Uncertainty attitudes as values in science

or: Does risk aversion undermine scientific objectivity?

1 April 2022

Bern

Joe Roussos

Institute for Futures Studies

This paper is about values in science and objectivity

- Objectivity serves as an aspirational goal for scientists
  - Science describes objective features of the world
  - Scientific results should be usable by anyone, regardless of their values
  - Science advisors should not engage in politics, but should support political decisions
- Many have thought that, to secure objectivity, we need value-free science
- I am interested in what exactly counts as a "value" when we talk about values in science, value-freedom, and objectivity
- My initial interest in this came from science-advising, but today's presentation is about core activities of scientific research, with a focus on climate science

Outline

- Inductive risk, through a decision theory lens
- Risk aversion as a value in science
- Does risk aversion challenge objectivity?
- Conclusion and way forward

Outline

- **Inductive risk, through a decision theory lens**
- Risk aversion as a value in science
- Does risk aversion challenge objectivity?
- Conclusion and way forward

Why worry? Democratic argument for value-free science

Claim: Objective science is required for decisions to be democratic

- In order to be democratic, decisions must respond appropriately to the values of the people (represented by the policymaker, political decision makers, etc)
- Scientific input is required in policy, as we want our policies to respond to how the world really is
- If science is value-laden then it “bakes in” some value judgements which are not those of the people, and these can influence which policy decisions are made
- This subverts democratic control, as scientists’ values dilute or supplant the values of the people
- Result: the “liberal epistemic division of labour” – scientists provide the facts, politicians provide the values
- Relevant notion of objectivity is freedom from values which would interfere with democracy

It is standard to consider categories of values in this literature

- ✓ Epistemic values: truth, accuracy, perhaps others
  - These are the core values for the scientific enterprise
- ✓ Cognitive values: simplicity, explanatory power, breadth, fruitfulness
  - These are properties that scientific theories and models might have, which help with the epistemic aims of science by making them easier to work with for beings with our cognitive capacities
- ✗ Moral values: wellbeing, happiness, just distributions, rights
  - These are classic practical values, critical to scientifically-informed policy and which should of course govern research (so that it is not itself unethical)
- What I have called moral values are the ones at issue in these debates
  - Moral values thought to interfere with democracy
  - Stronger forms of value-freedom call for no moral values whatsoever in science

Inductive risk, through a decision theory lens

- Science involves decisions, about questions, methods, analyses, representations, hypothesis acceptance, coarse-graining, and many more
- Decisions are a function of beliefs and preferences, or evidence and values
- The value-free idealist hopes that these decisions can be made with epistemic and cognitive values only
- Rudner (1953) argues that they must involve moral evaluations of the badness of error (inductive risks)
- Jeffrey (1956) responds that scientists shouldn't make decisions, but rather report probabilities
- Douglas (2009) argues that scientists face many decision points prior to final probabilities, and ought to weigh moral consequences when making them

A quick example: glacial melt rate

Claim: deciding to accept a hypothesis is value-laden

- Hypothesis R: melt rate is above relevant threshold *r*

	R is true, $\Pr(R) = 0.85$	R is false, $\Pr(\neg R) = 0.15$
Accept R	True Pos: Accept R and R true, dam will burst if unaddressed	False Pos: Accept R, but R false. Dam gets replaced unnecessarily
Reject R	False Neg: Reject R, R true. Dam bursts without warning, village destroyed	True Neg: Reject R, R false. No change, everyone safe

- Should R be accepted? i.e., is 85% certain enough?
- Claim: no purely scientific or epistemic answer to this question
- Scientists must consider the moral value of the outcomes above in deciding whether to accept R
- Highly stylised! Actual scientific decisions typically much further from "real-life" impacts and decisions are made with an eye to nearby instrumental goals

Philosophers have traced a range of more realistic cases

- Examples in climate science
  - Extreme event attribution methodology debate between storyline approach and fraction of attributable risk is in fact about how the two camps evaluate the risks of errors, where that evaluation is a moral matter (Winsberg, Oreskes, Lloyd 2020)
  - Choice of methods in adaptation studies should incorporate the values of the users of the climate service (Lusk and Parker 2019)
  - Representations of uncertainty in climate models do and must reflect scientist values (Winsberg 2010, Winsberg and Biddle 2012)
  - Model construction and evaluation reflect the priorities and purposes of the scientists involved (Intemann 2015)
- Often framed in terms of inductive risk (though not always)

Outline

- Inductive risk, through a decision theory lens
- **Risk aversion as a value in science**
- Does risk aversion challenge objectivity?
- Conclusion and way forward

Let's start with an aside about language: "risk"

- In English, "risk" is ambiguous between several meanings
  1. An unwanted event: the risk of getting cancer
  2. Probabilities of unwanted events: the risk that a smoker's life is shortened is 50%
  3. Expectation values of unwanted events (risk analysis): the risk of smoking is 12.5 life-years lost on average
  4. Spreads of possible events (economics): B is riskier than A
     - A: a lottery ticket which pays out 100kr for sure
     - B: a lottery ticket which pays out 0kr with 50% chance, and 200kr with 50% chance
- "Inductive risk" is a name given to the errors—FN and FP
  - When one makes an inference, one risks making these errors
  - Typically this is a usage of risk in sense 1 above

The IR discussion is limited to certain kinds of evaluations

- The values that are discussed in this debate take as their objects concrete outcomes—states of the world described without reference to chances
- E.g., The act of accepting the glacial melt rate hypothesis is evaluated in terms of the consequences of that act: TP, FP, FN, TN
  - These lead to downstream acts such as the building of a new dam, and those acts are in turn evaluated in terms of their consequences
  - Base of the evaluation: the goodness or badness of the state of the world in which people die due to a dam bursting
- Contrast this with attitudes to uncertainty: evaluations of the state of uncertainty under which the decision is made
- Little to no discussion of this in the values in science literature (to my knowledge!)

What are uncertainty attitudes?

- Risk aversion: the tendency of people to prefer decision situations with low uncertainty to decision situations with high uncertainty
- Classic economics e.g., consider these two bets
  - A. 100% chance of receiving 100 kr
  - B. 50% chance of receiving 0, 50% chance of receiving 200 kr
  - Assume for the moment that utility is linear in money
- These have the same expected utility, and so orthodox EU-maximising agents ought to be indifferent between them
- People who prefer A to B are risk averse
- Risk aversion changes how people make decisions, in addition to showing up as preferences between kinds of gambles

Note: In the paper I discuss two forms of uncertainty attitude: risk attitudes and ambiguity attitudes. In the talk I focus on risk aversion for simplicity, but my arguments really concern the whole family of uncertainty attitudes

Risk attitudes plausibly encode the value of chances

- Large literature on what risk aversion is, e.g., Stefánsson & Bradley (2019) and references therein
- On the view I favour, risk aversion is an aversion—a desire-like state, representative of an evaluative attitude
- Some candidate evaluative contents…
- Value of certainty: The agent values being in a state of complete certainty, disvalues being in states of uncertainty
- Value of having a chance: The agent values having a chance of winning a prize, over and above the value of the prize itself
  - Or perhaps more familiarly, despite getting a bad outcomes, the agent values having had a chance of getting a good outcome

Risk aversion makes a difference

- People have risk attitudes—there is a large empirical literature on this
- Agents with the same beliefs and the same desires, but different risk attitudes, will make different decisions
- If decisions can differ due to risk attitudes, then a fortiori so can scientific decisions
  - E.g., two scientists in the glacial melt case
  - All the same studies, same beliefs (same probabilities)
  - Due to AIR, they identify downstream consequences, and they evaluate them identically
  - Nevertheless one accepts, other rejects due to a difference of risk attitude
  - (In the paper I show that this occurs according to several theories of rational choice, for both risk attitudes and ambiguity attitudes)

How much of a difference?

- A typical scientific report or piece of expert advice will be the result of a sequence of many such decisions
  - Select a research question,
  - design a study,
  - select a method,
  - …,
  - classify data,
  - choose analysis techniques,
  - … ,
  - choose how to communicate a result
- If each faces the epistemic gap identified in the AIR, then in each case the scientist makes a practical decision and their uncertainty attitudes matter
- Cumulatively, over the sequence, these attitudes could make a big difference

This is just the "core" of science. Uncertainty attitudes will also matter a lot in raising the alarm about a potential policy issue

Risk aversion is a distinct evaluative stance affecting science

- Claim: we should consider risk aversion as a value affecting science
- Now, I said that this is new, but isn't the whole discussion about avoiding risks?
- Recall the ambiguity in the English term "risk". The IR discussion focusses on the value of the state of the world in which one has made the inductive error. This is a use of risk in the sense of bad outcome
- That is not what is at stake in risk aversion. RA agents aren't just valuing certain outcomes negatively, they are responding to the spread of outcomes across the possible states of the world
  - E.g., Biddle and Kukla (2017) offer a very wide-ranging typology of "phronetic risks", categorising kinds of risks. But there is no mention of attitudes to risk in my sense
  - E.g., Winsberg (2018) discusses uncertainty attitudes in the context of climate policy and IAMs in ch 8, but they do not show up in ch 9 on values in science
- My claim: wherever you worry about values in science, you should in the same breath talk about uncertainty attitudes

Outline

- Inductive risk, through a decision theory lens
- Risk aversion as a value in science
- **Does risk aversion challenge objectivity?**
- Conclusion and way forward

Uncertainty attitudes interfere with learning about the world

Moral disagreement

- Suppose Karim consults with two experts on the rate of agricultural technology development
- They have all the same evidence and the same prior empirical beliefs
- Still, Karim receives two different answers
- The difference is because they differed in their moral evaluations of the badness of the potential errors they might make
- This seems bad for Karim's project of learning facts about the world

Value-free science, OR
Use the right values

Risk-attitude disagreement

- Suppose Karim consults with two experts …
- They have all the same evidence, the same prior empirical beliefs, and evaluate the outcomes identically
- Still, Karim receives two different answers
- The difference is because one scientist is more risk averse than the other
- This seems bad for Karim's project of learning facts about the world

Risk-attitude free science, OR
Use the right values

Secondary issue: patterns of risk attitudes among disciplines

- One risk seems to be this: it is common to hear talk about academic disciplines inculcating certain risk attitudes
- Extreme e.g., safety engineers designing systems for air traffic or road traffic are institutionally risk averse—which is to say that they take low probability events very seriously, rather than acting on the expected outcome
- Or, a stereotype: economists are risk-seeking
- Information from different fields would not be equivalent and may be hard to compare and combine
- What does it do to our whole system of science if there are field-specific risk attitudes?
  - If they were all deliberate, as in the case of safety engineering, it might be fine—reflecting a societal agreement about areas where we want risk aversion
  - But if there are subconscious and unplanned patterns resulting from self-selection, as is perhaps the case with economists, it is potentially worrying

Response 1: Jeffrey's reply to Rudner

- "This is one more reason to be Bayesian! Merely report probabilities and avoid making decisions, which would be tainted by uncertainty attitudes"
- Rebuttal: Jeffrey's response seems ill-suited to modern, complex science
- Climate science is highly collaborative, relies on complex simulation models, and so any probabilities which are produced are not plausibly the credences of a specific scientist
- They are at best probabilities that a scientist endorses, and they were calculated and constructed on the basis of many decisions
    - Methods selection, model construction, weighing evidence, weighing different models, etc.

Response 2: Uncertainty attitudes are harmless values

- Recall that we have two large categories of acceptable values
- Cognitive values: facilitate thinking, understanding for agents with our cognitive capabilities
  - Examples I gave: simplicity, explanatory power, breadth, fruitfulness
- These concern properties of theories and models which facilitate our using them, so uncertainty attitudes don't seem to be candidates
- Cognitive values have a clearly positive valence, a direction which facilitates the epistemic mission of science; uncertainty attitudes don't seem to fit this pattern: none of risk aversion, neutrality, or seeking facilitates discovering truths

Response 2: Uncertainty attitudes are harmless values

- Epistemic values: constitute the truth-seeking mission of science
- I suggested that risk aversion reflects valuing certainty, isn't certainty an epistemic state?
- RA agents prefer decision situations where benefits and harms are "bunched up" rather than "spread out"—it is an attitude towards decision situations and patterns of outcomes, not states of knowledge
- RA agents can sometimes rationally turn down free information (Wakker 1988, Buchak 2010), though this is in very particular situations

Response 3: Isn’t there an obvious solution? Risk neutrality

- It might seem that there is an obvious solution: risk neutrality as value neutrality
- **Problem:** the name is misleading
- There is nothing neutral about risk neutrality
- Risk neutrality is an evaluative stance
    - It evaluates ($100 for sure) as equal in value to ($0, 0.5 ; $200, 0.5)
- Strong value-free ideal is that scientists should make no moral evaluations whatsoever
- Risk neutrality is not analogous to strong value-freedom
- Alternative lines of argument:
    - There is a risk attitude which does not interfere with democracy
    - There is an morally correct risk attitude for scientists to take
    - Separate question whether either of these is risk neutrality

Uncertainty attitudes are moral values

- In a policy context, scientists are involved in decisions made on behalf of others
- Are the uncertainty attitudes of social decision makers relevant?
- Ethicists say: plausibly, yes
  - Buchak (fc): Specific attitudes are morally required—social decision-makers should be risk avoidant but ambiguity neutral
  - Stefánsson (fc): Social decision makers ought to be more risk seeking than individuals would be if they were making the decisions themselves
- At minimum, uncertainty attitudes matter to moral decision making
  - The uncertainty attitudes of scientists may interfere with democratic decision making, just as their moral attitudes might
- At maximum, these are moral evaluations, e.g., of the value of having a chance of some good, like surviving or having clear air

Uncertainty attitudes interfere with stakeholder values

- The key point is that decisions aren't a only function of evidence and values
- These are filtered through attitudes to uncertainty when a decision is made
- Consider the "stakeholder engagement" model of managing values in science
  - E.g., Lusk and Parker (2019: 1647): "if choices must be made, they could be made in light of
  - the inductive risk preferences of the user or client: if it would be particularly bad for the user's purposes for uncertainty to be underestimated, then the provider might select those methodological options that will deliver a broader uncertainty estimate."
- The idea is roughly that scientists should act on behalf of users, so that the science reflects what they would do if they were in the scientist's epistemic position
- But incorporating user values while neglecting uncertainty attitudes will lead to decisions which still diverge from how stakeholders would make them if they could

Outline

- Inductive risk, through a decision theory lens
- Risk aversion as a value in science
- Does risk aversion challenge objectivity?
- **Conclusion and way forward**

So, where does that leave us?

- Some open options:
- Perhaps there are procedural constraints which should guide scientists uncertainty attitudes, e.g., stakeholder engagement and adopting user attitudes
- Perhaps disciplines should all have disciplinary attitudes, as safety engineers do, according to their position in society
- Perhaps there is a risk attitude which does not interfere with democracy
- Perhaps there is an morally correct risk attitude for scientists to take
- Could be that risk neutrality fulfils either or both of these, but that requires separate argument
- There will also be questions about how to aggregate such values

Conclusion

- The current discussion is philosophy of science is disconnected from both
  - the decision theoretic discussion of uncertainty attitudes and
  - the social ethics discussion about the moral status of uncertainty attitudes
- The aim of this paper is to establish this claim: we should pay attention to how the uncertainty attitudes of scientists influence their output in the same way that we pay attention to how their moral evaluations influence their output

Thanks!

Let’s see how risk attitudes make a difference in a toy example

- I will use the risk-weighted expected utility theory (REU) developed by Lara Buchak, but nothing here depends on this being a good/true theory
- To start: risk-sensitive agents reason in terms of the levels of benefit they can secure, and consider probabilities of exceeding a given level of benefit
- So far, this is formally equivalent to EU, it just focuses on exceedance probabilities rather than probabilities of outcomes

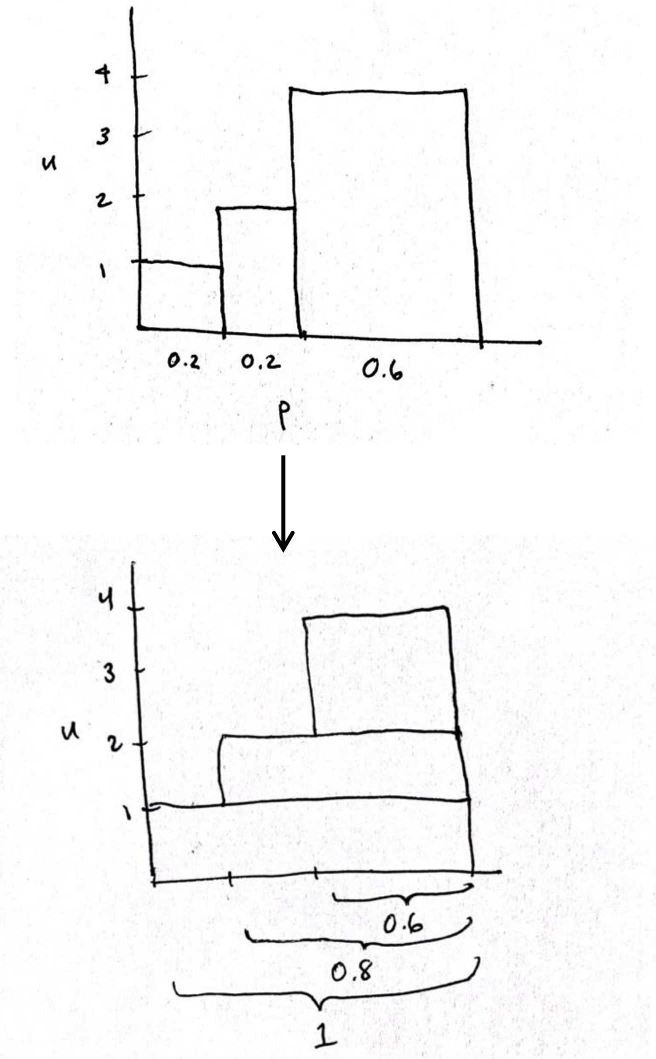
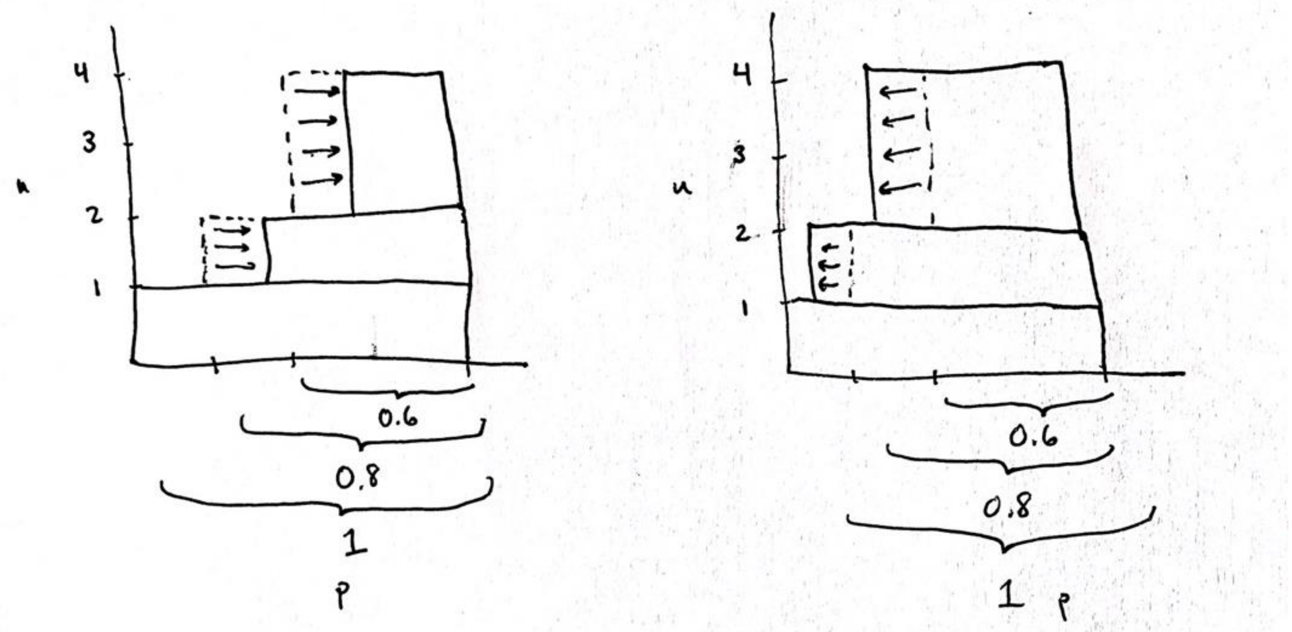


Image credit: Lara Buchak

Risk attitudes are represented by a function of the probabilities

- However, instead of valuing a level of benefit at its probability-weighted value, risk-sensitive agents weigh levels of benefits using a function of probabilities
- This is the risk function $r(p)$, and it is applied to exceedance probabilities
- Risk-averse agents have convex risk functions, e.g., $r(p) = p^2$
- Risk-neutral agents have $r(p) = p$
- Risk-loving agents have concave risk functions, e.g., $r(p) = p^{1/2}$



Toy example: New crop yields

- Two scientists have carried out the same studies regarding hypothesis H: that agricultural yield of a new genetically modified crop is above some threshold X
- Each faces the same decision whether to accept. Consider this as a decision under risk (i.e., assume they have quantitative statistics)
- Suppose that their epistemic and cognitive values don't determine the decision and so, by the AIR, they evaluate the outcomes of their decision using their moral values

	$P(H) = 0.6$	$P(\neg H) = 0.4$
Accept H	2	0
Reject H	1	1

- Suppose that they agree on the moral evaluations, so they both face the decision in the table above

	$P(H) = 0.6$	$P(\neg H) = 0.4$
Accept H	2	0
Reject H	1	1

The risk-averse scientist rejects *H*

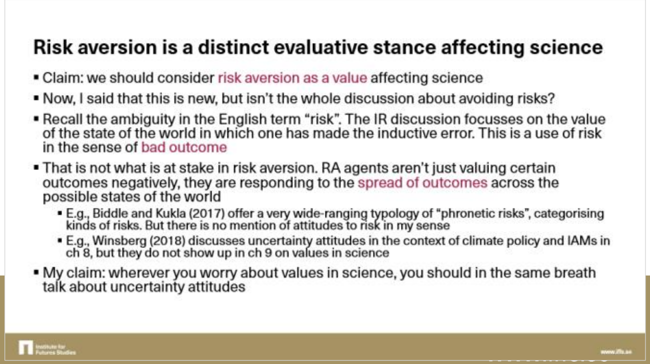
- They reason in terms of the levels of benefit they can secure, and consider probabilities of exceeding a given level of benefit.
- Consider Accept.
  - Worst outcome: 0. Get this no matter what happens, so exceedance-probability is 1.
  - Next-best: 2. How likely are they to get *at least* this level? 0.6.
- These are then weighted using the risk function $r_2(p) = p^2$
  - The worst-case is weighed fully, since $r(1) = 1^2 = 1$.
  - The next-best outcome is discounted, since $r(0.6) = 0.6^2 = 0.36$.
- We then calculate the risk-weighted expected utility, by multiplying these risk-weighted probabilities by utilities and summing.
  - $REU(A) = 0 \cdot 1^2 + 2 \cdot 0.6^2 = 0.72$
  - $REU(R) = 1$
- So, they Reject.

The risk-neutral scientist accepts *H*

- Scientist 1 is risk neutral, so $r_1(p) = p$

	$P(H) = 0.6$	$P(\neg H) = 0.4$
Accept H	2	0
Reject H	1	1

- For risk-neutral scientist 1, REU=EU, and it is easy to see they should Accept
- They had all the same data, and even agreed on the evaluations of all the outcomes
- Nevertheless, they arrive at two different scientific decisions
- Correspondingly, we the users see scientific disagreement
- What's at stake: their pure attitudes to risk



Risk aversion is a distinct evaluative stance affecting science

- Claim: we should consider risk aversion as a value affecting science
- Now, I said that this is new, but isn't the whole discussion about avoiding risks?
- Recall the ambiguity in the English term "risk". The IR discussion focusses on the value of the state of the world in which one has made the inductive error. This is a use of risk in the sense of bad outcome
- That is not what is at stake in risk aversion. RA agents aren't just valuing certain outcomes negatively, they are responding to the spread of outcomes across the possible states of the world
  - E.g., Biddle and Kukla (2017) offer a very wide-ranging typology of "phronetic risks", categorising kinds of risks. But there is no mention of attitudes to risk in my sense
  - E.g., Winsberg (2018) discusses uncertainty attitudes in the context of climate policy and IAMs in ch 8, but they do not show up in ch 9 on values in science
- My claim: wherever you worry about values in science, you should in the same breath talk about uncertainty attitudes

Brief aside: the same thing happens with ambiguity

- E.g., (Lusk and Parker 2019) Scientists will monitor concentrations of toxic air pollutants from vehicle exhausts in an outdoor shopping district; their findings will help the local government decide when and how to reduce traffic in the area
- Choice: install a type-A or type-B monitoring instrument
  - Neither is thought to be clearly more accurate or reliable than the other, but there is some evidence that they have different FP and FN error rates, though these cannot yet be quantified
- Inductive risk claim: scientists should/can/will consider downstream consequences
  - Underestimating high concentrations can be expected to result in insufficient traffic reduction and adverse human health effects, while overestimating concentrations can be expected to result in unnecessary traffic reduction and economic losses
- This is a decision under ambiguity: no quantified probabilities available. Perhaps representable as ranges of probabilities with vague boundaries
- Analogous attitudes to uncertainty: ambiguity aversion, ambiguity neutrality
- Again, scientists who differ only on these attitudes will make different choices