Externalist theories

Who cares about epistemic duties? (Not these guys.)



Al Goldman



Robert Nozick



Al Plantinga

BonJour on Externalism

"But in spite of this historical consensus, many recent epistemologists have argued that the internalist conception of justification is fundamentally mistaken, that epistemic justification can depend in part or perhaps even entirely on matters to which the believer in question need have no cognitive access at all, matters that are entirely *external* to his or her cognitive viewpoint."

Laurence BonJour, Epistemology, p. 203

"My conviction is that views of this kind are merely wrong-headed and ultimately uninteresting evasions of the central epistemological issues."

(BonJour, In Defense of Pure Reason, p. 1, n. 1.)

Basic challenge for externalism

"Indeed, if features of a belief that are in this way external to the believer's cognitive perspective can yield justification, **why could truth itself not play this role?** Surely the fact that a belief is true is, in a way, the best possible reason for holding it ...

In fact, no externalist is willing to go quite this far, but in a way that merely heightens the puzzling character of the externalist view: why should some external facts and not others be relevant to justification?"

BonJour, Epistemology, p. 204, emphasis added.

Warrant or Justification?

- In Martin chapter 4, externalists are presented as analysing "justification" in external terms.
 - E.g. "Thus Ayer denies access internalism; he is an externalist about justification."
- In contrast, Plantinga sees "justification" as too strongly tied to internalist notions of evidence, and epistemic responsibility, to analyse it in external terms:

"The term 'justification' suggests duty, obligation, permission, and rights—the whole deontological stable"

• (Plantinga, "Positive Epistemic Status and Proper Function", 1988)

Warrant or Justification?

- Some externalists think that justification (in the sense of epistemic responsibility) is an important concept in epistemology, even though it isn't always necessary for knowledge.
 - I.e. externalists **still need a word** for that concept.
- So, I agree with Plantinga that externalists should not give an external analysis of "justified", but should use a new term.
 - "Warrant" is probably the best choice here.

Externalist theories

- Causal theory
- Reliabilism
 - Counterfactual theories (e.g. truth tracking)
 - Process reliabilism
- Engineering/design standpoint (inspired by Reid)
 - Proper functionalism
 - Authoritarianism ("From the mint of nature")

Overview: Causal theory (Goldman)

• S knows that *p* iff the fact *p* is causally connected in an appropriate way with S's believing *p*.

(This is only the simplest case of an appropriate causal connection.)



Overview: Reliabilism

- Reliabilism says that a belief is justified (or "warranted"?) when it is *reliable* in some sense.
 - Counterfactual reliabilist theories (e.g. Nozick, Dretske) understand reliability in terms of facts like, "if P weren't true, then the subject would not believe that P" (owing to the laws of physics, and the construction of the cognitive process)
 - **Process reliabilism** (e.g. Goldman, Armstrong) says that a belief is justified by virtue of the reliability of the *process* that produced the belief.

The "engineering standpoint" (Plantinga, Reid)

- Whatever caused the formation of living organisms had a strong "bias" towards making systems that accomplish certain *functions* (e.g. selfreproduction, getting energy from the environment, sensation, and motion).
- Thus organisms are (on all accounts) "designed" in some sense to accomplish certain functions.
- "Design" is (almost) an *observation* here, not a theory.

E.g. Issus coleoptratus



The "engineering standpoint"

- Belief formation is also a biological process, and is therefore designed to work in a certain way, to accomplish specific functions.
- For a belief to have warrant, the cognitive process that produces it must be **working properly**, according to its design plan.

The causal theory

"A Causal Theory of Knowing", The Journal of Philosophy, 1967.

Simple form: S knows p iff S's belief in p is caused by the fact p.

N.B. Goldman considers chains of inferences, in a person's mind, to be causal chains. (Correctly, I think.)

- E.g. Smith returns home and finds a lot of sawdust and wood chips where a tree used to be. Smith believes that the tree was cut down.
 - Does he *know* that the tree was cut down?
- Scientific theories about events in the distant past (e.g. evolution) are inferred from data observable today.

Clarifications

- There is no need for a knower of p to be consciously **aware** of their grounds for believing p.
- 2. The analysis only provides truth conditions for knowledge, and is not a conceptual analysis.
- *3. Truth* conditions are distinct from *verification* conditions.
 - My analysis of "S knows p" does not purport to give procedures for finding out whether a person (including oneself) knows a given proposition

A common cause is enough

"Zeke tells Alice on Friday morning that he'll arrive in town on Saturday. He's very responsible and reliable, and on Friday Alice believes what he says. He arrives, as he said he would, on Saturday."



Knowledge of the future?

• Note that, in this case, Alice has knowledge of a *future* event.

• Goldman:

"If we ever can be said to have knowledge of the future, this is a reasonable candidate for it."

Goldman's final version

• S knows that p if and only if the fact p is causally connected in an "appropriate" way with S's believing p.

"Appropriate," knowledge-producing causal processes include the following:

(1) perception

(2) memory

(3) a causal chain, exemplifying either Pattern 1 or 2, which is correctly reconstructed by inferences, each of which is warranted (background propositions help warrant an inference only if they are true)

(4) combinations of (1), (2), and (3) (pp. 369-70)

"Correctly reconstructed by inferences"?

 $(p) \longrightarrow B_T(p) \longrightarrow A_T(p) \longrightarrow B_S(A_T(p)) \longrightarrow B_S(B_T(p)) \longrightarrow B_S(B_T(p)) \longrightarrow B_S(p)$ T believes that p S believes that T asserts that p the fact p

S believes that T believes that p

S believes that p

T asserts that p

Figure 2

• In this case ... S knows p because he has correctly reconstructed the causal chain leading from p to the evidence for p that S perceives, in this case, T's asserting (*p*). This correct reconstruction is shown in the diagram by S's inference "mirroring" the rest of the causal chain.

Problems with the causal theory

1. Knowing math

– Can facts (especially abstract facts) be causes?

- 2. The case of the mistaken gossip — The is a causal connection, but no knowledge here.
- 3. The fake barns case

Knowing mathematical facts

"Usually philosophers say that only *events* are causes – things that happen at a particular time. Now, surely we know truths of arithmetic, but the facts that make them true aren't events – they're eternal – so they can't be causes."

Two issues:

- 1. Can eternal facts be causes?
- 2. Can abstract facts be causes?

- 1. Can *eternal* facts be causes?
 - Can a nearby lake (that has existed for a million years) one day cause you to believe that it exists?
 - Can the value of a physical constant (such as G) one day cause you to believe it has the value 6.7×10^{-11} ?
- 1. Can *abstract* facts be causes?
 - This is a general problem in philosophy of math, as being non-causal is usually taken to be an essential feature of abstract objects.

More details on the Jack/Joan case

"Suppose that the reason the office gossip thought that Jack was going to get the job was that he saw a note on the big boss's desk saying 'TELL HEAD OFFICE THAT J. HAS THE JOB'. The reason the note was there was that *Jim* had been given the job; the office gossip didn't even know that Jim was in the running, and just assumed it was Jack. That's why he told Joan that Jack got the job."

Draw a causal diagram to represent this..

Causal diagram



 Actually, as some students pointed out (especially Patrick), Joan doesn't correctly reconstruct the causal chain, so this fails as a counter-example to the causal theory.

Fake Barns Case (Carl Ginet)

- "Henry is driving in the countryside with his young son, identifying landmarks for him as they drive past. 'Look, son— a cow! Over there—a tractor! There's a barn over there, in that field!'
- Henry's belief that there's a barn in the field is caused by his perceptual experience of the actual barn.
- He doesn't know this, but Henry is driving through Fake Barn County, where the zany locals have put up dozens of barn façades, false fronts that just look like barns when seen from the highway. It's sheer luck that Henry is right now looking at the one real barn in the region"

N.B. Not all philosophers have a clear intuition that Henry doesn't know "There's a barn over there, in that field".



- E.g. Ruth Millikan, "Naturalist Reflections on Knowledge", *Pacific Philosophical Quarterly*, Vol. 65 (4), 1984.
 - "Two of three teenage youngsters that I recently questioned do take them to be examples of knowledge ... I think they are defensible"

Verdict on the causal theory?

Reliabilism

- Reliabilism says that a belief is justified (or "warranted"?) when it is *reliable* in some sense.
 - Counterfactual reliabilist theories (e.g. Nozick, Dretske) say things like, "if P weren't true, then the subject would not believe that P" (owing to the laws of physics, and the construction of the cognitive process)
 - **Process reliabilism** (e.g. Goldman) says that a belief is justified by virtue of the reliability of the process that produced the belief.

David Armstrong: thermometer model

"... there must be a *law-like connection* between the state of affairs *Bap* [i.e., a's believing that p] and the state of affairs which makes 'p' true, such that, given *Bap*, it must be the case that p."

(Belief, Truth and Knowledge (1973), p. 166)

BonJour comments: "This is what Armstrong calls the "thermometer-model" of non-inferential knowledge: just as the readings of a reliable thermometer lawfully reflect the temperature, so one's basic beliefs lawfully reflect the states of affairs that make them true. A person whose beliefs satisfy this condition is in effect a reliable cognitive instrument; and it is, according to Armstrong, precisely in virtue of this reliability that these basic beliefs are justified."

Nozick and 'truth tracking'

- **TT**. S knows *p* iff
 - (i) p is true,
 - (ii) S believes p,
 - (iii) S's attitude toward *p* tracks the truth value of *p*:
 - —If p were **not** true, S would **not** have believed p; and
 - if matters had been different in a way that p remained true, S would still have believed p.
- (How does this handle the Gettier cases?)
 - Nogot/Havit
 - Sheep in the field
 - Fake barns

(iii) S's attitude toward *p* tracks the truth value of *p*:

—"Sensitivity"	$\neg p \Box \rightarrow \neg B(p)$
 "Stability" 	$p \Box \rightarrow B(p)$

- $\neg p \Box \rightarrow \neg B(p)$ means that in the closest world in which p is false S does not believe p.
- $p \square \rightarrow B(p)$ means that in the closest worlds where *p* remains true, B also believes *p*.

- Some add a "Safety" condition: $B(p) \Box \rightarrow p$.
 - If *S* were to believe *p*, *p* would be true

- Why is the *second* counterfactual (Stability) needed?
 - (if matters had been different in a way that p remained true, S would still have believed p.)

Nozick: "If someone floating in a tank oblivious to everything around him is given (by direct electrical and chemical stimulation of the brain) the belief that he is floating in a tank with his brain being stimulated, then ... he does not know that it is true."

Does this belief (that he's in a tank) track the truth?

- $\neg p \Box \rightarrow \neg B(p)$ holds
- $p \square \rightarrow B(p)$ fails

Lucky knowledge?

- Black is hard at work in her office. From time to time she looks up from her desk and computer to stretch her neck. On one such occasion she happens to glance out the window toward the street. Just at that moment she sees a mugging on the street. She has a clear view of the event. She is a witness.
- Black knows that a mugging has just occurred. Does this belief *track the truth*, however?
- N.B. "when p is true, S does believe p" is usually taken to mean that S believes p in all the "close" possible worlds where p is true.

Toward process reliabilism?

TT*. S knows p iff

- *i. p* is true,
- ii. S believes p,
- iii. S used *method M* to form the belief in *p*, and
- iv. when S uses **method M** to form beliefs about p, S's beliefs about p track the truth of p.
- This modification leads to the *generality problem* of deciding which "method" or "process" was active in producing the particular belief that *p*,
- Does this even fix the problem of lucky knowledge?

Variant of the fake barns case

 Smith sees what look like barns of various colours, then points to a (real) red one and forms two beliefs:

That's a barn, and *That's a red barn*

Do these beliefs track the truth?

- The first belief doesn't track the truth, but the second one does! (The fake barns cannot be coloured red.)
 - So only the second belief is knowledge?

Scepticism and externalism

- According to reliabilism, and externalism generally, having knowledge doesn't require you to be able to rule out sceptical scenarios, using the resources you have cognitive access to.
- Having knowledge may require that you "couldn't be wrong", but in the *objective* sense of possible, not the subjective sense.
- E.g. Nozick's externalism has interesting consequences:
 - I know there is one hand here
 - I don't know that I'm not a brain in a vat

Verdict on the tracking theory?

Process Reliabilism

- The justificational status of a belief is a function of the reliability of the process or processes that cause it, where (as a first approximation) reliability consists in the tendency of a process to produce beliefs that are true rather than false.
- N.B. If this "tendency" to produce true beliefs is understood as a probability, then it must be an *objective* probability, not a subjective one.

What is a 'reliable' process?

- "One that always (or almost always) works, across a wide range of circumstances."
- Thus there are two dimensions to reliability:
 - i. A high *probability* or frequency of giving correct results (e.g. "works in 99% of cases")
 - ii. Robustness: the probability in part (i) applies to a *wide range* of circumstances.

Range

- A reliable process is robust, i.e. it is almost always right in a wide (though limited) *range* of circumstances. E.g. a voltmeter:
 - has a maximum voltage before it gets fried.
 - works only in a certain temperature range,
 - gets affected by ionizing radiation.
 - withstands ordinary knocks, but won't work after getting walloped by a sledge hammer.

Is "lucky" knowledge reliable?

- Consider the earlier case, where Black (as a result of luckily glancing up at just the right time) knows that a mugging has occurred.
- Was this belief produced by a reliable cognitive process?
 - Yes

Degrees of reliability

- Reliability is obviously a matter of degree. No process is perfectly reliable, and some reliable processes are more reliable than others.
- Is this a problem for reliabilism?
- Reliabilists might say this is fine, since justification is a matter of degree as well.
 - But if knowledge is all-or-nothing, defining K = RTB is going to lead to more Gettier problems.

Goldman's formulation

- **1. Basic beliefs** (non-inferential beliefs) are justified to the extent that they result from reliable cognitive processes (e.g. perception, memory).
- 2. Non-basic beliefs are justified to the extent that they are produced by a reliable inferential process, using justified beliefs as premises.

Objections to reliabilism

- 1. Reliability is not *necessary* for justification
 - The **Brain in a Vat** (or Cartesian evil genius) case. These people are epistemically blameless, being diligent, careful and thorough. Hence their beliefs are justified. But they're not at all reliable.
- 2. Reliability is not *sufficient* for justification
 - **Clairvoyance cases**. These genuine clairvoyants have reliable beliefs, but the beliefs are unjustified because they have no reason to think they're true.
- 3. Generality Problem
 - Talk of "the process" that produced belief in p is fatally ambiguous. Reliability is defined only for process types, and each token process is an instance of many different types.

1. Evil genius case

"Imagine a group of people who live in a world controlled by a Cartesian evil genius ...

...The people in this position are, we may suppose, careful and thorough investigators. They accumulate large quantities of sensory evidence, formulate hypotheses and theories, subject their beliefs to careful experimental and observational tests, and so on ...

... Are the beliefs about their apparent world that the people in such a Cartesian demon world arrive at in these ways justified? ... From an intuitive standpoint, it seems hard (doesn't it?) to deny that they are."

(BonJour, *Epistemology*, p. 228)

2. Clairvoyance cases

... doesn't it seem as though Norman is being thoroughly irrational and so is not in fact justified in confidently accepting beliefs on this sort of basis?

(Think about this question on your own. One way to develop the issue further is to ask whether Norman **would be justified in acting** on one of these beliefs if an urgent occasion should arise: perhaps someone is trying to contact the president on an urgent matter and asks Norman if he knows where to find him.

(BonJour, p. 231)

Variants of the clairvoyance case

- In the case mentioned, Norman has independent evidence that his clairvoyant belief, that the President is in NYC, is false. (TV news, etc.)
- Bonjour also considers cases where Norman has scientific evidence that clairvoyance is impossible, all alleged clairvoyants have been exposed as fakes, etc.
- Such evidence makes no difference at all to Norman's reliability, yet it certainly makes his belief that the President is in NYC less justified!

3. Generality problem

"The **Generality Problem** is the problem of specifying exactly which process it is whose reliability determines how justified your belief is." (Jim Prior)

E.g. I look out the window, and believe "it's raining". Which process formed this belief? Is it:

- the process of forming beliefs on the basis of *perception*
- the process of forming beliefs on the basis of *vision*
- the process of forming beliefs about the weather on the basis of looking out a window
- the process of forming a belief that it's raining on the basis of seeing droplets splashing on the pavement
- etc.

Types or tokens?

- The generality problem takes it for granted that reliability values are defined only for process *types* rather than *tokens*.
- (After all, each token process produces either a true belief or a false one. There is no success rate!)
- Is this correct?
- (Most people now think that there are single-case chances.
- Can a reliable process be defined as one with a high chance of success, in that *exact* set of circumstances?)

Generality problem

- 1. If we define the process broadly, e.g. "vision", then the problem is that *some visual beliefs are a lot more justified than others*. Yet on this view they'd all be equally justified.
- 2. To avoid (1) we define the process as narrowly as possible. But then there may be only one belief that the process ever produces (no two perceptions are identical) and so it's meaningless to ask how often it produces true beliefs.
 - Appeal to single-case *chances*, rather than frequencies?
 - But since a reliable cognitive mechanism is also robust, in the sense that it works over a wide range of circumstances, the process must be defined broadly?

Generality Problem

Without some way of answering this question in a specific and nonarbitrary way, the reliabilist has not succeeded in offering a definite position at all, but only a general schema that there is apparently no nonarbitrary way to make more definite. Certainly some ways of specifying the relevant process are more natural than others; but the epistemological relevance of such naturalness is questionable, and even these more natural specifications are numerous enough to result in significantly differing degrees of reliability.

Though reliabilists have struggled with this problem, no solution has yet been found that even a majority of reliabilists find acceptable. (BonJour, p. 215)

4. The Range Problem

(The reliability of a machine is always limited to a certain range of circumstances, or possible environments.)

"The **Range Problem** is the problem of specifying *where* a process has to be reliable--in what range of possible environments?--in order for beliefs produced by it to count as justified."

• I.e. Knowledge seems to require some range of reliability, surrounding the actual circumstances. *What defines this range?*

5. Reliability is a matter of degree

• Linda Zagzebski (1994), "The Inescapability of Gettier Problems", *Philosophical Quarterly*.

"As long as the truth is never assured by the conditions that make the state justified, there will be situations in which a false belief is justified. I have argued that with this common ... assumption, **Gettier cases will never go away**"

- To avoid Gettier you'd need a theory that analysed knowledge as K = WB rather than K = WTB.
 - I.e. you need $W \Rightarrow T$.
 - This means that 'warrant' is *all or nothing*.

Verdict on process reliabilism?

- Does process reliabilism avoid the Gettier counterexamples? In the sheep case, for example, is there any unreliable process involved?
- If not, then that's a great loss. Externalism was supposed to be the radical solution to the Gettier problem. Other externalist theories (causal, truthtracking) have avoided at least the original Gettier cases.

The "engineering standpoint"

- Belief formation is a biological process, and is therefore designed to work in a certain way, to accomplish specific functions.
- For a belief to have high epistemic status (warrant?) the cognitive process that produces it must be working properly, according to its design plan.
- First proposed by Alvin Plantinga in *Warrant and Proper Function*, 1993.
- One of Plantinga's claims is that the "gap filler", that turns true belief into knowledge, is *warrant* rather than justification.
- K = WTB.

Plantinga's "proper functionalism"

Plantinga says that a belief, B, is warranted if:

- (1) the cognitive faculties involved in the production of B are functioning properly...;
- (2) your **cognitive environment** is sufficiently similar to the one for which your cognitive faculties are designed;
- (3) ... the design plan governing the production of the belief in question involves, as purpose or function, the production of true beliefs...; and
- (4) the design plan is a good one: that is, there is a **high** statistical or objective probability that a belief produced in accordance with the relevant segment of the design plan in that sort of environment is true.

Simplified ...

Plantinga says that a belief B is warranted iff it's produced by a cognitive mechanism M, where:

- (1) M is functioning properly
- (2) M is in the cognitive environment it's designed for
- (3) M is aimed at truth
- (4) M is reliable

Reliability isn't sufficient for warrant

Plantinga's "Epistemically Serendipitous Brain Lesion"

"Suppose that Sam suffers from a serious abnormality – a brain lesion, let's say. This lesion wreaks havoc with Sam's noetic structure, causing him to believe a variety of propositions, most of which are wildly false. It also causes him to believe, however, that he is suffering from a brain lesion. Further, Sam has no evidence at all that he is abnormal in this way, thinking of his unusual beliefs as resulting from an engagingly original turn of mind."

Accidentally reliable processes?

- Plantinga stipulates that the brain lesion is a *reliable* producer of the belief that one has a brain lesion, yet it seems that the belief isn't knowledge because the brain isn't *designed* to produce this belief.
- Does that sound like the right response?
- In that case, having a reliable true belief isn't *sufficient* for knowledge.
- How does this compare to BonJour's clairvoyance case?

Clairvoyant by design?

- What if Norman isn't merely a reliable clairvoyant, but is *designed* to be one? (The other perceptual mechanisms are designed, after all.)
- What if Norman's clairvoyant beliefs are "clear and distinct", assuring him of their reality? (This is of course a design feature as well.)
- We don't think ordinary perceptual beliefs need independent verification. (They're basic.) Why then would clairvoyant beliefs need this?

Reid's causal theory

- Thomas Reid didn't explicitly develop an externalist theory of knowledge.
 - But his response to scepticism allows us to guess what he might have said. (Feel free to make your own version.)
- Roughly speaking, *Knowledge is authorized belief*, i.e. a belief counts as knowledge if it is authorized by the "wise author of nature" (God).
- But what makes a human belief authorised?

Authorised mechanism

- A mechanism only carries the authority of its designer when it is *working as designed*.
 - This is similar to the way a written text carries the authority of the writer, but only so long as the text has not been altered or corrupted.
- So I think Reid would accept Plantinga's "proper functioning" condition for knowledge.

Authorised source

- Any causal *input* for the belief-producing mechanisms must also be authorised.
- Objective facts, according to Augustinian theism, are part of the divine viewpoint, and hence authorised.
- Thus, a perceptual belief must be caused by the corresponding fact (as the early Goldman said).
- Fact that $p \rightarrow$ Cognitive apparatus \rightarrow Belief that p

BonJour's question for externalists

- **Qu**: "why should some external facts and not others be relevant to justification?"
 - E.g. why not make the *truth* of *p* sufficient for *p* to be justified?
- **Ans**: If knowledge is *authorized* belief, then knowledge arises just when authorization is transferred down the causal chain to the agent's belief. External states of affairs are relevant to knowledge only when they're relevant to that acquisition of authority.

Questions about Reid's causal theory

- Does Reid's theory give correct answers in the Gettier cases?
 - JFB + JD \Rightarrow JTB cases?
 - Ginet's fake barns?
- What does Reid's theory say about the lottery problem?
- Does Reid's theory allow beliefs about the future?
- Belief in scientific theories?

Designed environment?

- Plantinga's theory was also inspired by Reid, and is similar to my version of Reid.
- Plantinga's theory doesn't define knowledge as authorised belief, or have a requirement that the belief that *p* be caused by the fact *p*.

But Plantinga **adds** a requirement:

(2) your cognitive environment is sufficiently similar to the one for which your cognitive faculties are designed

- Should this be part of Reid's theory?
 - Maybe to give a different answer in the fake barns case?

Epistemic duties for Reid?

- One cognitive mechanism that humans have is *reason*.
 - How is it designed to operate?
- Can't he say that one of its functions is to integrate multiple sources of information, especially when they conflict?
- In some cases at least, one might have a duty to doubt one's spontaneous beliefs, and verify them using other factors that are internally available (e.g. other beliefs)?
- If so, a Reidian/proper-functionalist can make some concessions to internalism.



- Unlike animals, humans can be aware of their own belief formation, and think critically about it.
- Arguably, this ability entails epistemic duties in some cases.

E.g. Stanislav Petrov saves the world

- In 1983 Petrov was the duty officer for a Soviet nuclear early-warning system when the system reported that 3-5 ICBMs had been launched from the United States.
- According to protocol, Petrov should have reported this attack to superiors. Had he done so, the USSR would probably have retaliated with its own nuclear weapons. ⁽³⁾
- Instead, Petrov judged that it was (probably) a false alarm, and did not report it. Why?
 - For the USA to launch so few missiles would be suicide
 - The automatic detection system was new and untested
 - Ground radar found no corroborative evidence

Stanislav Petrov case

- Note that, in this case, Petrov had multiple sources of information:
 - The automatic missile detection system
 - His background knowledge of nuclear conflict
 - Ground radar signals
- Given that they conflicted, he was (arguably) right to be sceptical of the missile warning.

Multiple cognitive mechanisms

- Like Petrov, every person has multiple cognitive faculties (vision, hearing, touch, reasoning, memory, etc.) which overlap in the sense that they can agree with or contradict each other.
 - E.g. you see (or seem to see) a pink elephant, floating in the air.
 - E.g. Norman has a clairvoyant experience, accompanied by assurance, yet science disavows clairvoyance.
- (What would a cognitive engineer design the agent to do in such cases?)

- In some cases at least, prior beliefs overrule perception.
- In other cases, perception overrules previously strong convictions.
 - (Your friend had a leg amputated. You saw the stump. But later you see him again with two legs—he says it grew back, and the doctors are baffled. Will you *ever* believe this happened?)
- In such clashes, what determines the winner?
 - Presumably reason is meant to decide such things.