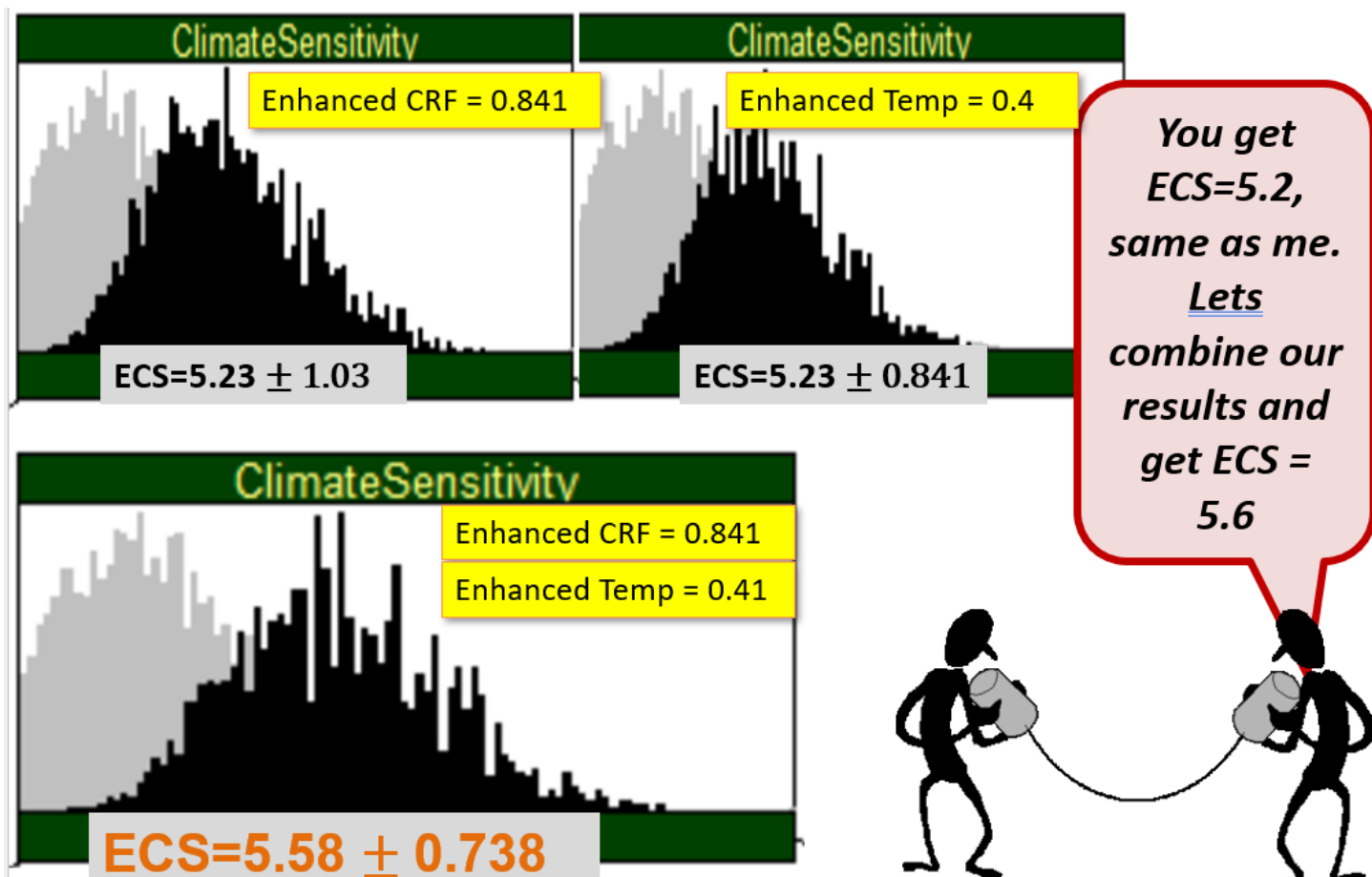
Trap 9

Agreement trap thinking agreement confers certainty

This one is very common and quite pernicious. Among experts, agreement is certainly no guarantor of certainty. The "range graph" below is from a 2019 [study in PNAS](#) of ice sheets contribution to sea level rise. 21 highly vetted experts answered 16 calibration variables from their field. For each variable the red vertical line denotes the true value, unknown to the experts. The experts gave their 5% – 95% confidence bands (horizontal black lines) and their 50% value (a black dot). Pictures like this, of which there are many, disabuse us of the notion that expert agreement, whatever that means, has predictive value.



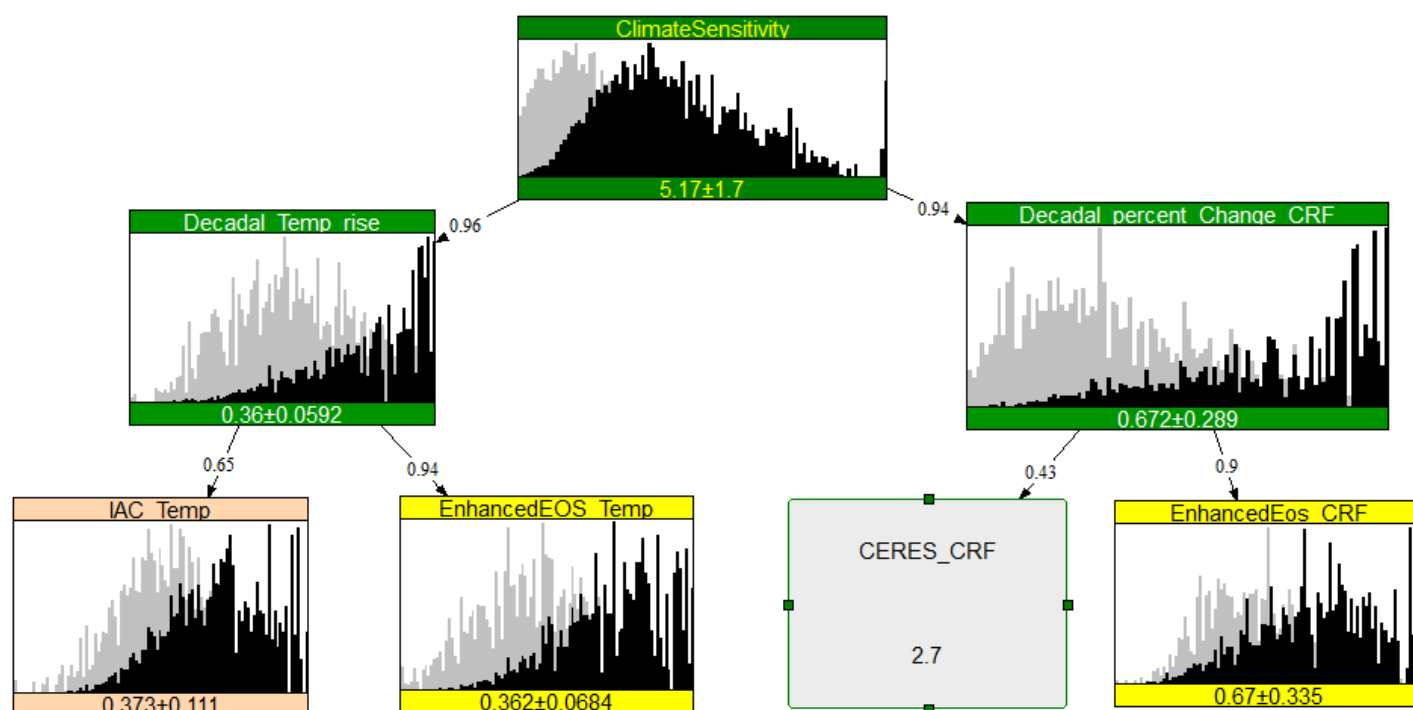
The ageement trap also concerns actual physical measurements. The graph below shows the results of measuring $CRF=0.841$ and $DTR = 0.4$ with the enhanced measurement systems. Coincidentally, they give the same prediction; the expected value of ECS after each measurement is 5.23. If we combine these measurements then the predicted value climbs up 5.58. Why? Each measurement drags the prior probability upward, though at different rates owing to their different error profiles. Note that the standard deviation after the CRF measurement is 1.03 whereas that of DTR is 0.841. By combining the measurements we combine their power to drag the prior distribution of ECS upwards, hence the prediction climbs from 5.2 to 5.6. At the same time the uncertainty goes down to 0.738.



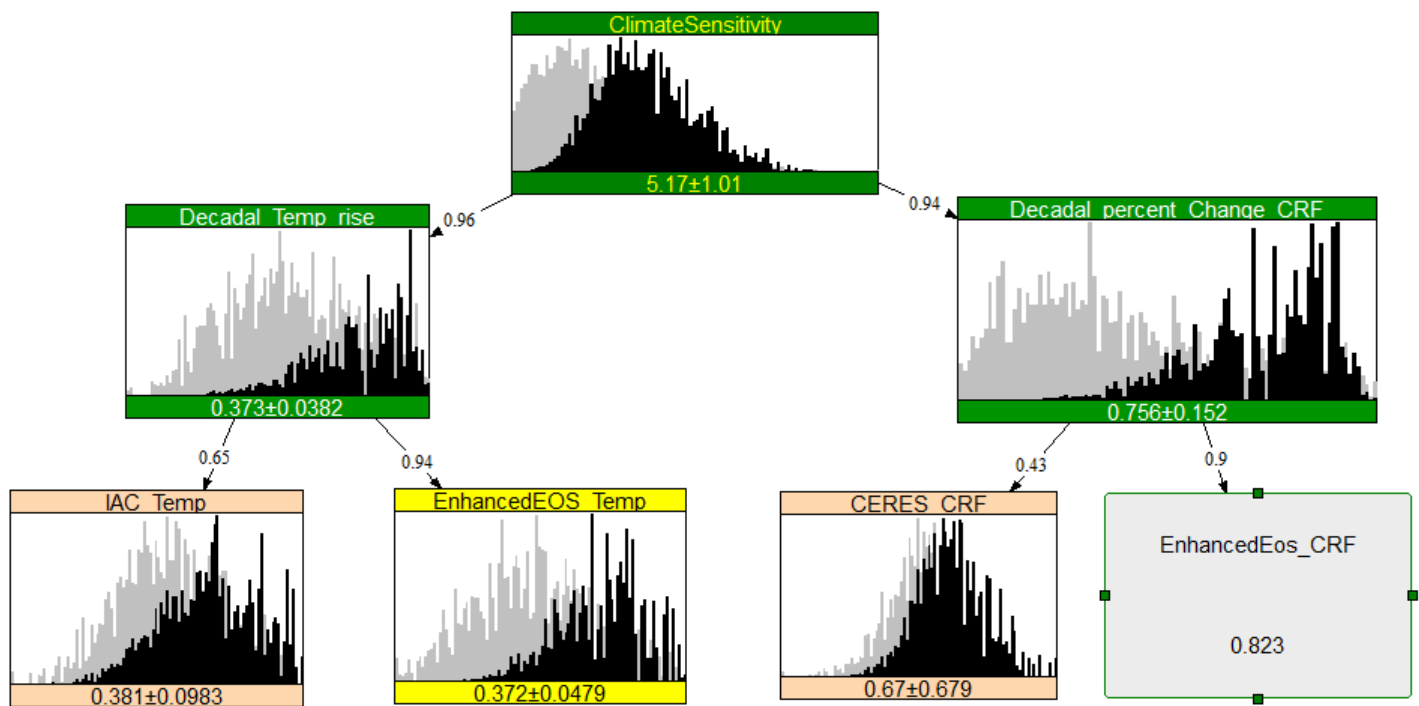
Such results are counter intuitive to folks who don't take the measurement errors and the prior distribution into account.

Trap 10

Negative Learning trap thinking new observations always reduce uncertainty



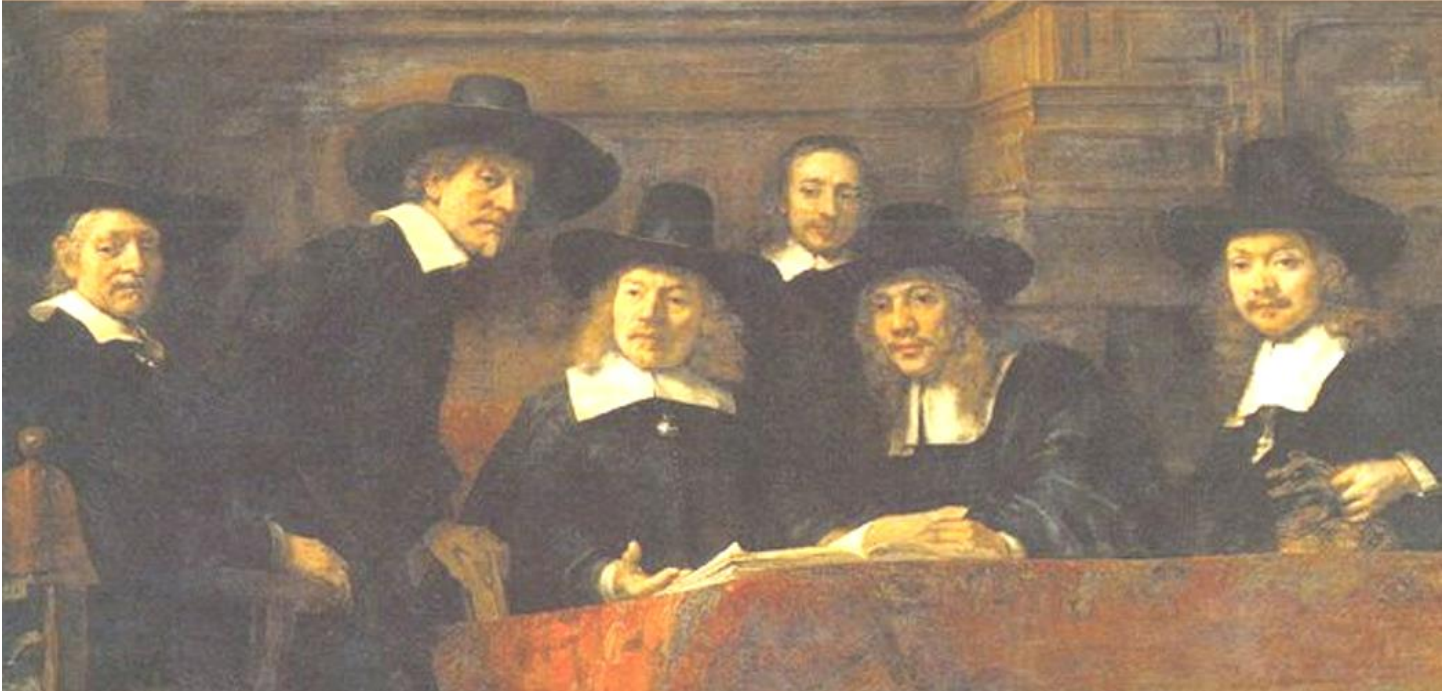
Here we show a situation in which the current measurement system for CRF returns a very high, but possible, value of 2.7 in 2050. Updating the joint distribution on this information drives all the other distributions upwards, even for un-preformed measurements of DTR and enhanced CRF. This shows that these measurements are NOT independent. They are correlated by the prior uncertainty on ECS. Note also that the uncertainty of ECS is now 1.7, up from 1.24 in the prior distribution. This is called negative leaning – after the observation we are more uncertain of ECS than we were before. The explanation is simple; conditionalizing on this very high value drags the prior distribution of ECS upwards and in so doing smears it out. Because of its relative large error it can't drag very effectively. The prior distribution persists somewhat. If its measurement error were smaller it would concentrate the distribution of ECS more tightly around the new measured value. Indeed, if we conditionalize the enhanced CRF on 0.823, we also predict a mean ECS of 5.17 but now with standard deviation 1.01. The negative leaning disappears.



Trap 11

The BOGSAT Trap (1961)

Uncertainty quantified by bunch of guys/gals sitting around a table



This is a very big one. When experts quantify their uncertainty in a structured expert judgment (SEJ) panel, the individual experts are validated against their answers on calibration questions from their field, as we saw in the Agreement trap. Their unvarnished uncertainties go straight to the problem owners. Committees of experts tasked with with "uncertainty characterizations" are never validated or held to account. They avail themselves of "calibrated uncertainty language" but their answers cannot be attributed to any specific member. The term BOGSAT was coined in derision and has had a checkered [history](#):

- First reference 1961
- Kent charts Calibrated Language 1964
- Calibrated Language adopted by US Defense Intelligence Agency 1976
- Calibrated Language abandoned by USDIA ~ 1978 for lack of validation
- Calibrated Language adopted by IPCC 2007
- Calibrated Confidence Language abandoned by IPCC 2014

In its Fifth Assessment Report (2014) the IPCC abandons the quantitative interpretation of confidence levels in the Fourth Assessment Report cited in the Confidence and Propagation traps.

IPCC 5th Assessment Report

Box 2.2 Quantifying Uncertainty

“Natural language is not adequate for propagating and communicating uncertainty ”

Confidence statements: “qualitative synthesis”

“qualitative synthesis of an author team’s judgment about the validity of a finding as determined through evaluation of evidence and agreement” Box TS1

The calibrated language of confidence statements is retained but shorn of any quantitative interpretation in favor of a "qualitative synthesis". How do we propagate syntheses through a chain of reasoning? Can we combine qualitative syntheses of combinations of separate findings or should we have a new qualitative synthesis for each combination of the thousands of confidence statements in the Assessment Report? The absence of guidance steers the unwary reader to believe that no guidance is necessary. BOGSAT rules.

This brings us to the final and most pernicious trap.

Trap 12

Communication Trap

trying to communicate what you don’t understand

We illustrate the Communication Trap with the IPCC's characterization of uncertainty in Equilibrium Climate Sensitivity (ECS). To recall ECS is the amount by which the Earth's surface eventually warms as a result of doubling the atmospheric concentration of CO₂. For context, celebrated climate scientist David Archer writes "A climate change of ... 5~6° C, would be catastrophic to human civilization." (David Archer (2009) *The Long Thaw*, p. 95). The last time the Earth was that warm, during the Paleocene–Eocene Thermal Maximum (PETM 56 million years ago), featured crocodiles on Spitsbergen.

According to the 5th Assessment report ECS is

“...extremely unlikely [<5%] less than 1°C (high confidence), very unlikely [< 10%] greater than 6°C (medium confidence).” (AR5 p.1111)

This says that a global catastrophe ($ECS > 6^{\circ}C$) has less than a one-in-ten probability of occurrence, and in that probability statement we have medium confidence. "Medium confidence" previously meant "about a 5 out of 10 chance" that the probability of catastrophe was greater than 10%. Now it means...? At least it means that the statement " $Prob(ECS > 6^{\circ}C) < 10\%$ " has declarative meaning, it can be either true or false. We can't have medium confidence in a meaningless string of words. The authors invest some degree of belief in " $Prob(ECS > 6^{\circ}C) < 10\%$ " but decline to say how much. Although they don't say so, everyone will infer that "high confidence" indicates a higher degree of belief than "medium confidence" in said statement. Were that not the case then their choice of words would be deception bordering on lying.

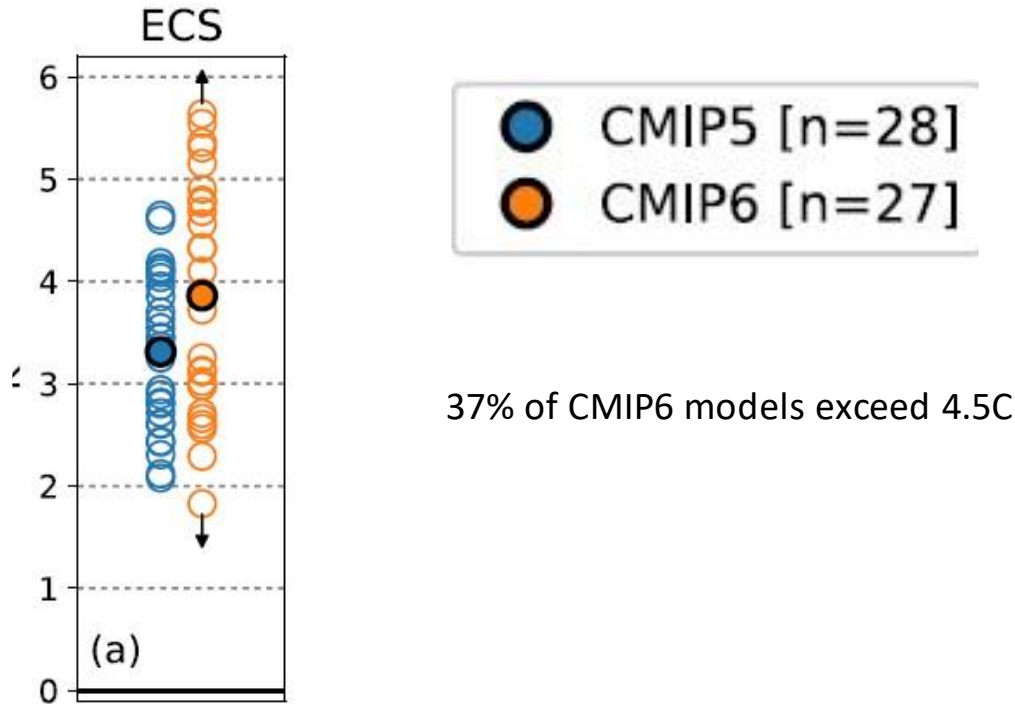
One is then left wondering why the degree of belief is lower on the high end ($> 6^{\circ}C$) than the low end ($< 1^{\circ}C$). Given that our very civilization hangs in the balance, one cannot help but ask how high that $6^{\circ}C$ would have to climb to garner the same high confidence they claimed at the low end. Do the IPCC authors have high confidence that the probability that ECS exceeds $9^{\circ}C$ is less than 10%? Alas the BOGSAT is adjourned.

Lets translate this uncertainty characterization to a context a bit closer to home:

Would you put your kids on an airplane in which authorities had medium confidence that the probability of a crash was less than 10%??



The 6th Assessment Report ran updated models in its Climate Model Intercomparison Project (CMIP6). Here's how they compare to CMIP5:



Very simply, if statements A and B are uncertain then also 'A and B' is uncertain³. You cannot characterize the uncertainty of 'A and B' simply by characterizing the uncertainty of A and the uncertainty of B. If you don't understand that, then you don't understand uncertainty. Many of the foregoing traps, eg the certainty trap, the confidence trap, the propagation trap, the chaining trap, the procrastination trap and the innumeracy trap, have the effect under estimating uncertainty. Does this improve decision making?

Saying that the communicators don't understand uncertainty is the charitable hypothesis. The alternative is that they DO understand but are deliberately teaching the public to reason incorrectly under uncertainty. That would be "libertarian paternalism" (see prologue) gone seriously sideways. Instead of preventing human fallibility it would be deliberately promoting it.

³ Unless of course A entails not-B.