

**What good are counterexamples?
Brian Weatherson**

Intuition is sometimes mistaken; we should be better at telling when that is the case.

1. Intuitions

Intuitions are intellectual seemings, not beliefs. Intuitions cannot be calibrated, not even different kinds against each other as in perception. They are often systematically mistaken, manifesting themselves in wrong beliefs: logical and probabilistic mistakes, moral mistakes (e.g. 'slavery is acceptable'), conceptual mistakes ('scratching one's head is neither voluntary nor involuntary', 'whales are fish').

2. Correcting Mistakes

A philosophical theory should be judged by four criteria (tests):

- 1 – It should have few counterexamples, and/or only ones based on weak intuitions, and/or ones whose origin is explained by the theory.
- 2 – It should have few unacceptable theoretical consequences, e.g. a normative theory should eliminate arbitrariness.
- 3 – Analysandum and analysans should be theoretically significant.
- 4 – Simplicity.

JTB theory of knowledge may only do badly on 1, but do better than its rivals on the remaining three tests.

3. Naturalness in a Theory of Meaning

If that was so, and the JTB theory was to be preferred as a result, we should take our Gettier intuitions to stand corrected.

Objection (a): Instead, we should take our acceptance of the theory to amount to a change in the meaning of 'knows'.

The objection relies on a faulty theory of meaning that would imply infallibilism about knowledge ascriptions (whenever we decide to change our usage of a term, we thereby change its meaning, so we can never stand corrected).

Consider a theory of meaning based on Lewis's notion of naturalness. The meaning of a predicate is a property, i.e. a set of possibilia. In assigning such a set to a predicate, we follow two criteria, corresponding to tests 1&2 and tests 3&4, respectively: first, an assignment should validate as many as possible of our pre-theoretic beliefs; second, it should assign a property that is simple and significant, where this is manifested in a high degree of naturalness (taken to be primitive). These desiderata should be weighed, and the best on-balance candidate should be adopted.

The JTB theory does much better than its rivals with regard to the second of these, and not much worse with regard to the first.

This balance theory of meaning can account for the possibility of mistaken intuitions, while still denying the possibility that intuitions about meaning can be systematically and radically mistaken.

It makes the right kind of predictions about when meaning will diverge from intuitions about meaning (cf sections 1, 2).

It can account for the fact that some, but not all disagreements are disputes about matters of fact (or values) as opposed to disputes about matters of meaning. A shared language is a collection of closely-enough overlapping idiolects. If someone's usage differs only slightly from others', the assignment which does best at capturing that person's usage according to both criteria is the same as the assignment which does best at capturing other people's usage. The reason is that highly natural properties are rare.

This is why acceptance of the JTB theory would not amount to a change in the meaning of 'knows'. It would change usage considerably, but since reasonably natural properties are rare and unevenly distributed across the space of possible linguistic dispositions, that change does not amount to a change in meaning. There is no reasonably natural property in the vicinity of our disposition to use 'knows'. (most natural in vicinity given by JTB?)

Objection (b): conceptual intuitions should be given greater privilege than logical, ethical or semantic intuitions. It is wrong to generalize from systematic fallibility of the latter to systematic fallibility of the former.

If this is right, should consider a group that systematically assents to 'If A then B implies if B then A' to be expressing a mistaken logical but a correct conceptual intuition (?). This has various odd consequences: implies that we cannot be mistaken about our logical intuitions; implies that we cannot say that they are making a mistake; implies that it is vitally important in debates about philosophical logic whether one is engaging in logical or conceptual analysis.

4. Keeping Conceptual Analysis

Bad way to respond to objection (a) would be to reject conceptual analysis altogether: no 'is' from a 'means'. That is, could just reply that what we are really after is an analysis of knowledge, not of the meaning of 'knows', as indicated by our usage.

The reason this would be a bad way to respond is that we need to attend to facts about meaning to find out about the world. It would hardly be coherent, while still being interested, to do otherwise (e.g. to claim that one has found a deep truth about knowledge that is not connected to the meaning of 'knows', or to what everyone else means by 'knows').

Consider the following argument:

P1: 'Knowledge' refers unequivocally to the relation *justifiably truly believes*.

P2: 'Knowledge' refers unequivocally to the relation *knows*.

C: The relation *knows* is the relation *justifiably truly believes*.

Assuming P1 or other facts of that form could, in principle, be known through conceptual analysis, opponents of any such argument (i.e. opponents of conceptual analysis) must reject either P2 or deny that C follows from P1 and P2. That is, they must reject either (an instance of) the R-schema

't' refers to t,

or (an instance of) the R-inference pattern (NB t ranges over tokens of referring terms):

P1: t refers unequivocally to α .

P2: t refers unequivocally to β .

C: $\alpha = \beta$.

5. Against the Psychologists

Objection (c): why think that short lists of necessary and sufficient conditions are preferable? Haven't psychologists shown that our concepts don't have simple necessary and sufficient conditions? (Relevant because necessary and sufficient conditions are paradigmatically short/simple theories.)

Do those experiments show that ordinary terms, and perhaps terms of interest to philosophers (e.g. 'knowledge'), have Wittgensteinian family resemblance conditions rather than dictionary-like necessary and sufficient conditions?

No, for three reasons.

(1) Experiments are unlikely to generalize from ordinary terms to terms of philosophical interest. If it is easy, for any candidate list of necessary and sufficient conditions for a concept F , to find some thing that is an F but that lacks some of the listed properties, that is evidence for F 's being a cluster concept. But with 'knows', for example, it is, on the contrary, much easier to find cases which satisfy the analysis but which are not cases of knowledge.

(2) It is not a matter of accident, or magic, that some thing is an F , for pretty much any concept F . Usually we can answer the question why a thing is an F , precisely by listing necessary and sufficient conditions for being an F .

(3) There is something wrong with experiments cited, since fruit is clearly not a cluster concept. A fruit is "the edible product of a tree, shrub or other plant, consisting of the seed and its envelope".

Note that it does not quite do merely to respond by charging the proposal with incoherence. Of course suitably resembling some paradigmatic member of the family is itself a necessary and sufficient condition, but it may well, in some cases, involve a pretty unnatural and gerrymandered property.

6. In defence of analysis

At best, have shown that if 'knows' has a short analysis, it is that given by the JTB theory, notwithstanding Gettier intuitions. But (objection (d)) perhaps 'knows' does not have an analysis at all.

'Knows' might fall into one of three broad camps:

- (a) it might be rather unnatural,
- (b) it might be fairly natural in virtue of its relation to other, more natural properties,
or
- (c) it might be a primitive natural property, one that does not derive its naturalness from anything else.

Respond to objection by arguing against (a) and (c), and for (b).

Against (c): We usually judge whether someone in a certain situation has knowledge by examining the circumstances, which are the facts in virtue of (s)he has/doesn't have knowledge in that situation. That is not inductive or abductive reasoning; we feel that those facts *constitute* the person's having (or not having) knowledge.

Against (a): Well, this is a live possibility, and it does happen with many terms, including colour terms, and terms denoting (what Sider calls) maximal properties such as being a

coin: something is a coin* iff it is suitable to be a coin in every respect save that it may massively overlap a coin, i.e. a coin* is a piece of metal (or suitable substance) that is (roughly) coin-shaped and is (more or less) the deliberate outcome of a process designed to produce legal tender. Assuming any collection of atoms has a fusion, there are trillions of coins* in the vicinity of any coin, and since coins do not in general overlap, the property being a coin arbitrarily and thus unnaturally picks out exactly one of these. But, we do not take 'knows' to be as indeterminate as 'blue' or 'coin', despite the existence of some rather strong grounds for indeterminacy in it. Also, we take disputes between users of 'knows' to be genuine.

The following all seem plausible but are jointly inconsistent:

- 1- Knowledge supervenes on evidence: if two people (not necessarily in the same possible world) have the same evidence, they know the same things.
- 2- We know many things about the external world.
- 3- We have the same evidence as some people who are the victims of massive deception, and who have few true beliefs about their external world.
- 4- Whatever is known is true.

There surely is a fact of the matter which of these is false. If 'knows' denoted a highly unnatural relation, why should we accept these claims, or have disputes about which one should be rejected?

Unlike with disputes about whether something is blue, or is a coin, we take two people with different intuitions about Gettier cases to be genuinely disagreeing.

Finally, we are prepared to use 'knows' in inductive and analogical arguments; we would not do this if 'knows' denoted a highly unnatural relation.

For (b): all it requires is that there be some connection, perhaps only discoverable a posteriori, perhaps not even humanly comprehensible, between knowledge and other more primitively natural properties. These properties need not be denoted by any terms of English, or any other known language. It need not involve a decomposition, only a means by which knowledge is linked into the network of natural properties. Perhaps taking it to be the most general factive mental state, or the norm of assertion, can do the job.

7. Naturalness and the JTB theory

Have argued that intuition is untrustworthy in cases where following it slavishly would mean concluding that some common term denoted a rather unnatural property while accepting deviations from intuition would allow us to hold that it denoted a rather natural property. But perhaps what the Gettier cases show is that *being a justified true belief* is not a natural property (but a disjunction of *being justified because true* and *being justified, and also true*), so that this is not one of those cases.

This is an example of a good (critical) way to use intuition. It is unlikely to generalize. And it probably shows something more interesting than that the JTB theory is mistaken, namely that naturalness is not always transferred via conjunction (assuming that *being a justified belief* and *being a true belief* are natural). Then we may have to take naturalness to be primitive, as opposed to analysing it in the other two ways suggested by Lewis.