"Click!" Bait for Causalists

HUW PRICE AND YANG LIU
University of Cambridge

Causalists and Evidentialists can agree about the right course of action in an (apparent) Newcomb problem, if the causal facts are not as initially they seem. If declining $1,000 *causes* the Predictor to have placed $1m in the opaque box, CDT agrees with EDT that one-boxing is rational. This creates a difficulty for Causalists. We explain the problem with reference to Dummett's work on backward causation and Lewis's on chance and crystal balls. We show that the possibility that the causal facts might be properly judged to be non-standard in Newcomb problems leads to a dilemma for Causalism. One horn embraces a subjectivist understanding of causation, in a sense analogous to Lewis's own subjectivist conception of objective chance. In this case the analogy with chance reveals a terminological choice point, such that either (i) CDT is completely reconciled with EDT, or (ii) EDT takes precedence in the cases in which the two theories give different recommendations. The other horn of the dilemma rejects subjectivism, but now the analogy with chance suggests that it is simply mysterious why causation so construed should constrain rational action.

KEYWORDS: decision theory, Newcomb's problem, Dummett, subjectivism, causation, chance.

1. *Euthyphro* for Causalists

Causality and rational means–end deliberation seem to go hand-in-hand, in normal circumstances. Putting it roughly, *A* causes *B* if and only if, other things being equal, it would be rational for an agent who desired *B* to *do* or *bring about A,* in order to realise *B.* Much of the interest of Newcomb problems lies in the fact that they seem (to some!) to provide counterexamples – cases in which rational means–end reasoning goes one way, causal connectedness goes another. 'Evidentialists' argue the case for such counterexamples, while 'Causalists' oppose it.

Newcomb problems aside for the moment, let's think about the nature of the relationship between causality and rational action. Let's ask the *Euthyphro* question. Is it the causal connection between *A* and *B* that makes it rational to do *A* to achieve *B?* Or does the rationality of the latter somehow constitute or ground the fact that *A* causes *B?* Let's call the first option the *objectivist* view, the second the *subjectivist* view.[1]

---

1. These are not necessarily the happiest terms for the views in question. We use them here because they are the terms in use in the case of chance, to which we want to draw an analogy. Elsewhere (Price, 2001, 2007) we have recommended the term 'pragmatism' for what we here call 'subjectivism'.

The standard understanding of Causal Decision Theory (CDT) seems to presuppose objectivism. CDT is offered as a formalised prescriptive theory of rational decision, couched in terms of an agent's beliefs about objective relations of causal dependence. It remains a subjective theory in the sense that it is couched in terms of what the agent *believes*, rather than in terms of how the world actually *is.* But it is objectivist *about causation*, in that the crucial beliefs concern objective causal relations.[2]

In our view, this understanding of CDT is not compulsory. Indeed, it cannot be compulsory so long as subjectivism about causation is itself a live option. A subjectivist will agree that CDT gets things right *in a descriptive sense,* formalising the relation between causal beliefs and rational deliberation. But since the subjectivist takes this relation to be constitutive of causal belief, CDT is in effect being 'read backwards', characterising causality in terms of rational action.[3]

Among the advantages of this subjectivist version of CDT is that it allows a pleasing reconciliation between CDT and Evidential Decision Theory (EDT) – in effect, a dissolution of Newcomb's problem. This reconciliation is the end goal we have in mind, but our main concern here will be one step back: we want to present an objection to the orthodox objectivist version of Causalism. The objection turns on a dilemma that confronts Causalism, when presented with Newcomb problems. So far as we know, this dilemma has not been noted in this form, either by Causalists or their Evidentialist opponents.

One horn of the dilemma is to embrace a subjectivist understanding of the relation between causality and rational deliberation – unacceptable for an orthodox objectivist Causalist, but a welcome option for subjectivists themselves. As we'll see, this option either brings CDT into complete alignment with EDT, in standard Newcomb cases, or explains why EDT takes precedence, in versions of the view that allow some small divergence. (The difference between these versions depends on a terminological choice about the use of 'causation', in difficult cases.)

The other horn of the dilemma retains objectivism about causality, but at the cost, as we'll argue, of making it mysterious why causality should be thought to be the arbiter of rational action, in the way that CDT proposes.

It turns out that the structure of the options here is closely parallel to those in the case of chance, in a tradition whose subjectivist roots trace back (for our purposes) to the kind of theories of objective chance proposed by D H Mellor and David Lewis. (We take our subjective/objective terminology from Lewis.) This analogy between causation and chance has not been widely noticed,[4] perhaps because objectivist intuitions are even more prevalent for causation than for chance, so that it has been hard to see the pressure for a degree of subjectivism that flows from decision theory (a pressure to which both Mellor and Lewis were sensitive in the case of chance).

---

2. That is, for our purposes, relations whose nature is to be understood independently of their role in CDT itself.

3. As we shall see in §4 below, this is compatible with the view that CDT functions prescriptively in normal cases, where the relevant causal relations are not in doubt.

4. Lewis himself missed it, apparently – a fact reflected, as we'll explain, in his unwavering commitment to two-boxing in the classic Newcomb problem.

To exhibit this pressure in the causal case, we'll begin with the observation that Newcomb problems raise the possibility that the facts of causation are not what they seem (in particular, not what the orthodox Causalist assumes them to be), and that this explains the appearance that CDT and EDT offer different recommendations. The idea that Newcomb problems might involve non-standard causal relations certainly isn't new. (Like so much else in this field, it goes back to Nozick's original (1969) paper, and is explored in mote detail in Mackie (1977).) What is new, so far as we know, is a full appreciation of how this possibility turns into a dilemma for orthodox objectivist Causalism, once the *Euthyphro* issue is in view. We'll approach the issue via the work of Michael Dummett, who gets close to the points we want to make.

2. Bringing about the past?

Dummett published two famous early papers about the direction of causation, ten years apart (Dummett, 1954, 1964). Dummett has a similar cluster of concerns in both papers: why we take causes to *precede* their effects in time (or are at least to be *no later than* their effects, as in cases of simultaneous causation); why we think it makes sense to act for the sake of *later* ends but not for *earlier* ends; and, especially, whether the latter principle might coherently admit exceptions – cases in which it would make sense to act to *bring about* an earlier end.

In the 1954 paper Dummett satisfies himself that causation itself is unidirectional. He argues that causes are the *beginnings* of explanatory chains or processes. Their effects are later events in these chains, and hence lie later in time.[5]

> The temporal direction of causation, from earlier to later, comes in because we regard a cause as *starting off* a process: that is to say, the fact that at any one moment the process is going on is sufficiently explained if we can explain what began it. Causes are simultaneous with their immediate effects, but precede their remote effects. (1954, 29)

Dummett then asks whether, "given that *in general* causality works in the earlier-to-later direction, we could not recognise a few exceptions to this general rule."

> If we find certain phenomena which can apparently be explained only by reference to later events, can we not admit that in these few cases we have events whose causes are subsequent to them in time?

"I think it is clear that we cannot", Dummett replies:

---

5. Dummett thinks that the continuation of such processes does not require explanation, once they get going. He doesn't raise the question whether the past–future asymmetry of the picture is merely a conventional matter, on a par with Hume's stipulation that causes are the *earlier* and effects the *later* of pairs of constantly conjoined events. Another possibility would be to tie it to some fundamental directionality in time, but this would sit oddly with his determination *not* to take past–future asymmetries for granted.

> One event is causally connected with another if it is either its immediate cause or it is one of its remote causes. To be its immediate cause, it must be simultaneous with it. If it is a remote cause of it, then it is so in virtue of being the immediate cause of the beginning of some process whose continuance is not regarded as requiring explanation, and whose arrival at a certain stage is in turn the immediate cause of the event in question. . . . An event subsequent to the event whose occurrence we were wishing to explain could fall into neither of these two categories. (31)

However, Dummett continues,

> [t]his explanation why an effect cannot precede its cause does not . . . end the matter. We may observe that the occurrence of an event of a certain kind is a sufficient condition for the previous occurrence of an event of another kind; and, having observed this, we might, under certain conditions, offer the occurrence of the later event, not indeed as a causal, but as a *quasi-causal* explanation of the occurrence of the earlier.[6] (31–32, our emphasis)

Dummett then discusses the possibility of such future-to-past 'quasi-causation' at considerable length. He imagines the following objection:

> [T]o suppose that the occurrence of an event could ever be explained by reference to a subsequent event involves that it might also be reasonable to bring about an event in order that a *past* event should have occurred, an event previous to the action. To attempt to do this would plainly be nonsensical, and hence the idea of explaining an event by reference to a later event is nonsensical in its turn. (34–35)

In response, Dummett argues that the idea of acting to bring about a past is far from nonsensical. He concludes that in exceptional circumstances it would be not merely coherent but entirely rational to act in this way. The 1954 paper closes with this example of the kind of case he has in mind.

> Imagine that I find that if I utter the word "Click!" before opening an envelope, that envelope never turns out to contain a bill; having made this discovery, I keep up the practice for several months, and upon investigation can unearth no ordinary reason for my having received no bill during that period. It would then not be irrational for me to utter the word "Click!" before opening an envelope in order that the letter should not be a bill; it would be superstitious in no stronger sense than that in which belief in causal laws is superstitious. Someone might argue: Either the envelope already contains a bill, or it does not; your uttering the

---

6. It is the lack of a suitable embedding in a (past-to-future) process that prevents it from being *genuinely* causal, in Dummett's view.

word "Click!" is therefore either redundant or fruitless. I am not however necessarily asserting that my uttering the word "Click!" changes a bill into a letter from a friend; I am asserting (let us suppose) that it prevents anyone from sending me a bill the previous day. Admittedly in this case it follows from my saying "Click!" that if I had looked at the letter before I said it, it would not have been a bill; but from this it does not follow that the chances of its being a bill are the same whether I say "Click!" or not. If I observe that saying "Click!" appears to be a sufficient condition for its not being a bill, then my saying "Click!" is good evidence for its not being a bill; and if it is asked what is the point of collecting evidence for what can be found out for certain, with little trouble, the answer is that this evidence is not merely collected but brought about. Nothing can alter the fact that if one were really to have strong grounds for believing in such a regularity as this, and no alternative (causal) explanation for it, then it **could not but be rational** to believe in it and to make use of it. (1954, 43–44, emphasis added)

2.1. From Dummett to Newcomb

Dummett's "Click!" example is easily transformed into a Newcomb problem. We simply need to suppose that Dummett incurs a small cost, say tuppence, for saying "Click!". Not saying "Click!" then *dominates* saying "Click!" – whether or not the envelope contains a bill, Dummett does better by saving his tuppence. Presumably Dummett will still feel that "it could not but be rational" to say "Click!", despite the expense. (We are assuming that the usual expense of Dummett's bills, averaged over all his incoming mail, was more than tuppence per envelope – even in 1954.)

If we construe the example in this way, Dummett counts in modern terms as an Evidentialist. What will he say to someone who disagrees about the rationality of paying to say "Click!"? With some violence to his own Oxford idiolect, we can imagine him offering the classic Evidentialist response: "If you're so smart, why ain'cha rich?" And such an opponent – a Causalist, in our terms – will offer the standard reply. In each individual case Dummett could have saved his tuppence. As Dummett himself says, 'it follows from my saying "Click!" that if I had looked at the letter before I said it, it would not have been a bill' (44). So, claims the Causalist, Dummett always chooses the sub-optimal option. *Contra* Dummett, it could not but be *irrational* to waste one's money on the "Click!"

Causalists shouldn't relax, however. As we noted, Dummett himself calls what the "Click!" case displays 'quasi-causation'. By his later piece he's dropped the 'quasi', and talks explicitly about 'backward causality':

When, however, we consider ourselves as agents, and consider causal laws governing events in which we can intervene, the notion of backwards causality seems to generate absurdities. If an event C is considered as the cause of a preceding event D, then it would be open to us to bring about C in order that the event D should have occurred. But the conception

> of doing something in order that something else should have happened
> appears to be intrinsically absurd: it apparently follows that backwards
> causation must also be absurd in any realm in which we can operate as
> agents. (1964, 340)

The rest of the 1964 paper is devoted to rebutting this appearance of intrinsic absurdity, covering similar ground to the earlier piece. Dummett does not re-use the "Click!" example, but he does offer a more complicated example with essentially the same structure.

But if we allow this shift in terminology (from 'quasi-causation' to 'causation' simpliciter), then Causalism will recommend one-boxing, apparently, in the same case. Dummett imagines that by saying "Click!" he *brings it about* that the envelope does not contain a bill. The 1964 paper is called 'Bringing About the Past', and Dummett's main claim is that if there were such an exceptional case, we would be entitled to regard ourselves as bringing about the past in precisely the same sense as, ordinarily, we regard ourselves as bringing about the future.

We can now see why Dummett represents such a challenge to conventional two-boxing Causalism. By putting the causal terminology itself into play in this way, he threatens to defend one-boxing with the Causalist's own weapons. The Causalist can object to the terminological move, of course, but this won't be straightforward. She will be trying to defend causal orthodoxy in a case in which the rationality of an associated action is in dispute. Nancy Cartwright (1979) taught us to associate the distinction between causal and merely associative correlations with the distinction between effective and ineffective strategies. But in the present case, Dummett and the two-boxer Causalist disagree about whether paying to say "Click!" is an effective strategy, so there's no easy way to apply Cartwright's criterion to resolve the terminological issue. (This thought will lead us to our dilemma.)

True, there's still the matter of the counterfactuals. Could they get the Causalist off the hook? We'll come back to this (as Dummett himself did, in later work), but is worth noting that even in the 1954 paper, Dummett notes that our counterfactual intuitions can't be trusted in these kind of cases:

> [W]here we are concerned with a regularity which works counter to
> ordinary causal regularities, our normal methods of deciding the truth of
> a contra-factual conditional break down. (1954, 38)

2.2. *Dummett's Dilemma*

Dummett's trajectory between 1954 and 1964 thus highlights a difficulty for orthodox CDT – one that gets far too little attention in discussions of Newcomb problems, in our view. It turns on the fact that Newcomb problems are, by their nature, cases in which the relevant causal facts may be disputed. When we unpick the options in such a dispute, we find the dilemma that confronts Causalism.

Why isn't this issue already well known? Perhaps because neither side has seen a motivation to explore it. It is an unwelcome option for the orthodox objectivist

Causalist, who has an obvious interest in assuming that the causal facts are sufficiently solid to ground rational decision theory, in difficult cases; and an apparent irrelevance for the Evidentialist, who professes not to be interested in causation in the first place. So it lies in something of a blindspot, an unappealing corner from both points of view. Yet in that blindspot lies something important for both sides, in our view: the prospect of a reconciliation about Newcomb's problem that endorses the key insights both of Causalism and Evidentialism.[7]

To expose the dilemma for Causalism, imagine the kind of case Dummett points us to, where we have two views of causation on the table. They differ in extension, so cannot both be a true guide to rational decision, by CDT's lights, if 'rational' is to remain univocal. How are we to adjudicate? The Causalist wants the understanding of causation that gives the rational choice, when plugged into CDT. But how are we to know which is the rational choice, given that we subscribe to CDT, and we are unclear about the meaning of 'cause'? We have an equation with too many variables.

At the heart of this problem lies our dilemma. On one horn, the Causalist attempts to meet the challenge to her use of causal terminology in objectivist terms. In other words, she relies on some principle *other than an explicit rationality criterion* to adjudicate the use of 'cause' and related terms. For example, she insists that causes be no later than their effects. To defend CDT, she then faces the problem of explaining why a rule for rational action should always respect this criterion. Dummett's "Click!" case will be a direct challenge.[8]

On the other horn, the Causalist relies on an explicit rationality criterion to fix the extension of causal terminology – in effect, she opts for subjectivism, reading CDT 'back to front', as an analytic constraint on the notion of causal dependence. But now CDT is no use at all as a guide to the rational course of action, in disputed cases. On the contrary, a ruling on the application of causal notions must lie *downstream* of the resolution of such disputes, in contested cases. (Again, "Click!" provides an example.) And if the subjectivism itself weren't bad enough, the only available account of rationality will be that of the Evidentialist, presumably.[9] Swept away downstream, Causalists risk being washed ashore in the land of one-boxers.

In summary, this is the dilemma. If the principle criterion of causality is not that of supporting rational action as such, then the issue can be raised as to whether (and if so why) the criterion always keeps step with rational action. If the criterion is rational action, then the causal structure cannot be resolved while rational action is in dispute – and that way lies subjectivism, and arguably Evidentialism.

In honour of Dummett, we'll call this *Dummett's Dilemma* (DD). Dummett himself

---

7. More importantly still, this reconciliation is the clue to an important ingredient of an understanding of causation itself, in our view.

8. Dummett's method seems to be applicable to whatever criterion the Causalist proposes at this point. We simply imagine a case in which we encounter a reliable correlation that doesn't meet the criterion in question, but which nevertheless seems to be of practical value in the way that saying "Click!" is in Dummett's case.

9. A subjectivist can't characterise rationality in causal terms, at least initially – that would simply be circular.

does not describe DD explicitly. Indeed, as we'll see in a moment, he makes a move that prevents him from developing its full potential as an objection to Causalism. But he does exemplify it, in the sense that the problem he represents for an orthodox two-boxing Causalist lines up with the first horn in 1954 and the second horn in 1964. In 1954 the implicit challenge is: Why is *Causal* Decision Theory rather than *Quasi-causal* Decision Theory the appropriate guide to rationality in the "Click!" case, if the difference between the two rests merely on a criterion couched in terms of the direction of time? In 1964 it is: On what grounds *independent of a knowledge of causal structure* can we declare the "Click!" policy irrational, in order to deny that there is backward causation?

2.3. *Dummett on counterfactuals*

Let's return to the counterfactuals. Do they provide a middle way for objectivist Causalism, an objective criterion for causal connection that wears its link to rationality on its face? It may seem so. Imagine this objection to Dummett's reading of the "Click!" example. "But Dummett," says the opponent,

> if you had not said "Click!" the envelope would still not have contained a bill, and you would have been tuppence the richer. The truth of that counterfactual shows that there's no causation involved, and explains the sense in which you did the irrational thing!

This is the argument for irrationality involved in the standard Causalist response to the "Why ain'cha rich?" challenge (see Lewis, 1981).

We know how Dummett would have responded to this objection, for he discusses Newcomb problems in two papers (1986, 1987) from the mid-1980s. He retains his commitment to one-boxing, and his main response to the two-boxer's appeal to counterfactuals is to deny that ordinary counterfactual reasoning is a helpful guide to rationality in such cases.[10]

> I thus have a choice between doing something that will, with a very high
> probability, result in my getting $1,000 and doing something that will,

---

10. In general, Dummett thinks that counterfactuals are a great deal more obscure than many philosophers tend to assume:

> [T]he meaning of a subjunctive conditional is enormously obscure; they have been a thorn in the side of analytic philosophers for many decades. Even more obscure than their meaning is their point: when we have devised some theory to explain their truth-conditions, we are at a loss to explain for what purpose we want to have sentences with those truth-conditions in the language at all. Very important things may hang on them: whether the accused is guilty of murder may depend on whether the deceased would still have died if the accused had not acted as he did. But that is no answer to the question about their point: for, in such a case, we have given the subjunctive conditional such a point, and the question is why we should do so. (1986, 355)

with a very high probability, result in my getting $10,000. Plainly, . . . the rational thing for me to do is the second. After I have done it, the rules governing the assertion of counterfactual conditionals may entitle me to assert, "If I had taken both boxes, I should have got $11,000"; *but that is only a remark about our use of counterfactual conditionals.* Before I make my choice, I should be a fool to disregard the high probability of the statement "If I take both boxes, I shall get only $1,000". That is not merely a remark about our use of the word "probability", nor even about our use of the word "rational", but about what it is rational to do. (1986, 375, emphasis added)

In our view, however, this leaves Dummett vulnerable to the objection that the ordinary usage of counterfactuals and the ordinary usage of the notion of rationality are part of the same package. For example, it might be held to be an essential criterion for the rationality of an action that one not be in a position to know in advance that if one does it, one would have done better by doing something else. This claim might be disputed, but while it is in play, Dummett will not be free to dismiss the relevance of ordinary counterfactual claims, without also threatening the ordinary use of claims about rationality. In other words, his own claim about "what it is rational to do" will be just as suspect (despite his insistence to the contrary).

At one point Dummett makes what we take to be the right move against the Causalist appeal to counterfactuals. Describing a slightly different version of the Newcomb problem, Dummett says this:

Smith [the two-boxer], when his turn comes, argues, "The money is already in the boxes. Whether each contains £10 or £1,000, I shall get twice as much if I choose to open both boxes"; he accordingly does so, and receives £20 in all. Jones [the one-boxer] says to him, "You were a fool: you should have chosen to open only one box", but Smith replies, "If I had opened only one, I should have got no more than £10". (1986, 356)

So far this is the standard exchange, but Dummett now notes that the one-boxer can challenge the two-boxer's counterfactual:

Jones may well retort, "Not at all: if you had opened only one, the psychologist would previously have put £1,000 in each".[11] (1986, 356)

---

11. Causalists may object that Jones is using a 'backtracking' counterfactual, and so not really disagreeing with Smith, who had the regular kind of counterfactual in mind. But this misses the point. Jones' counterfactual is indeed backtracking in the strictly temporal sense, but (at least as we are reading Dummett here) the question as to whether it is backtracking in the causal sense is precisely what is at issue. (The term 'backtracking' is unhelpfully ambiguous in this context.)

The same point is missed in a classic discussion by Horgan (1981). Horgan points out that there is a kind of conceptual circularity in the Causalist's defence of the rationality of her decision rule, since it always comes back to a counterfactual: "If you'd made the other choice, you would have done better." He argues that Causalists have no non-circular way to justify such an account of

But Dummett doesn't press the advantage home, and slides back to his dismissive line about the ordinary use of counterfactuals:

> Now which [of Smith and Jones] is right? If the question is, which of their judgements about the truth of the counterfactual conditionals best accords with our established ways of judging such matters, it is arguable that Smith has the better case. If so, all that follows is that he has a more accurate grasp of the existing use of such sentences: it does not show that Jones will be acting foolishly when his turn comes, and he chooses to open only one box, obtaining £1,000. (1986, 356)

As we said, this lays Dummett open to the charge that counterfactuals and rationality are part of the same package, and that he can't dismiss one in this way without undercutting his own right to use the other "full voice".

A better move, in our view, is to press the dilemma for counterfactuals, just as it has been pressed for causation. In other words, it is to point out that the counterfactual characterisation of the case, like the causal characterisation, is simply contested. The objectivist Causalist now faces the same dilemma as before. If she seeks to resolve the contest by appealing to some independent criterion for the relevant kind of counterfactual dependence – e.g., a temporal asymmetry principle – the question will arise, as in Dummett's classic discussions in (1954) and (1964), as to whether and why that criterion should constrain rational deliberation. But the alternative is subjectivism, with the same consequences here as in the causal case. So the appeal to counterfactuals adds nothing new.

Dummett himself shows the same scepticism about the relevance of our use of the term "cause" as about that of counterfactuals – in both cases, he thinks that there is a danger of missing the interesting and important point, which is about whether a certain kind of practical action makes sense:

> Particularly is this so if you concentrate on how the word "cause" is used, and how it's connected with a temporal direction. Now, if someone wants to know whether it's reasonable for him to do something when his motive is that something should have happened—which some people regard as absolutely ridiculous—it's a cheat to fob him off with explanations in terms of how we use the word "cause". That's not the question. Whether it would be called a cause, or is rightly called a cause, may be an interesting question, but it is not the question he was asking. If it's not called a cause, then all right, perhaps we'll call it something else. But the question is: Is there any sense in doing this thing? (1986, 369–370)

---

rationality, as opposed to one based on the Evidentialist's 'backtracking' counterfactuals. He is right, but he misses the stronger option of putting in play the 'causal' notion of counterfactual dependence itself – of regarding that notion as contested in Newcomb cases. This option is the one on which the counterfactual version of DD depends.

Dummett seems just to take it for granted that the latter question ('Is there any sense in doing this?') is to be answered in probabilistic terms. As he puts it in one of the passages above:

> I thus have a choice between doing something that will, with a very high probability, result in my getting $1,000 and doing something that will, with a very high probability, result in my getting $10,000. Plainly, ... the rational thing for me to do is the second.

Here Causalists may feel that Dummett is simply blind to the main motivation for Causalism, namely, the observation that not all probabilistic dependencies between Acts and Outcomes are sound bases for effective strategies. (They may also feel that his formulation here – "something that will, with very high probability, *result in* ..." – simply builds in the causal notions that they themselves strive to make explicit.) So Evidentialists who wish to exploit DD need to make clear that they can steer clear of these objections. We shall now propose a way to do this, and at the same time to take maximum advantage of DD. The key is to allow the Causalists themselves to do the heavy lifting.

3. Dummettian Evidentialism

Let us introduce a character we'll call the Dummettian Evidentialist, or DEVI, for short. DEVI's strategy for dealing with Causalists goes like this. When offered something claimed to be a Newcomb problem – i.e., a case in which Evidentialism and Causalism are claimed recommend different courses of action – she initially plays dumb, claims to prefer the one-box option (or its analogue). Preparing to try to trap the Causalist in Dummet's Dilemma, she looks ahead to a hypothetical series of cases and says, "Why won'cha get rich?"

In some cases, Causalists will be able to explain why DEVI wouldn't get rich (or richer than they themselves would), after all. In the Smoking Gene case, for example, the correlation between smoking and cancer is taken to be explained by a gene that predisposes to smoking and cancer – it is a common cause of both. This means that someone who would otherwise choose to smoke, but who gives up smoking "to lower their probability of getting cancer", as they put it, will have the gene (and hence get cancer) just as often as any smoker. In this case, their reason for giving up smoking screens off the correlation between smoking and the gene. So no advantage accrues to such a person, in this kind of case, to offset the disadvantage of losing the pleasure of smoking. Giving up smoking is an *ineffective* strategy for avoiding cancer, in this case.

In this example, choosing not to smoke is the analogue of one-boxing – in other words, it is the option favoured (or said by Causalists to be favoured[12]) by Evidentialists.

---

12. The qualification is important. On the Evidentialist side a long tradition is devoted to arguing that EDT is not committed to one-boxing (or its analogue) in many claimed Newcomb cases, including so-called medical Newcomb problems such as Smoking Gene. The well-known Tickle Defence is one argument of this sort. The beauty of DEVI's strategy is that she makes the Causalist do all this work.

Unlike in the classic Newcomb problem, however, there is simply no analogue of the "Why ain'cha rich?" challenge for the Causalist to answer. It is not that folk who give up smoking "to lower their probability of getting cancer" do better than Causalists in having lower rates of cancer, though worse than they themselves might have done if they had not needlessly deprived themselves of the pleasure of smoking. They simply do *worse* than Causalists, full stop. They have the same rate of cancer, and less pleasure from smoking.

In a case of this kind, then, the Causalist is able to argue that there is no gap between the two horns of Dummett's Dilemma. Even a subjectivist will have to accept that choosing not to smoke does not have a causal influence on rates of cancer. To argue this way, the Causalist needs to be able to establish that one-boxing (i.e., in this case, refraining from smoking) is an ineffective strategy, *without appealing to* (what she takes to be) *the known facts about the causal structure of the situation.* Why? Because otherwise it will remain an option for Devi to argue that the Causalist is simply wrong about the causal facts, as below.[13]

In cases in which Causalists are able to respond in this way to "Why aint'cha rich?" – i.e., in which they can reliably predict that Devi will not get rich, *without relying on disputed causal claims* – Devi helps herself to the explanation, modifying her Evidentialism if necessary. In this way, Devi allows Causalists to do the kind of work that has been done by her own less lazy Evidentialist cousins, of explaining why EDT doesn't in fact recommend one-boxing, in some (claimed) Newcomb problems, such as Smoking Gene.

A second class of cases are those Newcomb problems in which the Causalist needs another response to "Why aint'cha rich?" – cases in which both sides agree that the Evidentialist does better in the long run, but in which Causalist maintains that the Evidentialist *would have* done even better, had she two-boxed. In this case Devi plays Dummett's card.[14] She says: Here is an alternative view of the causal and counterfactual facts, according to which one-boxing is rational, by your Causalist lights. (This is the same as Dummett saying that saying "Click!" *causes* rather than merely *quasi-causes* the fact that the envelope does not contain a bill.)

The Causalists now faces Dummett's Dilemma. If they opt for objectivism, relying on some criterion other than the rationality of two-boxing as the basis for their view of the causal facts, they meet the analogue of Dummett's question about the criterion that causation only works 'forwards': Why should rational action be so constrained?

---

13. More generally, when Cartwright argues for claims of this form:

> [C]ausal laws cannot be done away with, for they are needed to ground the distinction between effective strategies and ineffective ones. ... [T]he difference between the two depends on the causal laws of our universe, and on nothing weaker. (1979, 420)

she requires that the fact that a strategy is effective or ineffective be *prior* to the causal facts that she takes to explain why it is one rather than the other. Otherwise she would have no distinction between *explanans* and *explandandum.*

14. Or, to put it more accurately, a card we found in Dummett's hand. We saw that Dummett himself does not play the hand to best advantage.

Whereas if they opt instead for subjectivism, then they are unable to dismiss Devi's alternative construal of the causal facts, while the issue of the rationality of one-boxing remains in dispute.

Devi always has the upper hand here, because the riches are on her side. The standard Causalist response to "Why aint'cha rich?" won't work, so long as the counterfactuals are just as much in play as the facts about causation. "On my understanding of causation, and hence counterfactuals," Devi insists, "It simply isn't true that had I two-boxed I would have been richer – on the contrary, I would have been poorer."

It is no use Causalists trying to stipulate their way out of the problem, by saying in their description of the Newcomb problem, "We assume that causation works in the ordinary way." For Devi will respond: "You're not entitled to assume that causation works in the normal way, in proposing examples that put this kind of pressure on the ordinary notion of causation. My point is that so long as you concede that the one-boxer will get rich, you've conceded that the case is in that contentious territory. At that point, my move is on the table, and you can't avoid it."[15]

True, there may well be cases in which it is simply unclear who gets rich. Perhaps there are limits on the long run, or confounding correlations to take into account, or perhaps the example is simply under-described or incoherent in some crucial respect. No matter, from Devi's point of view. So long as the cases remain unclear or incoherent they don't constitute a counterexample to the general method; and whichever way they fall, Devi's method will accommodate them.

4. The analogy with chance

We began with the *Euthyphro* question about causation. Is it the causal connection between *A* and *B* that makes it rational to do *A* to achieve *B?* Or does the rationality of the latter somehow constitute or ground the fact that *A* causes *B?* There is analogous question about chance. Is it the fact that there is a high chance that *P* that makes it rational to hold a high credence that *P?* Or is the fact that there is a high chance that *P* somehow *constituted by* or *grounded in* the fact that it is rational to be confident that *P?*[16] Again, we might call these options *objectivism* and *subjectivism* about chance, respectively.

In the case of chance, one writer sensitive to these issues is D H Mellor. Mellor (1971) defends a version of what we are here calling the subjectivist option, calling it 'personalism'. Following Kneale (1949), he insists that personalism is compatible with

---

15. The Causalist can certainly stipulate that the case involves no unusual causation *in her preferred objectivist sense.* But she thereby places herself on the objectivist horn of DD, facing the challenge of explaining why causation in *that objectivist sense* should be a good guide to rational action, in an unusual case of the kind presented (a case in which both sides agree that it is the Evidentialist who gets richer).

16. It may seem that in the case of causation we are asking about a connection with rational *action,* whereas in the case of chance we are asking about a connection with rational partial *belief.* As we shall see in a moment, however, this difference turns out to be superficial.

the view that chances are real and objective – it is just that in saying what they are, we need to begin with rational degrees of belief, or credences.

> [C]an we not analyse full belief that the chance of heads on a coin toss is 1/2 without reference to some supposedly corresponding partial belief that the coin will land heads? The reason for denying this is the fact to which Kneale himself draws attention (p. 18) "that knowledge of probability relations is important chiefly for its bearing on action". It follows as Kneale says (p. 20) that "no analysis of the probability relation can be accepted as adequate ... unless it enables us to understand why it is rational to take as a basis for action a proposition which stands in that relation to the evidence at our disposal". Similarly with chance. It must follow from our account that the greater the known chance of an event the more reasonable it is to act as if it will occur. This concept of a quantitative tendency to action is just that of partial belief [i.e., credence] as it has been developed by the personalists. It is thus available to provide in our account of chance that necessary connection with action on which Kneale rightly insists. *A great difficulty facing other objective accounts of chance, notably the frequency theories, has been to build such a connection subsequently on their entirely impersonal foundations.* (Mellor, 1971, 3, emphasis added.)

Lewis (1980) also defends a form of subjectivism. Like Mellor, he takes chance to be real and objective, but takes it to be definitive of chance that it plays a distinctive role in guiding credence. As he says, he is "led to wonder whether anyone *but* a subjectivist is in a position to understand objective chance!" (1986 [1980], 84). Returning to this theme in later work, he criticizes rival approaches on the grounds that they pay insufficient attention to this connection between chance and credence: "Don't call any alleged feature of reality 'chance' unless you've already shown that you have something, knowledge of which could constrain rational credence," he says.[17]

Lewis gives a name to the principle that knowledge of chance properly constrains rational credence. In view of its centrality to an understanding of chance – "this

---

17. It is worth quoting the full passage in which Lewis makes this remark:

> Be my guest—posit all the primitive unHumean whatnots you like. ... But play fair in naming your whatnots. Don't call any alleged feature of reality "chance" unless you've already shown that you have something, knowledge of which could constrain rational credence. I think I see, dimly but well enough, how knowledge of frequencies and symmetries and best systems could constrain rational credence. I don't begin to see, for instance, how knowledge that two universals stand in a certain special relation N* could constrain rational credence about the future coinstantiation of those universals. (Lewis, 1994, 484)

Lewis's point here is closely analogous to the objectivist horn of DD. A Causalist who rejects subjectivism needs some other criterion for causal connectedness. Following Dummett's lead in the case in which the criterion invokes temporal direction, DD then asks the Causalist to explain why *that* criterion should constrain rational action. This is the same challenge that Lewis here offers to his own objectivist opponents.

principle seems to me to capture all we know about chance", as he puts it – he calls it the Principal Principle (PP). His original formulation of PP is as follows:

> Let *C* be any reasonable initial credence function. Let *t* be any time. Let *x* be any real number in the unit interval. Let *X* be the proposition that the chance, at time *t*, of *A*'s holding equals *x*. Let *E* be any proposition compatible with *X* that is admissible at time *t*. Then $C(A|XE) = x$. (1986 [1980], 87)

(The notion that Lewis here calls the 'admissibility' of the evidential proposition *E* will play a crucial role in what follows, and we'll return to it shortly.)

PP may seem at first sight to be a psychological principle, but this is misleading. Credence is here taken to be *defined* in behavioural terms – "a quantitative tendency to action", as Mellor puts it.[18] So PP is actually a principle describing the rational bearing *on action* of beliefs about objective chance.

With this reminder in place, it is possible to see how, as Price (2012) proposes, there is an analogy between PP and CDT. PP tells us how a rational agent will act in a certain range of decision problems, in the light of her beliefs about objective chance. CDT tells us how a rational agent will act in a different range of decision problems,[19] in the light of her beliefs about objective causal dependency. In both cases we have a principle that connects a belief about what we might call a modal fact, on the one hand, to rational action, on the other.

The analogy between PP and CDT has been obscured by two factors, in our view. Both are products of the particular history of the relevant part of decision theory, and especially of the long rivalry between CDT and EDT. The first factor is that there is no analogue, agreed on all sides, of the behaviourally-defined notion of credence. In conditional decision theory, where an agent needs to consider the possibility that States might depend on Acts, we are usually offered instead two *different* accounts of the relevant notion of conditional dependence: an evidential one associated with EDT and a causal one associated with CDT. Comparing this to the case of chance, it is as though the behaviourally-defined notion of credence had dropped out of unconditional decision theory, and we were being offered two rival accounts of how beliefs about something else constrained action: PP, in which the 'something else' is objective chance, and some rival principle, featuring some more evidential rival to chance.

This first factor is easily remedied. We simply need to introduce into our notation a notion of conditional dependence that is understood in the required behavioural, descriptive sense – explicitly neutral between CDT and EDT, who can now disagree about how these behaviourally-defined conditional credences (BCCs) are rationally constrained by beliefs about something else.

---

18. Similarly, Lewis stresses that credences simply *are* behavioural dispositions: "No wonder your credence function tends to guide your life. If its doing so did not accord to some considerable extent with your dispositions to act, then it would not be your credence function. You would have some other credence function, or none." (1980, 288 in original version)

19. In fact a *broader* range.

In the unconditional case Savage's Decision Theory (SDT) does the job of formalising a definition of credence in terms of choice behaviour. Accordingly, we can then think of PP as a sum of two components: the first a strictly psychological principle (call this PP$_\psi$) relating beliefs about chances to credences; and the second SDT, connecting credence to choice behaviour. We thus have

$$PP = PP_\psi + SDT.$$

To make the analogous move in the conditional case, we need to think of CDT as a sum of two components. The first (call it CDT$_\psi$) describes how a rational agent aligns her BCCs with her beliefs about relevant causal dependencies. The second is a decision theory that does the job done by SDT in the unconditional case, of representing BCCs in term of decision behaviour. Fishburn (1964) developed a conditional version of SDT which is a generalization of Savage's system, and this appears to be what we need. Calling it FDT, this gives us the following analogue of the decomposition of PP:

$$CDT = CDT_\psi + FDT.$$

We said earlier that there are two factors that seem to have obscured the analogy between CDT and PP. The second factor is the comparative rarity, within the philosophy of causation, of the kind of subjectivism represented in the case of chance by major figures such as Mellor and Lewis; perhaps combined with the view, from the CDT side in the CDT/EDT debate, and relatedly from Cartwright, that objectivism about causation does useful work, in distinguishing effective from ineffective strategies. This factor (or combination of factors) is harder to remedy, but we have offered DD and the analogy with chance by way of therapy. Combined with the salutary analogue of Mellor's and Lewis's insistence that unless an account of chance puts the link with action in at the beginning, it will have trouble getting it out at the end, DD leaves the objectivist Causalist in an unenviable position: unable to explain why a decision maker should *care* about relations of causal dependence, as the Causalist conceives them (at least in the disputed cases in question).

This therapy is combined with the offer of what we hope will prove a palatable alternative for the Causalist, at least with the analogy with chance in mind. Provided we approach causation in a subjectivist spirit, we can enshrine CDT as the, or at least *a,* central commitment of a theory of causation, as Lewis does for PP in the case of chance. Causal relations become objectified relations of conditional credal dependence, the latter understood from the agent's point of view.

Note that this allows CDT a prescriptive and explanatory role, in normal circumstances. In such circumstances we can read off our rational BCCs from our knowledge of causal relations, just as PP enables us to read off our rational (unconditional) credences from knowledge of chances.

What about abnormal circumstances? Here, too, the analogy with chance is helpful. Lewis takes PP to admit exceptions: unusual kinds of evidence can give an agent reason to hold credences that do not align with her beliefs about chance. Lewis calls

this 'inadmissible evidence'[20] – the evidence that a crystal ball would provide, say, if it showed us the result of a coin toss in advance. The coin might be known to be fair, but we would be foolish to hold a credence of 0.5 in Heads and bet accordingly if the crystal ball reveals that it will in fact land Tails. Lewis thus allows exceptions to PP. But as Hall (1994, 2004) points out, there is another alternative. For Hall, cases of (so-called) inadmissible evidence are just cases in which the chances are strange. PP holds with full generality.

Hall and Lewis agree about the rationality of taking the crystal ball into account. They both disagree with a hardliner who neither modifies her view of chance, nor admits exceptions to PP, in the light of inadmissible evidence. "To be sure", this hardliner says, "My crystal ball shows me that this fair coin will land Tails on the next toss, and I have no reason to doubt it. But it is a fair coin, so the rational credence in Heads is 0.5. I'll accept an evens bet on Heads."[21]

Note that in disagreeing with the hardliner, Lewis and Hall cannot be relying on PP, for in the cases in question they disagree both about PP's applicability and about the appropriate theory of chance. They therefore share some conception of practical rationality that is for both of them 'prior' to the theory of chance. Again, this is the analogue of the point on which Devi's strategy relies, that even an objectivist needs a non-causal criterion for the distinction between effective and ineffective strategies.

For a subjectivist about causation the same options exist with respect to CDT in Newcomb cases. One option (Dummett's choice in 1954) is to allow that, like PP for Lewis, CDT is not exceptionless. In strange cases such as "Click!" and the classic Newcomb puzzle, it may be rational to ignore CDT. The other option (Dummett's choice in 1964) is to say as Hall says for chance that causation may be weird in these strange cases, so that CDT recommends one-boxing (and paying tuppence to say "Click!"). Note that just as Lewis and Hall agree that it is rational to let crystal balls influence one's credences, these two options agree that one-boxing is the rational choice in the classic Newcomb problem. The disagreement is entirely about a terminological matter, the very matter on which Dummett changed his mind between 1954 and 1964.

The peculiarity of the causal case is that the hardline position – unoccupied, so far as we know, in the case of chance – remains the orthodoxy for causation. Such is the grip of objectivism in this case. Against these hardliners we have offered the pragmatic subjectivist wisdom of Dummett in the causal case, and of Mellor, Lewis and others in what we have urged is a closely analogous case, that of chance. We conclude that objectivist Causalism is philosophically as well as financially impoverished, and that subjectivism takes home the bacon, in both respects.[22]

---

20. An infelicitous name, perhaps, since Lewis is not denying that we should admit it and use it. It is inadmissible only as a constraint on chance, in Lewis's view.

21. Confronted with the "Why ain'cha rich?" objection, the hardliner makes the same response as the Two-Boxer in the classic Newcomb problem: "The reason why [I am] not rich is that the riches were reserved for the irrational" (Lewis, 1981, 377).

References

Cartwright, N. (1979). Causal laws and effective strategies. *Noûs*, 13(4):419–437.

Dummett, M. A. E. (1954). Can an effect precede its cause? *Proceedings of the Aristotelian Society, Supplementary Volume*, 28:27–44.

Dummett, M. A. E. (1964). Bringing about the past. *The Philosophical Review*, 73(3):338–359.

Dummett, M. A. E. (1987). Reply to D H Mellor. In Taylor, B., editor, *Michael Dummett: Contributions to Philosophy*, pages 287–298. Kluwer Academic Publishers.

Dummett, M. A. E. (1993 [1986]). Causal loops. In Flood, R. and Lockwood, M., editors, *The Nature of Time*, pages 135–69. Oxford: Blackwell. Reprinted in Dummett, M. A. E., *The Seas of Language,* Oxford: Oxford University Press (1993), 349–375. (Page references are to the latter version.).

Fishburn, P. C. (1964). *Decision and Value Theory*. John Wiley and Sons, New York.

Hall, N. (1994). Correcting the guide to objective chance. *Mind*, 103(412):505–518.

Hall, N. (2004). Two mistakes about credence and chance. *Australasian Journal of Philosophy*, 82(1):93–111.

Horgan, T. (1981). Counterfactuals and Newcomb's problem. *The Journal of Philosophy*, 78(6):331–356.

Kneale, W. C. (1949). *Probability and Induction*. Clarendon Press.

Lewis, D. (1981). 'Why ain'cha rich?'. *Noûs*, 15(3):377–380.

Lewis, D. (1986 [1980]). A subjectivist's guide to objective chance. In *Philosophical Papers*, volume II, pages 83–132. Oxford University Press (originally published in *Studies in Inductive Logic and Probability*, Vol. II, ed. Richard C. Jeffrey, Berkeley: University of California Press 1980).

Lewis, D. (1994). Humean supervenience debugged. *Mind*, 103(412):473–490.

Mackie, J. L. (1977). Newcomb's paradox and the direction of causation. *Canadian Journal of Philosophy*, 7(2):213–225.

Mellor, D. H. (1971). *The Matter of Chance*. Cambridge University Press.

Nozick, R. (1969). Newcomb's problem and two principles of choice. In Rescher, N., editor, *Essays in honor of Carl G. Hempel*, pages 114–146. Springer.

Price, H. (2001). Causation in the special sciences: the case for pragmatism. In Costantini, D., Galavotti, M. C., and Suppes, P., editors, *Stochastic Causality*, pages 103–120. CSLI Publications.

Price, H. (2007). Causal perspectivalism. In Price, H. and Corry, R., editors, *Causation, Physics and the Constitution of Reality: Russell's Republic Revisited*, pages 250–292. Oxford University Press.

Price, H. (2012). Causation, chance, and the rational significance of supernatural evidence. *Philosophical Review*, 121(4):483–538.