Is it rational to have faith?

Looking for new evidence, Good's Theorem, and risk aversion.

Lara Buchak

UC Berkeley

buchak@berkeley.edu

What is Faith?

Typical faith ascriptions:

- "I have faith that God exists."
- "He has faith that his spouse isn't cheating on him."
- "She has faith that her car will start when she leaves for work this morning."
- "I have faith in you."
- "He acted on faith."
- "It was an act of faith."

Faith ascriptions typically involve a proposition to which the actor acquiesces.

The truth or falsity of this proposition is typically a matter of importance to the actor.

- # "I have faith that the Nile is the longest river in Egypt."

Faith typically involves an action.

- Faith can be context-dependent: I have faith that my car will start when what is at stake is being late for work, but I don't have faith that my car will start when what is at stake is carrying a life-saving organ to the hospital.

What is Faith?

We make assertions of faith only when the outcome of the proposition involved is yet uncertain, or our evidence doesn't yield complete certainty.

- "I have faith that you passed the exam." (when your friend is worried about the outcome of an exam)
- # "I have faith that you passed the exam." (when she shows you that she got an A)
- # "I have faith that 2 + 2 = 4."

It is possible to make faith assertions if we had no evidence for the proposition, or strong evidence against it.

- Although she's spilled all the secrets I've told her so far, I have faith that this time will be different."
- "I don't think there's any evidence that God exists, but I have faith that God exists."
  - (But we may be inclined to think these are cases of misplaced faith.)

First pass at an analysis

- A person has faith that X, expressed by *A*, **only if** that person performs act *A*, and *A* is better if X is true than it is if X is false, and there is some alternative act *B* that is available to the person such that *B* is better than *A* if X is false.
  - Captures the idea that the truth or falsity of X is a matter of importance to the actor, and the idea that doing A must be based on X in some way.

- Derivative uses:
  - A person p has faith in another person q if and only if there is some act *A* and some proposition(s) X that is (are) *about* q such that p has faith in X, expressed by *A*.
  - A person performs an act of faith (or acts on faith) if and only if he performs some act *A* such that there is a proposition X such that he has faith in X, expressed by *A*.
- What we're still missing: faith in X seems to involve going beyond the evidence in some sense.

Epistemic Rationality

- Proportion one's beliefs to the evidence.
  - Includes updating only in response to new evidence.
- Have consistent degrees of belief.
- Don't be dogmatic. $p(X)$ can't be 0 or 1 except for logical or analytic truths.

Faith and Evidence I

- First thought: Faith in X requires believing X to a higher degree than one thinks is warranted by the evidence.
  - Moorean problems; it's hard to actually have faith in this sense.
  - Also, clearly irrational.
- Second thought: Faith in X requires being determined not to abandon one's belief in X under any circumstances.
  - Seems too strong in religious case and in ordinary friendship cases.
  - Distorts the phenomenology of acting on faith: acting on faith feels like a risk, but if one is certain that X is true, then doing A isn't a risk at all.
  - Incompatible with the phenomena of losing faith: on this picture, if one abandons one's faith, one never had faith in the first place.
  - Again, irrational: seems to require assigning $p(X) = 1$.
- Third thought: Faith in X requires treating the circumstances in which one would abandon one's belief in X as epistemically impossible.
  - Better, because it seems to explain how a person could lose faith: he could find himself in circumstances he previously took to be epistemically impossible.
  - But still too strong; I could have faith in X while still recognizing that ~X is compatible with my evidence.
  - Incompatible with realizing at the present time that one might later lose one's faith.
  - Again, conflicts with Bayesian rationality: seems to require having degree of belief 0 to encountering evidence against X. E.g. requires that if one holds that $p(X \mid C) < p(X)$, then one must hold that $p(C) = 0$. => $p(X) = 1$.

Faith and Evidence II

- Better thought: Faith in X requires *not actively looking for evidence for or against X.* That is, not actively engaging in an inquiry whose only purpose is to figure out whether X is true or false.
  - Example: Hiring a private detective to investigate one's spouse.
  - Example: Being presented with an envelope which one knows contains evidence for or against one's spouse's constancy.
  - Example: Doubting Thomas.
- Doesn't entail assigning p(X) = 1 or having degree of belief 0 in encountering evidence against X.
  - Not actively looking for evidence for or against X is compatible with thinking that evidence against X might be out there. It is also compatible with accidentally coming across evidence against X, and then changing one's mind about what to do on the new evidence (losing faith).
- That a person could lose his faith in response to evidence seems entirely compatible with the claim that he really has faith; but that he actively seeks out evidence that he knows might make him revise his beliefs does not.
- This also explains why we think of faith as involving a commitment: it involves a commitment to not look for more evidence.

Faith

- A person has faith that X, expressed by *A*, only if that person performs act *A*, and *A* is better if X is true than it is if X is false, and there is some alternative act *B* that is available to the person such that *B* is better than *A* if X is false, **and the person refrains from gathering further evidence to determine the truth or falsity of *x* (or would refrain, if further evidence were available), even if the evidence is cost-free.**

Practical Rationality

Maximize expected utility

EU({O1, E1; O2, E2; … ; On, En}) =

$$\sum_{i=1}^{n} p(E_i)u(O_i)$$

A Challenge to Expected Utility Maximization

- Expected utility theory does a poor job of handling *risk aversion*.
- Most people are risk averse in that they prefer $x to a gamble that yields $x on average. E.g. most people prefer $50 to a coin flip between $0 and $100.
- Expected utility theory handles this by allowing that the utility function diminishes marginally. This locates the badness of risk in the outcomes: risk aversion is a property of how an agent evaluates particular outcomes.
- Intuitively, though, risk is a property of a gamble as a whole: a gamble is risky insofar as it has a high variance, a low minimum, and so forth. Risk is a "global" property of a gamble, rather than a "local" property. And intuitively, risk aversion has to do with how an agent takes global features of gambles into account: for example, if an agent is averse to risk, he cares more about the minimum than an agent who is risk neutral. (In the extreme case: an agent might care only about the minimum.)
- In other words: EU theory says "care about the average value." But there are lots of other features of a gamble to care about.

Risk Aversion: Some Examples

- **First Example: Gambles with Independent Goods**
- Consider two goods that are independent in the sense that having one does not increase or decrease the value of the other. E.g., non-independent goods include a right-hand glove and a left-hand glove, or a dollar and a second dollar. Independent goods might include a pair of gloves and a nice dinner.
- In utility terms, two goods are independent when u(A&~B) + u(B&~A) = u(A&B), assuming u(status quo) = 0.
- Consider the choice between the following deals. People tend to prefer deal 2 to deal 1, but there are no utility values we can assign to the goods such that they are independent and deal 2 is preferred to deal 1.

	HH	HT	TH	TT
Deal 1	Dinner	Dinner and gloves	Nothing	Gloves
Deal 2	Dinner	Dinner	Gloves	Gloves

Risk Aversion: Some Examples

- **Second Example: Allais Paradox**
- Consider the choice between A and B, and the choice between C and D, where the gambles are as follows:

A: $5,000,000 with probability 0.1, $0 otherwise.
B: $1,000,000 with probability 0.11, $0 otherwise.

C: $1,000,000 with probability 0.89, $5,000,000 with probability 0.1, $0 with probability 0.01.
D: $1,000,000 with probability 1.

- People tend to choose A over B, and D over C, but there are no utility values we can assign to $0, $1M, and $5M such that these choices maximize expected utility.

(Ask me for more examples if you remain unconvinced…)

Generalizing the Theory to Include a Subjective Risk Function

- Again, what these examples show is that people tend to care about global properties of gambles, e.g. the minimum (what happens in the worst state), the maximum, the variance, etc.

Standard theory:

{A, p; B, 1 – p} A gamble that yields A with probability *p* and B with probability 1 – *p*

Expected utility: u({A, p; B, 1 – p}) = p(u(A)) + (1 – p)u(B)

= u(B) + p[u(A) – u(B)]

- Decision makers have subjective utilities and probabilities.
- Decision makers are risk averse if the utility function is concave. (Or, more generally, if the goods involves are not independent.) This implies that risk aversion is a property of how an agent values individual outcomes.

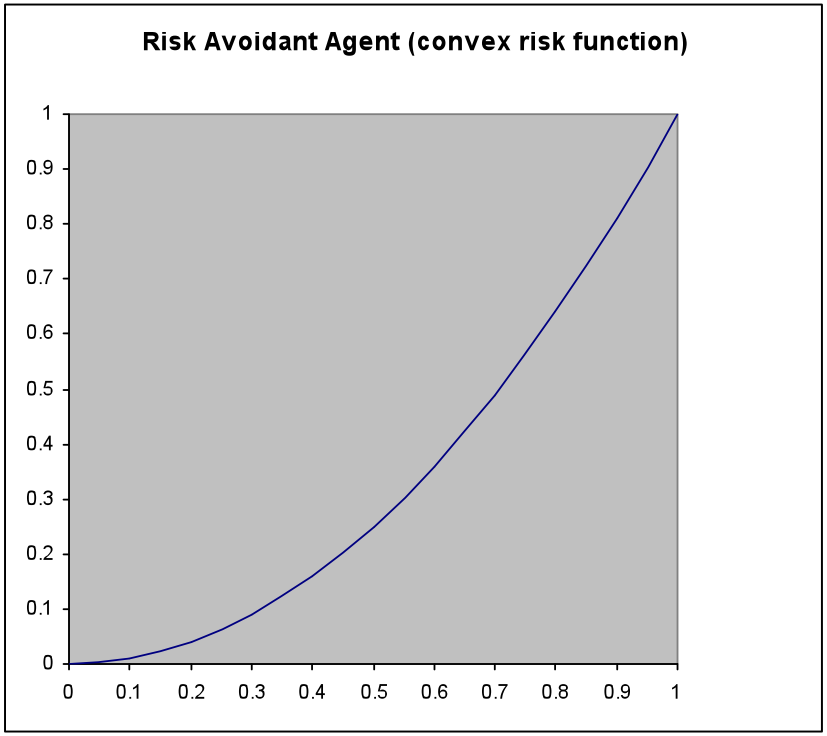
Generalization:

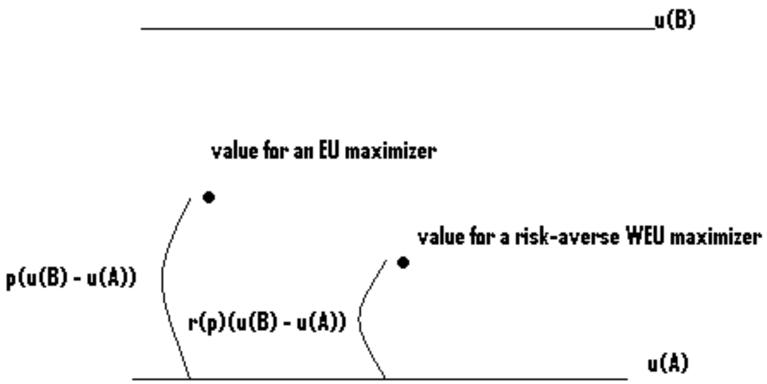
If A is at least as good as B:

Weighted-expected utility: u({A, p; B, 1 – p}) = u(B) + r(p)[u(A) – u(B)],

- r:[0, 1] → [0, 1] is a function that represents how the agent takes risk into account.
- Decision makers have subjective utilities, probabilities, and evaluation of risk.

Example of a Risk Function





Risk Function

- Decision makers are risk averse if the risk function is convex. (Adding probability to a state is better the more likely that state already is.) This implies that risk aversion is a property of how an agent values *arrangements of outcomes across the possibility space.*
- Decision makers are risk seeking if the risk function is concave. Decision makers are risk neutral if r(p) = p – they are ordinary expected utility maximizers.
- If r(p) is convex, the minimum is weighted more heavily than for the risk-neutral agent. Limit case: r(p) = 0 if p ≠ 1, the agent who uses maximin. If r(p) is concave, the maximum is weighted more heavily than for the risk-neutral agent. Limit case: r (p) = 1 if p ≠ 0, the agent who uses maximax.
- The introduction of a risk function allows for rational agents to have the preferences in the two original examples. It also allows for people to prefer $50 to a coin flip between $0 and $100, while still valuing these amounts of money linearly. And it explains why the St. Petersburg gamble is worth so little.

Gambles over multiple options, e.g. $\{O_1, p_1; O_2, p_2; \ldots ; O_n, p_n\}$

Expected utility:

$$u(O_1) + (\sum_{i=2}^{n} p_i)(u(O_2) - u(O_1)) + (\sum_{i=3}^{n} p_i)(u(O_3) - u(O_2)) + \ldots + (p_n)(u(O_n) - u(O_{n-1}))$$

Weighted-expected utility:

$$u(O_1) + r(\sum_{i=2}^{n} p_i)(u(O_2) - u(O_1)) + r(\sum_{i=3}^{n} p_i)(u(O_3) - u(O_2)) + \ldots + r(p_n)(u(O_n) - u(O_{n-1}))$$

Is refraining from gathering cost-free evidence irrational?

Good's Theorem

If an agent is choosing between three acts, A_1, A_2, and A_3, we can calculate the expected utility of each act:

$$\sum_i p(S_i)u(A_j \mid S_i)$$

The expected utility for a rational agent will be the maximum expected utility among the acts:

$$\max_j \sum_i p(S_i)u(A_j \mid S_i)).$$

For each possible experimental result E_k, the expected utility of getting that result and using it will be:

$$\max_j \sum_i p(S_i \mid E_k)u(A_j \mid S_i)).$$

Therefore, the expected utility of performing the experiment will be:

$$\sum_k p(E_k)\max_j \sum_i p(S_i \mid E_k)u(A_j \mid S_i).$$

Good proves that unless the act recommended is the same no matter what the result of the experiment is – that is, if the same act maximizes expected utility for each E_k – then this value is always higher than the agent's expected utility before making the observation.

Good’s Theorem: Example 1

- Consider a person with the following degrees of belief in X and R:

p(X) = 0.9	p(X | R) = 0.99	p(X | ~R) = 0.05	p(R) ≈ 0.904
p(~X) = 0.1	p(~X | R) = 0.01	p(~X | ~R) = 0.95	p(~R) ≈ 0.096

- The agent is deciding between two acts, A and ~A, associated with the following utility payoffs:

u(A&G) = 10	u(A&G) = 1
u(A&~G) = 0	u(~A&~G) = 1

- If he makes the decision now, he will do A, with EU 9.
- If he carries out the experiment, then his expected utility values would be:

If R comes out true: EU(A) = 9.9; EU(~A) = 1. He will do A, with EU 9.9.

If ~R comes out true: EU(A) = 0.5; EU(~A) = 1. He will do A, with EU 1.

- Since he will get result R with credence 0.904 and ~R with credence 0.096, the EU of performing the experiment is:

EU(E) = (0.904)(9.9) + (0.096)(1) ≈ 9.05

- So he should perform the experiment, as the theorem predicts.

Good's Theorem: Example 2

- Consider a person with the following degrees of belief in X and Q:

p(X) = 0.9	p(X | Q) = 0.99	p(X | ~Q) = 0.2	p(Q) ≈ 0.886
p(~X) = 0.1	p(~X | Q) = 0.01	p(~X | ~Q) = 0.8	p(~Q) ≈ 0.114

- This experiment differs in that the negative result tells less strongly against X.
- Again, if he makes the decision now, he will do A, with EU 9.

- If he carries out the experiment, then his expected utility values would be:

If Q comes out true: EU(A) = 9.9; EU(~A) = 1. He will do A, with EU 9.9.

If ~Q comes out true: EU(A) = 2; EU(~A) = 1. He will do A, with EU 2.

- So no matter what he sees, he will do A.
- Since he will get result Q with credence 0.886 and ~R with credence 0.114, the EU of performing the experiment is:

EU(E) = (0.886)(9.9) + (0.114)(2) = 9.

So he should be indifferent between performing the experiment and not.

Risk Avoidant Agent

- The conclusion of Good's Theorem doesn't hold for non-expected utility maximizers.
- Example: risk avoidant agent with $r(p) = p^2$
- Example 1:

p(X) = 0.9	p(X | R) = 0.99	p(X | ~R) = 0.05	p(R) ≈ 0.904
p(~X) = 0.1	p(~X | R) = 0.01	p(~X | ~R) = 0.95	p(~R) ≈ 0.096
u(A&G) = 10	u(A&G) = 1		
u(A&~G) = 0	u(~A&~G) = 1		

- If he makes the decision now:

REU(A) = 0 + $(0.9)^2$(10 – 0) = 8.1

REU(~A) = 1 + $(0.9)^2$(1 – 1) = 1

- If he performs the experiment and gets result R:

REU(A) = 9.8; REU(~A) = 1; do A.

- If he performs the experiment and gets result ~R:

REU(A) = 0.025; REU(~A) = 1; do ~A

- So performing the experiment **is equivalent to taking a gamble that yields 10 utiles if R&X, 0 utiles if R&~X, and 1 utile if ~R:** 0 with credence 0.009, 1 with credence 0.096, 10 with credence 0.895.

REU(E) = 0 + r(p(~R)+p(R&X))(1 – 0) + r(p(R&X))(10 – 1)

= $(0.096 + 0.895)^2(1) + (0.895)^2(9) = 8.2$

- So he should perform the experiment.
- The experiment makes the agent really unlikely to get a payoff of 0, since p(R&~X) is very low, but still very likely to get a payoff of 10, since p(R&X) is high.

Risk Avoidant Agent, Example 2

- Example 2:

p(X) = 0.9 | p(X | Q) = 0.99 | p(X | ~Q) = 0.2 | p(Q) ≈ 0.886

p(~X) = 0.1 | p(~X | Q) = 0.01 | p(~X | ~Q) = 0.8 | p(~Q) ≈ 0.114

u(A&G) = 10 | u(A&G) = 1

u(A&~G) = 0 | u(~A&~G) = 1

- If he makes the decision now:

REU(A) = 0 + $(0.9)^2$(10 – 0) = 8.1

REU(~A) = 1 + $(0.9)^2$(1 – 1) = 1

- If he performs the experiment and gets result Q:

REU(A) = 9.8; REU(~A) = 1; do A.

- If he performs the experiment and gets result ~Q:

REU(A) = 0.4; REU(~A) = 1; do ~A

- Again, performing the experiment is equivalent to taking a gamble that yields 10 utiles if Q&X, 0 utiles if Q&~X, and 1 utile if ~Q: 0 with credence 0.009, 1 with credence 0.114, 10 with credence 0.877.

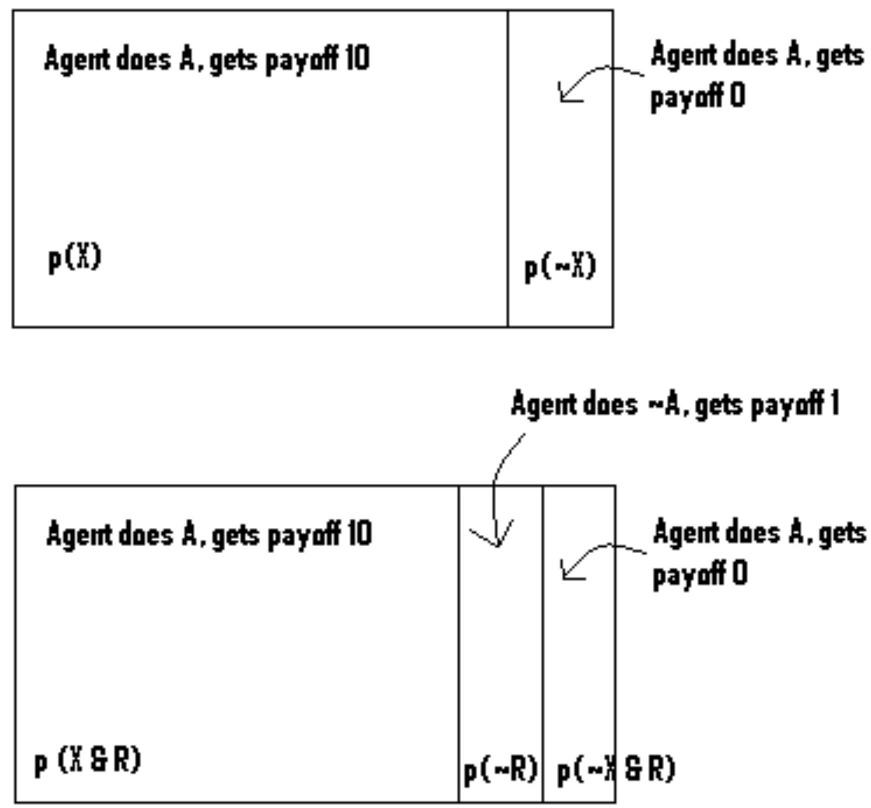
REU(E) = 0 + r(p(~Q)+p(Q&X))(1 – 0) + r(p(Q&X))(10 – 1)

= $(0.114 + 0.877)^2$(1) + $(0.877)^2$(9) = 7.9

- So he should not perform the experiment.
- Again, the experiment makes the agent really unlikely to get a payoff of 0, since p(R&~X) is very low, but it also decreases the agent's likelihood of getting a payoff of 10, since p(R&X) is a lot lower than p(X).

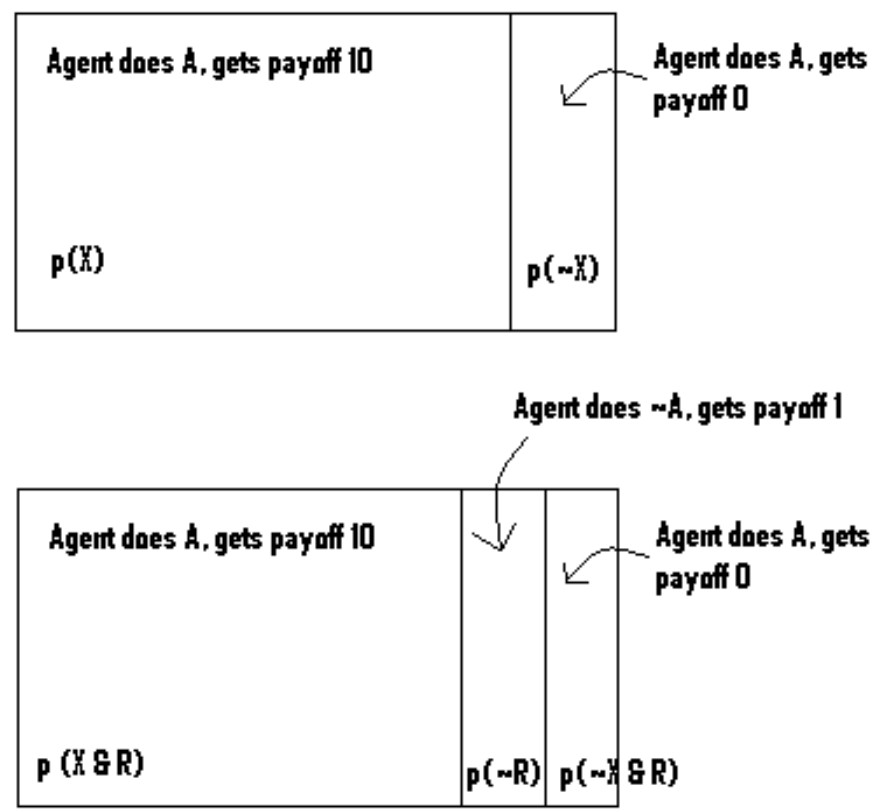
How experiments shift possible payoffs

- Some credence is shifted from the region in which the agent gets 0 to the region in which the agent gets 1: namely, p(~R & ~X), the credence associated with a “correct” negative result.
- Some credence is shifted from the region in which the agent gets 10 to the region in which the agent gets 1: namely, p(~R & X), the credence associated with an “incorrect” negative result.
- For the EU maximizer, these shifts never result in an all-things-considered worse gamble, and they sometimes result in an all-things-considered better gamble.
- For the EU maximizer, the possibility of improving from 0 to 1 (with a particular probability) always makes up for the possibility of worsening from 10 to 1 (with a particular probability).



How experiments shift possible payoffs

- However, for risk avoidant agents, these credence shifts won't always result in an all-things-considered better gamble.
- In order to result in an ATC better gamble, improvements need to be greater to offset declines.
- Part of what it means to be risk avoidant is to not be willing to accept the possibility of a loss in exchange for an equivalently-sized possibility of a gain.
- Rough characterization of situations in which performing an experiment lowers the REU:
  - Agent already has fairly high credence in X.
    - => positive result won't help very much.
  - The negative result will not be conclusive evidence in favor of X
    - => A significant amount of credence will shift away from the p(X) region.



What's going on in non-formal terms

- If someone believes in a proposition X with near certainty, then evidence in favor of that proposition won't drastically alter her degree of belief in that proposition, so it won't drastically alter her REU.
- On the other hand, evidence for ~X will cause a drastic change.
- In the situation above, A is a risky option: it does very well if X is true and very poorly if ~X is true. So the risk avoidant agent needs a fairly high credence in X in order to perform it.
- Therefore, evidence for ~X will put the risk avoidant agent in a situation in which it is rational to perform ~A. But from her (new) point of view, there's still a fairly high chance (0.5) that this is misleading evidence: *evidence that leads her to do the action that, though rational, in the end gives her a lower payoff*.

What's going on in non-formal terms

- "Misleading" evidence: evidence that makes it prudent to do ~A even though, as it turns out, X is true; evidence that leads the agent to miss out of the possibility of something great, namely doing A when X is true.
- p(~R & X): probability of getting misleading evidence for ~X.
- To emphasize: this needn't involve any irrationality on the agent's part; nor does it involve not wanting to know the truth.
- For the risk avoidant agent, the risk of misleading evidence is not made up for by the benefit of non-misleading evidence.
- Example: If a private investigator turned up evidence that your spouse was cheating, it would indeed be rational to leave her, but, given your previous credence in her faithfulness, *there is still a significant chance that by doing the rational thing you would be missing out on some great good,* namely a relationship with a spouse who is *in fact faithful*.

Faith Redux

- Faith (analyzed as requiring not looking for further evidence) and risk avoidance rise or fall together: if risk avoidance is rationally permissible, then so is faith.
- Modus tolens?
- We can now separate cases in which faith is rational from those in which faith is irrational ("misplaced"):
  - A is a risky act, relative to the available alternatives.
  - There is already a fairly high chance of X.
  - The negative result of the experiment one is considering performing is not conclusive enough in favor of ~X.
- Caveat: whether faith is rational depends on the experiment to be performed. But our ordinary notion of faith is not relative to a particular experiment.
  - Furthermore, the "conclusive experiment" will always be rational to perform.
- Resolving the dilemma?
  - Make faith relative to a particular search one is considering?
  - Limit faith to actual acts? Then faith will be rational in cases in which there are no experiments that would be conclusive enough.