



OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

ANALYSIS

If You're so Smart Why Are You Ignorant? Epistemic Causal Paradoxes

Author(s): Adam Morton

Reviewed work(s):

Source: *Analysis*, Vol. 62, No. 2 (Apr., 2002), pp. 110-116

Published by: [Oxford University Press](#) on behalf of [The Analysis Committee](#)

Stable URL: <http://www.jstor.org/stable/3329249>

Accessed: 23/02/2012 18:55

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Oxford University Press and *The Analysis Committee* are collaborating with JSTOR to digitize, preserve and extend access to *Analysis*.

<http://www.jstor.org>

If you're so smart why are you ignorant?

Epistemic causal paradoxes

ADAM MORTON

1. Fantasia epistemologico-theologica

Once before a time there was a god, who liked to be worshipped. So he liked his creatures to have enough intellectual extravagance to conjecture his existence. In fact, he liked his creatures to have intellectual extravagance. Such creatures reminded him of himself and relieved his loneliness. So he was willing to reward intellectually adventurous animals, whose tiny animae dared to conjecture about things they had no intrinsic hope of understanding. He was happy to reward such creatures by giving them a small but not quite miniscule chance of true belief. So created he animals, of a first and a second kind, and he created a universe for them to puzzle over. Animals of the first kind had intellectual extravagance, and he contrived the universe so that it bore a remote resemblance in some respects to the conjectures that the more daring of the animals would make, given the evidence they were capable of collecting. And to express his contempt for lowly life forms that stick to the bare facts he equipped animals of the second kind not only with an aversion to conjecture but also with pattern-recognition capacities that were likely to miss many subtle twists of the evidence. Their predictions were likely to be wrong.

You are one of these animals, though you don't know which kind. You have foraged some evidence. A daring explanation of it zips vaguely through your mind; you think thoughts about curled up multidimensional spaces and particles moving backwards in time. The explanation sums up patterns in the evidence that you can think of no other way to summarize. You are poised between the extravagant conclusion that the universe is in fact as conjectured, and the timid conclusion that future evidence is likely to conform to the pattern on which the conjecture gives a handy if dubious hold. You realize that this is just the sort of situation that the god was anticipating. (That there is such a god with such preferences is of course another daring conjecture, but you leave that out of the calculation.) So if you accept the extravagant conclusion you are most likely an animal of the first kind. Your explanation has a small chance of being true and a slightly larger chance of delivering true predictions. And if you accept the timid conclusion you are most likely an animal of the second kind. There are

doubts even about your predictions of future data. This suggests that you maximize your chances of truth by believing the daring conjecture. But an alternative line of reasoning also occurs to you. If you are an animal of the first kind then your timid conclusion is more likely to be true than your extravagant conclusion. So too is it if you are an animal of the second kind. So either way ...¹

2. *Dominance versus maximization*

The competing factors behind considerations about what belief to hold in this case should have a familiar air. On the one hand we have a maximization argument: considerations pointing towards a greater chance of truth for one belief than another. On the other hand we have a dominance argument: there are only two possible cases and, in either one, one belief is better than the other. And the two seem to conflict. Put this way, the example is reminiscent of the cases which motivate the switch from evidential to causal decision theory, such as Jeffrey's nicotine addiction case or Newcomb's problem, or the examples in the influential paper by Gibbard and Harper. (See Jeffrey 1983: 15–23, Gibbard and Harper 1978, Joyce 1999. I am presupposing familiarity with these examples. A good collection is Campbell and Sowden 1985, but the literature has continued to grow since.) Like many of those cases, the example depends on a probabilistic feedback between the state whose rationality is being expressed and the factors on which its rationality depends. Moreover, the example generates temporal asymmetries much like those that causal decision cases often do. Under suitable conditions it is – arguably – rational to acquire a disposition to perform future acts (choose 1-boxishly, drink toxins, cooperate in Prisoners' Dilemmas), although when the time to act arrives it will be rational to act contrary to those dispositions. Similarly, if in the case above we give an individual a choice of becoming an animal of either type then there is a clear case for becoming the adventurous type one, a case which does not wipe out the reasons for accepting the more cautious belief at the epistemic crunch.

(Another way of putting it: the ur-type of causal decision case is the Calvinist argument for good behaviour. Virtuous acts are evidence that you are – were always – one of the elect, so although sins do not cause damna-

¹ David McCarthy points out to me that the example could also be made in terms of perception. Animals of the first kind have accurate colour vision and are inclined to make colour judgements even when the light is not adequate. Animals of the second kind have less accurate colour vision but are inclined to stick to what is clearly visible. So if you find yourself making a somewhat unlikely colour judgement you are both straining at the limits of perceptual knowledge and acquiring 'evidence' that your judgement is accurate.

tion sinners ought to regard their crimes as very bad news. The case in §1 is an epistemic analogue of this situation.)

The existence of epistemic analogues of causal decision cases is puzzling, though. Those cases show that it can be good news that one is doing something, though it is the less good thing to do. Now certainly it can be good news that one believes something irrational. It might be a sign that one is likely to be accepted by some religious community, for example, or that some anti-intellectual person might love one. But that is an instrumental kind of good news. What a purely epistemic analogue would require would be that it be good evidence of the truth of your belief that you had irrationally acquired it. So why would the acquisition then be irrational? Some clarification is needed.

3. *Simpsonian statistics and reliable beliefs*

Behind many causal decision cases there is a common statistical pattern. It is usually called Simpson's paradox – see Cartwright 1983: 36–38 – and amounts to the following surprising but non-paradoxical fact. We have a population of individuals, which we divide into two sub-populations. There are two groups of individuals, A and B, found in both sub-populations, and an attribute X that such individuals can have. Then it is possible that in both sub-populations the proportion of As that are X is greater than the proportion of Bs that are X, but in the population as a whole the proportion of As that are X is *smaller* than that of Bs that are X.

Here is an example, which should both make it less puzzling that such statistics can occur and make the connection with the topic of this paper. You have submitted a research paper for an interdisciplinary journal. Your paper is on decision theory and so could be thought of as philosophy, economics, statistics or psychology. The editor of the journal assigns papers to two teams of referees, those for humanities and those for social science. These teams then both label each paper with the name of a discipline, for use in their own record-keeping, and, independently, evaluate its quality, leading to eventual acceptance or rejection. You don't know which set of referees your paper has gone to, let alone whether it has been accepted. But you do learn, via an email indiscretion, that it has been labelled as philosophy. And you know from fairly robust past evidence that disproportionately many 'philosophy' papers are rejected. So you feel pessimistic, inclined to believe that your paper will be rejected.

However, there are facts that you do not yet know. The papers are accepted or rejected entirely on the basis of quality, and 'philosophy' papers are generally of high quality. In fact, both the humanities and the

social science referees tend to accept disproportionately many papers they have labelled 'philosophy'. The reason that disproportionately many 'philosophy' papers are rejected overall is that the humanities referees both label more of the papers assigned to them as 'philosophy' and reject more of the papers assigned to them than the social science referees do. Thus being in the pool of papers with the higher rejection rate correlates with being labelled 'philosophy', even though in either pool being labelled 'philosophy' correlates with acceptance.²

Now consider again your belief that your paper will be rejected. It is based on definite statistics, and not plainly irrational. But there is something deviant about it. For the fact that your paper has got the 'philosophy' label is correlated in both humanities and social science with a greater chance of acceptance. So in that way learning its labelling is *good* news. There are two ways to express the deviance of your belief. First, the reasoning on which the belief is based, though not fallacious, is not a reliable source of true beliefs. It is not the case that had your paper not been labelled 'philosophy' it would have been less likely to be rejected. In the nearest worlds in which it is not labelled 'philosophy' it is more likely to be rejected, since in the nearest worlds it is still with whichever referees actually have it, and 'philosophy' labelling tends to be associated with a judgement of quality that goes with acceptance.³ Second, the belief is unstable in relation to future evidence. You may learn tomorrow that your paper is being considered by the humanities referees, or you may learn that it is in the hands of the social scientists. Suppose that you have learned the full underlying statistics of acceptance and rejection. Then if you learn that the humanities referees have your paper you will find the fact that it has been labelled 'philosophy' comforting, a suggestion of acceptance. And so will you if you learn the social scientists have it. So the grounds for your earlier belief that the paper will be rejected are going to be undermined either way the future information comes in, as you can now know.

² If you want numbers: There are 1000 papers being considered by each set of referees. The humanities referees give the label 'philosophy' to 100, and accept 100, including 12 they have so labelled. The social scientists give the label 'philosophy' to 10 and accept 200, including 3 they have so labelled. So overall 0.15 of papers are accepted but only 0.13 of those labelled 'philosophy'. But in humanities 0.1 of all papers are accepted and 0.12 of 'philosophy', and in social science 0.2 of all papers are accepted and 0.3 of 'philosophy'. (The numbers alone will make the statistics Simpsonian, but we need to know the causal set-up before we can make a suitable case out of it.)

³ I mean this only as a rough gesture to the kind of counterfactual consideration needed. Getting the counterfactuals right is a hard problem. But it is one that any externalist epistemology must face, as must any causal decision theory.

It is important to distinguish between two contexts in which to assess the problematic quality of the belief that your paper will be rejected. Before you know the full statistics, knowing only that ‘philosophy’ is correlated with rejection, your belief is unreliable in that there is not a suitable counterfactual link between it and its evidence, and in that given true but unknown fuller statistics it can be undermined however future evidence comes in. At this stage we can diagnose the situation as: internalistically justified but externalistically unjustified. (This is analogous to the situation of a person in a Jeffrey-style case, in which there is a common gene that disposes both to nicotine addiction and to lung cancer, who knows only that smoking is correlated with lung cancer. Such a person’s reasonable choice of not smoking is not a reliable means of preventing the disease.) Suppose that you continue to hold the belief even later, when you know the full statistics. Then the status of your belief is more puzzling. It is not simply irrational: after all, the ‘philosophy’ label is a sign that the paper has probably gone to the humanities referees, who reject more papers. But it is still unreliable, as you can see, and as a sign of this you should be able to see that whichever way information about the referees comes in, when it does you will be able to cite the paper’s labelling as a reason for greater confidence in its acceptance. (Again an analogy with Jeffrey’s case may help. Suppose that if you have the relevant gene smoking makes lung cancer less likely, as it does if you do not, although overall smoking is correlated with a greater risk of lung cancer. This could be because, although the gene is the greatest cause of lung cancer, stress, which smoking reduces, is a minor cause. Then on finding yourself smoking you could either say, ‘Oh-oh: that suggests I have the fatal gene’, or you could say ‘Good, whether or not I have the gene this will give some measure of protection.’)

The important point here is not whether the belief in question is, all things considered, justified. (My own conviction is that ‘justified’, ‘rational’ and the like are too crude labels to do justice to the issues here. So for that matter are externalist alternatives such as ‘reliable’, or ‘counterfactually linked’, without a lot of fine-tuning.) The important point is the similarity of the issues to those that arise with causal decision cases. With many puzzling epistemic and decision-theoretic cases the root of the puzzle is Simpson-shaped statistics where one set of correlations has counterfactual force and the other does not.

4. *One-coining/two-coining*

There are thus systematic reasons for thinking that some aspects of the causal decision cases have epistemic analogues, in fact that they are fundamentally epistemic phenomena. So is there an epistemic analogue of the most intuition-dividing causal decision case, Newcomb’s problem? I think there is.

An infallible predictor predicts your inferences in the following situation. You see a film of two coins being tossed. One of them – coin A – comes down heads ten times in a row, and the other – coin B – is roughly even H/T. You have to come to a conclusion about the eleventh toss, which has occurred and whose result is hidden from you. The predictor manipulates your chances of making a true prediction as follows. He has a collection of coins; some are biased to varying degrees to H or T and some are fair. All are tossed ten times and the sequences filmed. If the predictor's prediction is that you will believe that the eleventh toss of coin A is heads and that the eleventh toss of coin B is also heads, then the sequence that you see is chosen so that coin A has a heads-always bias and coin B has a bias to heads although in the chosen sequence it came down roughly even H/T. If the predictor's prediction is that you will believe just that the eleventh toss of A is heads (and have no belief about that of B) then the sequence that you see is chosen so that both coin A and coin B are fair. As a result, if you are (infallibly!) predicted to make the more extravagant prediction then at least part of your prediction is almost certainly true (and the other part has a more than 50/50 chance). But if you are predicted to make the more modest prediction then that prediction has just a 50/50 chance of truth. (So the probability of a true belief in the first case is more than 0.5, and in the second case it is 0.5.) That can be taken as an argument for the 2-coin prediction. But on the other hand the coins have already been filmed, and whichever ones they are your belief will be reasonably safe if you predict just one coin and verging on risky if you predict for both.

Is possibly irrational epistemic extravagance in this case analogous to possibly irrational practical restraint in the Newcomb situation? There is certainly a formal similarity, and if you try the case out on your friends you will find that they are divided on what is the 'right' belief to accept.⁴

University of Oklahoma
Norman, OK 73019, USA
adammorton@ou.edu

References

- Campbell, R. and L. Sowden. 1985. *Paradoxes of Cooperation and Rationality*. Vancouver: University of British Columbia Press.
- Cartwright, N. 1983. *How the Laws of Physics Lie*. New York: Oxford University Press.
- Gibbard, A. and W. Harper. 1978. Counterfactuals and two kinds of expected utility. In *Foundations and Applications of Decision Theory*, vol. 1, ed. C. Hooker, J. Leach

⁴ I have had invaluable help from James Joyce, and helpful comments from James Bell, and Remy Debes. An audience at Ohio State was very patient when I bumbled and mangled a central part of the argument.

- and E. McClennen, 125–62. Dordrecht: Reidel. Repr. in Campbell and Sowden 1985.
- Joyce, J. 1999. *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.
- Nozick, R. 1969. Newcomb's problem and two principles of choice. In *Essays in Honor of Carl G. Hempel*, ed. N. Rescher, 114–46. Dordrecht: Reidel. Repr. in Campbell and Sowden 1985.

A test for theories of belief ascription

BRYAN FRANCES

If publication rates are an accurate guide to popularity, then the two currently most popular approaches to belief ascription are Millianism and Contextualism. The former rules out ordinary Frege cases such as Lois believing that Superman flies while failing to believe that Kent flies. The latter approach, as I will define it for the purposes of this essay, includes all theories according to which standard uses of belief reports of the form 'A believes that *a* is *F*' and 'A believes that *b* is *F*', where '*a*' and '*b*' are co-referential proper names or general terms, may differ in truth-value even though the proposition that *a* is *F* is identical to the proposition that *b* is *F*. For those who enthusiastically concur with current conventional wisdom, the primary argument of this essay, centring on a puzzle about belief reports, is modest; for those who don't buy the fashionable views, the argument is challenging. Facts about the puzzle case and commonsensical principles about belief entail that one of the two approaches must be correct, although the puzzle is perhaps best thought of as a test case for all theories of belief ascription. If one thinks that any version of either Millianism or Contextualism must flout important intuitive principles regarding belief, then one is left with the conclusion that *any* adequate theory of belief ascription will have to be significantly revisionary with regard to semantic intuitions. If so, then we can forget about letting those intuitions have the weight that they are commonly accorded in theory construction for belief ascription (and, as a consequence, belief).

I present the argument by altering, elaborating and examining a puzzle that I introduced in a rough-and-ready way in earlier papers (1998, 1999). If the reader has had enough of puzzles about belief or belief ascription, I