

Decision-Theoretic Pluralism

Causation, Evidence, and Indeterminacy*

By Adam Bales

A prominent philosophical debate concerns whether we should accept causal decision theory (CDT) or evidential decision theory (EDT) as our best theory of rational choice. However, instead of accepting one of these theories at the expense of the other, an alternative would be to accept that both theories play a partial role in the true account of rational choice. In this paper, I defend a pluralist account of this sort. In particular, I argue that rational permissibility is an indeterminate notion, with EDT and CDT each corresponding to one sharpening of this notion.

Keywords Causal Decision Theory, Evidential Decision Theory, Indeterminacy, Newcomb's problem.

We're doing rational choice theory wrong. This is the claim that I will defend in this paper. We're doing rational choice theory wrong and if we do it right then we can irenically resolve some long-standing disagreements.

*Thanks to Arif Ahmed, Matteo Bianchetti, Kenny Easwaran, Richard Holton, Anna Mahtani, Jack Spencer, Reuben Stern, Shyane Siriwardena, Robbie Williams, audiences at the Moral Sciences Club and several anonymous reviewers. This is a draft of an article, the final version of which appears at <https://academic.oup.com/pq/advance-article/doi/10.1093/pq/pqy021/5039118>.

My interest in this paper will be in the decision-theoretic debate that was kickstarted when Nozick (1969) introduced *Newcomb's Problem*: the debate pitting causal decision theory (CDT) against evidential decision theory (EDT).¹ Nozick himself came to think that both the causal and evidential perspectives captured elements of the truth about rationality. On the other hand, philosophical orthodoxy holds that at most one of these perspectives is right (and much discussion has concerned which should be abandoned and which retained).

In the first half of this paper, I argue that the existing literature already gives us strong grounds to reject philosophical orthodoxy and accept Nozick's claim that both perspectives capture part of the truth about rationality. In the second half of the paper, I then make this claim precise. In particular, I argue that it is sometimes an indeterminate matter whether a decision is rationally permissible, with the evidential and the causal perspectives corresponding to different sharpenings of this indeterminate concept. This result ironically resolves the decision-theoretic debate.

1 Two perspectives on rationality

The jumping-off point for this paper is *Newcomb's Problem*:²

An agent faces two boxes, one transparent (which contains \$1000) and one opaque (which is either empty or contains \$1,000,000). The agent may either take both boxes (two-box) or just the opaque box (one-box). However, the opaque box was filled yesterday based on a perfect prediction of the agent's behaviour,

¹CDT and EDT aren't the only two positions in the debate but I focus on these theories here.

²In all cases discussed in this paper, it should be assumed that the agent *believes* that the case has the features described.

such that \$1,000,000 was placed in the opaque box if the agent was predicted to one-box (otherwise the box was left empty).

Newcomb's Problem highlights the distinction between two perspectives on rational choice: a causal perspective and an evidential perspective.

To get to the causal perspective, consider an argument for two-boxing: (1) the agent can't causally influence the contents of the opaque box, as this box is already filled and the case is stipulated to involve no backwards causation; (2) however this box is filled, the agent ends up \$1000 richer if she two-boxes than if she one-boxes (as she will inevitably get the opaque box's contents but will also get \$1000 from the transparent box if she two-boxes); and therefore (3) the agent should two-box.

This argument suggests a causal perspective on rationality, on which what an agent ought to do depends centrally on her beliefs about what she can and cannot causally influence (after all, this argument appeals to the fact that the agent can't causally influence the opaque box's contents). This causal perspective is normally spelled out in the form of causal decision theory (CDT), which in informal terms holds that a decision is rationally permissible if its expected causal effects are at least as good as the expected causal effects of any other decision. CDT endorses two-boxing in *Newcomb's Problem*, as this causes the agent to gain the contents of both boxes, rather than causing her to gain the contents of just the opaque box.

These same points can be made more formally. Now, in formal terms, CDT entails that a decision is rationally permissible if it maximises causal expected utility (CEU), where the CEU of a decision, D , is defined as follows:

$$CEU(D) = \sum_S Cr(S \setminus D)U(S \wedge D)$$

Here, S ranges across a set of world states (that is, possibly ways that the world might be). U is then a utility function, which assigns a real-number utility value to an outcome (that is, the conjunction of a state and a decision). This number represents how desirable the agent finds each outcome, such that a higher number represents a more desirable outcome. Finally, $Cr(S \setminus D)$ is a credence function that assigns credences (that is, subjective probabilities) to a state S after accounting for the causal influence of D . There is some disagreement about how precisely this should be spelled out.³ However, the basic idea is that this credence will be high if the agent either thinks that D causes S to be likely to hold or thinks that S is likely to hold regardless of the agent's decision.⁴ So CDT labels a decision as permissible if it maximises a causal-credence-weighted sum of utilities.

Now we can apply this theory to *Newcomb's Problem*. First, let E represent the opaque box being empty and M represent the opaque box containing \$1,000,000. Then let T represent the agent two-boxing and O represent her one-boxing. Now, because the agent's behaviour does not causally influence the box contents $Cr(E \setminus T) = Cr(E \setminus O) = Cr(E)$ and $Cr(M \setminus T) = Cr(M \setminus O) = Cr(M)$. Assuming for simplicity that the utility of outcomes is equal to the monetary value of these outcomes, we can now calculate the CEU of each decision as follows:

$$\begin{aligned} CEU(T) &= Cr(E \setminus T)U(E \wedge T) + Cr(M \setminus T)U(M \wedge T) \\ &= Cr(E) * 1000 + Cr(M) * 1,001,000 \end{aligned}$$

³Indeed, CDT is often formalised in *counterfactual* (rather than causal) terms but this distinction is unimportant for my purposes.

⁴For more details about this credence, see Joyce 1999: 161–80.

$$\begin{aligned}
CEU(O) &= Cr(E \setminus O)U(E \wedge O) + Cr(M \setminus O)U(M \wedge O) \\
&= Cr(E) * 0 + Cr(M) * 1,000,000 \\
&= CEU(T) - 1000
\end{aligned}$$

As the CEU of two-boxing (T) is higher than the CEU of one-boxing (O), it follows that CDT will endorse the permissibility of just two-boxing. So CDT, the causal perspective on rationality, endorses two-boxing.

Now we can turn to the evidential perspective on rationality. To reach this, consider an argument for one-boxing: (1) if the agent two-boxes then she will have been predicted to two-box and so the opaque box will be empty (and so the agent will end up with just the \$1000 from the transparent box); (2) if the agent one-boxes then she will have been predicted to one-box and so the opaque box will contain \$1,000,000 (and so the agent will end up with \$1,000,000); (3) \$1,000,000 is more than \$1000; and therefore (4) the agent should one-box.

This argument can lead to an evidential perspective on rationality, which is typically spelled out in the form of evidential decision theory (EDT). In informal terms this theory holds that a decision is rationally permissible if it provides the agent with evidence that is at least as desirable as the evidence that any other decision would provide. EDT sides with the above argument and endorses one-boxing in *Newcomb's Problem*, as one-boxing provides the desirable evidence that the agent was predicted to one-box and hence that the opaque box contains \$1,000,000, whereas two-boxing provides the undesirable evidence that this box is empty.

Again, this point can be made formally. In particular, EDT labels a decision as permissible if it maximises evidential expected utility (EEU). The EEU of a decision is calculated in the same way as its CEU, except

that $Cr(S \setminus D)$ is replaced with the conditional credence $Cr(S | D)$, which we can think of as capturing the evidence that D provides about S. So we calculate the EEU of a decision, D, as follows:

$$EEU(D) = \sum_S Cr(S | D)U(S \wedge D)$$

So EDT takes a decision to be permissible if it maximises a conditional-credence-weighted sum of utilities.

Now, because of the predictor's perfect accuracy, $Cr(M | O) = Cr(E | T) = 1$ and $Cr(E | O) = Cr(M | T) = 0$. We can now calculate the EEU of the decisions in *Newcomb's Problem* as follows:

$$\begin{aligned} EEU(T) &= Cr(E | T)U(E \wedge T) + Cr(M | T)U(M \wedge T) \\ &= 1000 \end{aligned}$$

$$\begin{aligned} EEU(O) &= Cr(E | O)U(E \wedge O) + Cr(M | O)U(M \wedge O) \\ &= 1,000,000 \end{aligned}$$

As the EEU of one-boxing exceeds the EEU of two-boxing, EDT will endorse one-boxing. So according to EDT, the evidential perspective, one ought one-box in *Newcomb's Problem*. We have two perspectives on rationality and these perspectives part way on *Newcomb's Problem*.

2 Perspectival Pluralism

In the decades since Nozick introduced *Newcomb's Problem* to the philosophical community, there has been an ongoing debate about which of these

perspectives is correct. However, Nozick (1993: 45) has spoken out against the assumption, tacit in this debate, that at least one of these perspectives must be rejected outright:

I suggest that we... say not merely that we are uncertain about which *one* of these two [perspectives] is (all by itself) correct, but that both of these [perspectives] are legitimate and each must be given its respective due.

Call this broad approach *pluralism*. That is, according to pluralism there are either: (a) unified norms of rational permissibility *simpliciter*, such that these norms refer to considerations from both the causal and evidential perspectives; or (b) distinct norms of causal rationality and evidential rationality (which are captured by the causal and evidential perspectives in turn).⁵ An example of the first type of pluralism would be Nozick's own view, according to which a decision is permissible if it maximises a weighted sum of the CEU and the EEU. There is a clear sense in which this view takes permissibility to depend both on considerations captured by the causal perspective (the CEU) and considerations captured by the evidential perspective (the EEU). An example of the second type of pluralism would then be the view discussed in Horgan 1985: Horgan suggests that perhaps causal and evidential rationality are simply distinct normative domains and are not competing attempts to capture truths about the same normative domain.

Setting aside, for now, the distinction between these two types of accounts, one motivation for pluralism follows from the fact that both the argument for two-boxing and that for one-boxing are compelling; this is what

⁵In this paper, I talk about rational *permissibility* rather than *requirement*. In doing so, I follow the treatment of indeterminacy and choice in various other discussions (cf. Dougherty 2014 and Rinard 2015). In any case, given that an agent is required to ϕ just if she is both permitted to ϕ and not permitted to $\neg\phi$, the discussion could easily be adapted so as to comment on rational requirement.

makes *Newcomb's Problem* puzzling. One possible response to this fact is to accept pluralism: that is, accept that the two perspectives that follow from these arguments both play a role in an adequate account of rationality.⁶

Nor is *Newcomb's Problem* the only case that supports pluralism. For example, consider *the Psychopath Button* (from Egan 2007):

Paul is faced with a button that, if pressed, will kill all of the psychopaths in the world. Paul must now choose between two options: he can either *press* the button or he can *refrain* from doing so. Now Paul strongly desires to live but he has a weak desire to rid the world of psychopaths. Further, while Paul is almost certain that he isn't a psychopath, he thinks that only a psychopath would be likely to press the button. Should Paul press the button?

As with *Newcomb's Problem*, there are two arguments here, to contradictory conclusions.

Egan points to one of these. In particular, Paul thinks that only a psychopath is likely to press the button and so thinks that if he himself presses the button, he is almost certainly a psychopath. If Paul is a psychopath and presses the button then he will cause his own death. Consequently, Paul pressing the button provides evidence that he will almost certainly cause his own death. Under such circumstances, Paul should not press the button when he has the risk-free option of refraining from pressing.

The evidential perspective is responsive to this argument. After all, if Paul presses then this provides evidence that he's a psychopath and so evi-

⁶Both sides of the debate have responses to the argument that they reject (cf. Joyce 1999: 151–4; Ahmed 2014: 195–9). However, insofar as the arguments are intuitively plausible, one might reject the responses rather than the original argument. It is surprising that more people haven't chosen tollens over ponens here.

dence that he will die. On the other hand, if Paul refrains from pressing, this provides evidence that no-one will die and so that Paul himself will survive. As the second lot of evidence is more welcome for the agent than the first, the evidential perspective will endorse refraining from pressing.⁷

Now, while many people find the argument for refraining compelling, some do not. Further, there is in fact another argument available here. After all, Paul is almost certain that he isn't a psychopath. If Paul isn't a psychopath then he receives his most desired outcome by pressing the button and receives a less desired outcome if he refrains from pressing the button. So Paul is rationally required to press the button.⁸

The causal perspective is responsive to this argument. After all, Paul is almost certain that he's not a psychopath and so almost certain that pressing will cause all psychopaths to die but will not cause his own death. On the other hand, refraining will not cause any deaths. Insofar as Paul wants to kill all psychopaths, the expected causal impact of pressing is then better than the expected causal impact of not pressing. Consequently, the causal perspective will label pressing as rationally required.⁹

So, as in *Newcomb's Problem*, in *the Psychopath Button* there are compelling arguments to contradictory conclusions. Then, once again, a natural response is to accept that both the causal and the evidential perspectives play a role in an adequate account of rational choice. After all, the evidential perspective is responsive to one of the plausible arguments in *the Psychopath Button* and the causal perspective is responsive to the other. Given this, we can make sense of both arguments by accepting that both perspectives play

⁷More formally, this result follows from EDT because $Cr(\text{Psychopath} \mid \text{Press})$ is high.

⁸It is stipulated in *the Psychopath Button* that an appropriately small risk of death can be outweighed by the higher chance of successfully killing all psychopaths. More formally, this result follows from CDT because $Cr(\text{Psychopath} \setminus \text{Press}) = Cr(\text{Psychopath})$ and, as specified, this credence is low.

⁹For a more formal discussion, see Egan 2007: 99–100.

a role in our best account of rational choice and so can make sense of both of these arguments by accepting pluralism (further, I will argue in §5 that pluralism can also account for why the first of the above arguments might strike us as more compelling than the second). So, along with *Newcomb's Problem*, *the Psychopath Button* supports pluralism. Consequently, we have two initial reasons to accept pluralism. In the remainder of the paper, I will discuss further reasons to do so, in the course of discussing a particular form of pluralism: indeterminacy pluralism.

3 Indeterminate Rationality

Pluralism can come in many flavours. This raises the question of which form of pluralism we should accept. Well, here it's natural to start with Nozick, my paradigm pluralist, whose account I outlined earlier. According to this account, a decision is permissible if it maximises a weighted sum of the CEU and EEU. Now the full details of this account will not matter for my purposes. Here's what does matter: Nozick accepts a view on which there are unified, overarching norms of rationality. Nozick's view is then pluralist insofar as he holds that these unified norms make reference to considerations from both the causal and the evidential perspectives. This is the first of the two types of pluralism that I mentioned earlier.

However, assuming that we are motivated to accept pluralism in part by reflection on *Newcomb's Problem* and *the Psychopath Button*, this sort of hybrid view is problematic. After all, we can now ask what the overarching norms of rational permissibility *simpliciter* entail about rational permissibility in any particular version of these cases. Take *Newcomb's Problem*. Here there are three options. First, these norms might entail the permissibility of just two-boxing. However, the view then fails to be appropriately re-

sponsive to the earlier argument for one-boxing. Second, these norms might entail the permissibility of just one-boxing. However, the view then fails to be appropriately responsive to the earlier argument for two-boxing. Third, the norms might entail the permissibility of both decisions. However, the view then fails to be appropriately responsive to either the argument for one-boxing or the argument for two-boxing. After all, these arguments establish not just that one decision is rational but also that the other is *not*. Given this, taking both decisions to be permissible is in conflict with both arguments. Consequently, any account that spells out pluralism in terms of unified, overarching norms of rationality will struggle to account for both arguments in *Newcomb's Problem*. Assuming that accounting for these is one of our motivations for accepting pluralism, we have grounds to reject such a version of pluralism.¹⁰

It is time to turn, then, to an alternative type of pluralism. To get to such a view, I start with a comment made by Horgan (1985: 229):

[The stalemate over *Newcomb's Problem*] is frustrating...because it suggests the disturbing possibility that there is really no such thing as the rational act *simpliciter* but only the [evidentially-rational] act and the [causally-rational] act.

In other words, Horgan is a pluralist not in the sense that he takes the causal and evidential perspectives to both play a role in norms of rational permissibility *simpliciter* but in the sense that he accepts the existence of distinct norms of causal rationality and norms of evidential rationality.¹¹

¹⁰One rejoinder: perhaps one-boxing is required in just some versions of *Newcomb's Problem* and two-boxing is required in others. Perhaps the force of the above arguments then results because they imperfectly mimic more nuanced, true arguments. Still, absent a substantial argument for this claim, it is reasonable to set this form of pluralism aside.

¹¹Outright rejection of the existence of rational permissibility *simpliciter* is the most extreme form of this view. A more moderate view would accept that there are norms of

This was the second type of pluralism that I mentioned above.

This raises a further question: if there are merely norms of causal rationality and of evidential rationality then what should we say about *talk* of rational permissibility simpliciter? Here's an attractive view: our concept of rational permissibility admits of indeterminacy, with the evidential perspective (EDT) corresponding to one sharpening of this concept and the causal perspective (CDT) corresponding to another.¹² Now, the full details of this view will depend on the account of indeterminacy adopted. However, for concreteness I will focus here on the view that we are led to if we adopt supervaluationism regarding indeterminacy.¹³ Consequently, on the view under consideration: (a) a decision is rationally permissible simpliciter if and only if it is rationally permissible from both the causal and evidential perspectives (that is, according to both CDT and EDT); (b) a decision is rationally impermissible simpliciter if and only if it is rationally impermissible on both the causal and evidential perspectives (again, according to both EDT and CDT); and (c) a decision is indeterminately permissible if and only if it is permissible on one perspective and impermissible on the other (so if it is permissible according to one of EDT and CDT and impermissible according to the other).¹⁴ In the remainder of this paper, I will argue for

rational permissibility *simpliciter* but that they fall silent in a small number of cases. In these cases, the overall notion will not apply and instead we can apply only the notions of the evidentially and the causally rational.

¹²Here, I assume that we do have a single concept of (instrumental) rational permissibility. This claim strikes me as intuitively plausible. Further, proponents of EDT and CDT do typically take it that they are attempting to provide an account of the same sort of means-end rationality. Then, on top of all of this, EDT and CDT come apart only in strange cases that were not encountered in the development of our concepts. As such, it is unlikely that we will have distinct concepts to pick out these two classes of norms (given that they overlap in the cases that shaped our conceptual world).

¹³I will assume that indeterminacy occurs at the semantic level rather than in the world itself. Alternatives to supervaluationism include epistemicism and subvaluationism and which account one adopts here will likely depend on their broader views about indeterminacy. I find supervaluationism generally plausible and so focus on this account here.

¹⁴Might there be other sharpenings of permissibility? I return to this question later. However, I note that if indeterminacy pluralism is to be plausible then any other sharpen-

this indeterminacy pluralism.¹⁵

4 Four Arguments

In order to do so, I will provide four arguments for this view.

4.1 Indeterminacy Worries

As a starting point, then, I note that some people might be suspicious from the get go about the idea of indeterminate permissibility.¹⁶ Consequently, it's worth starting out by deflating this suspicion.

In order to do so, I note that there are various reasons, independent of the considerations in this paper, to take the possibility of indeterminate permissibility seriously. For example, Williams (2014; 2016) has argued that there are some cases where it is indeterminate whether we will receive the things that we care about (perhaps because it is indeterminate whether we will be the person that receives these things) and has suggested that indeterminate permissibility might arise in such cases. Similarly, Rinard (2015) has suggested that cases involving imprecise credences are cases where a person has indeterminate beliefs and has suggested that indeterminate permissibil-

ings will need to coincide with EDT and CDT in everyday cases, where it is implausible to declare that some decision is indeterminately permissible. Why think that this will turn out to be the case? The basic idea, which I will discuss further in §4.2, is that reflection on the factors that shape our concepts suggests that indeterminacy will only arise in strange cases here.

¹⁵Alternatively, perhaps 'rational permissibility' is ambiguous. Despite some apparent differences, I think there is little to separate the ambiguity and indeterminacy views, so discuss indeterminacy for concreteness but remain open to both views.

¹⁶It might be worried that an evaluation of a decision as indeterminately permissible cannot guide us. However, it's not clear that this is true (cf. Williams 2014). Further, even if it is, the indeterminacy theory can guide us when decisions are determinately permissible and can play useful non-guidance roles (like providing insight into the normative world).

Another concern: does permissibility lack the features of paradigm examples of vagueness (like baldness and redness)? Perhaps but it is similar to another such case: the vagueness of counterfactual dependence. Further, indeterminacy extends beyond vagueness (see Williams 2012) and so indeterminacy pluralism need not be about vagueness at all.

ity naturally arises here. Finally, Schoenfield (2015) and Dougherty (2014), among others, have suggested that moral permissibility can be indeterminate. For example, Dougherty considers that it might be morally permissible to save your friend rather than, say, two strangers, morally impermissible to save your friend rather than, say, 1000 strangers and has suggested that there might be some number of strangers in between where it is indeterminately permissible to save your friend rather than the strangers. In the light of this, it is plausible that rational permissibility too might sometimes be indeterminate.

We already have good reasons, then, to think that rational permissibility can be indeterminate. Consequently, the mere fact that the theory under consideration involves indeterminate permissibility is no mark against it.¹⁷

4.2 Semantic Considerations

Following on from this, it's worth saying something to motivate the specific indeterminacy theory under discussion. Note, then, that indeterminacy pluralism denies that there are any norms of rational permissibility *simpliciter* (or, at least, denies that these norms must always give unequivocal guidance). What this theory explains, then, is our *talk* of such permissibility. Consequently, the sort of indeterminacy under discussion here is semantic indeterminacy, where such indeterminacy arises when the conventions that govern our language don't settle some case.

Now, as noted in Horwich 1985, these conventions evolve so as to be useful in our actual environment and so it's entirely plausible that these conventions will not settle strange cases like *Newcomb's Problem*. Consequently, Horwich notes, EDT and CDT (which agree in most everyday cases) will

¹⁷Nor does the literature on indeterminate permissibility undermine the novelty of the current discussion (which outlines a new source of indeterminacy of permissibility).

plausibly do an equally good job of capturing the settled conventions about rational permissibility.¹⁸ As such, it is natural to think that permissibility might sometimes be indeterminate and that the evidential and causal perspectives might correspond to different ways of resolving this indeterminacy. We have initial grounds, then, to be sympathetic to the indeterminacy theory.

4.3 The Indeterminacy of Counterfactuals

More speculatively, it can be argued that the indeterminacy theory follows from the indeterminacy of counterfactuals. In particular, which decisions are permissible plausibly depends on what would be the case if each decision were made (that is, permissibility is connected to counterfactual reasoning).¹⁹ However, counterfactual dependence is indeterminate (cf. Lewis 1979: 457). Now if permissibility depends on what counterfactuals hold and this is an indeterminate matter, then indeterminacy of permissibility naturally arises.

Taking this argument slowly, start with Lewis's (1973) account of counterfactuals. In simplified terms, this account holds that a counterfactual is true if its consequent holds in the possible world where the antecedent holds that is *most similar* to the actual world. Now similarity is indeterminate: one world might be more similar to the actual world in terms of its laws of nature while another might be more similar in terms of its past history and there doesn't seem to be a determinate answer as to which of these worlds is overall more similar to the actual world. Now, as similarity is indeterminate

¹⁸Horwich (1985) concurs but argues for extending our concept of permissibility in accordance with EDT. The indeterminacy theorist stops one step earlier: as we're already committed to indeterminate permissibility elsewhere, there is little cost to stopping here.

¹⁹This view does not unfairly favour CDT (which can be formalised in counterfactual terms) because EDT too can be construed in such a way (Horgan 1981: 343–7).

and counterfactuals hold based on similarity considerations, it is at least sometimes indeterminate whether a counterfactual holds.

Now, we can recognise the importance of counterfactuals to choice by framing decision theory not in terms of conditional or causal credences but in terms of counterfactual credences. So, on this view, decision theory appeals to the agent's credences that if she were to choose D, S would be the case. A decision is then permissible if it maximises counterfactual expected utility. Now this theory's guidance depends on how we resolved the indeterminacy of counterfactual dependence: different guidance will result from different resolutions. Consequently, if we don't insist on using a single resolution of this indeterminacy in decision theory, decisions will sometimes be indeterminately permissible.

But will CDT and EDT correspond to sharpenings of permissibility, as per indeterminacy pluralism? Well, they will if each results from counterfactual decision theory by resolving counterfactual indeterminacy in some manner. And indeed, CDT is often presented as counterfactual decision theory, with a non-backtracking resolution of counterfactual indeterminacy (roughly, a resolution that insists on sameness of past).²⁰ Further, per Horgan (1981: 345–7), EDT too can be construed in this manner, but with a conditionalised resolution of indeterminacy. So CDT and EDT will naturally correspond to sharpenings of the indeterminate notion of permissibility.

Still, this hardly provides a watertight argument for indeterminacy pluralism. After all, while counterfactuals plausibly play a role in rational choice, it could be insisted that we must resolve all counterfactual indeterminacy to make sense of this role. If so, indeterminate permissibility will not arise.²¹ So if it doesn't provide a watertight argument, what does the

²⁰I set aside here a complication involving chancy cases. See Lewis 1981.

²¹Alternatively, someone could adopt the extreme position that *all* resolutions of coun-

above discussion establish? It establishes a natural way that indeterminate permissibility can arise and shows how we can formalise indeterminacy pluralism. This is far from a decisive argument for the view. Nevertheless, it complements the other arguments discussed.

4.4 Resolving the Problem Cases

Finally, the plausibility of indeterminacy pluralism is bolstered by the fact that this theory can make sense of both of the scenarios discussed in the previous section. The solution in both cases will be basically the same. After all, in *Newcomb's Problem* the intuitions and arguments favouring one-boxing and two-boxing are both taken seriously by indeterminacy pluralism because each intuition and associated argument gets things right on one sharpening of the notion of rational permissibility. Then we can make the same move in relation to *the Psychopath Button*: the intuition favouring pressing is captured by one sharpening of permissibility and the intuition favour refraining is captured by another. Consequently, indeterminacy pluralism can make sense of both *Newcomb's Problem* and *the Psychopath Button*. There are four reasons, then, to accepting indeterminacy pluralism.

5 Psychopaths and Context

Still, at this point a problem arises for indeterminacy pluralism. After all, many people find the argument for refraining in *the Psychopath Button* far more compelling than the argument for pressing the button in this scenario.

Consequently, these people conclude that an agent is straightforwardly re-

terfactual indeterminacy play a role in rational choice. However, on such an account it is likely that most decisions will be indeterminately permissible. As such, I take this account to be implausible given the view, in §4.2, of the sort of indeterminacy under discussion here.

quired to refrain from pressing the button here. This raises two challenges to indeterminacy pluralism. First, this theory needs to be able to account for why the argument for refraining has greater intuitive force than the argument for pressing, given that indeterminacy pluralism entails that each of these arguments are right on one sharpening of permissibility. Second, the proponent of this theory needs to be able to explain away the intuition that refraining is straightforwardly rationally required, given that indeterminacy pluralism instead entails that this decision is indeterminately permissible.²²

Fortunately, a simple response to both challenges is available. In particular, it is plausible that contextual factors help determine which sharpenings of permissibility are psychologically salient in some circumstance.²³ Further, if contextual factors make the evidential perspective more psychologically salient than the causal perspective in *the Psychopath Button* then this would explain why people tend to find the argument for refraining more compelling than the argument for pressing and, as a result of this, would explain why people (mistakenly) take refraining to be determinately rationally required.

What contextual factors might play the desired role? Well, for a start, the stakes (that is, the difference in expected value between the two available decisions) in *the Psychopath Button* are higher on the evidential perspective than on the causal perspective (which is to say that there is a greater gap between the EEU of pressing and of refraining than between the CEU of pressing and of refraining). After all, on the evidential perspective pressing is very bad, because pressing kills Paul. On the other hand, on the causal perspective pressing is a far weaker good, as it is merely likely to rid the

²²A number of philosophers deny the force of the intuition that refraining is rationally required (cf. Arntzenius 2008; Joyce 2012). So one might well doubt that this objection is as forceful as I have taken it to be.

²³That is, I am arguing that context helps determine our intuitive judgements. An alternative view: perhaps rational permissibility itself has contextual parameters that sometimes pick out just one sharpening.

world of psychopaths.²⁴ So the stakes are higher on the evidential perspective than on the causal perspective. Further, stakes are key contextual factors that plausibly inform our intuitive judgements and so make the evidential perspective psychologically more salient here.²⁵ In addition to this, moral consideration plausibly influence our intuitive judgements about rationality.²⁶ That is, we are plausibly more likely to judge a decision as rational if it is moral and more likely to judge it as irrational if it is immoral. Now insofar as pressing involves mass murder and so is morally bad, we're likely to be psychologically more responsive to arguments that label pressing as irrational. Consequently, this factor too will make the evidential perspective psychologically more compelling than the causal perspective in the case at hand.

By reference to these factors, indeterminacy pluralism can account for the greater intuitive force of the argument for refraining than the argument for pressing in *the Psychopath Button*: contextual factors are present in *the Psychopath Button* that make the evidential perspective (and hence the argument for refraining) more psychologically salient than the causal perspective (and hence the argument for pressing). Consequently, not only can indeterminacy pluralism capture the sense in which there are two compelling arguments in *the Psychopath Button* but it can also make sense of why one of these arguments exerts greater intuitive weight than the other.

²⁴The case stipulates that Paul values survival over ridding the world of psychopaths.

²⁵For evidence that stakes inform our decision-theoretic intuitions, see MacCrimmon and Larsson (1979: 393) and Anand (1990). For previous discussions that have relied on the assumption of such stake sensitivity, see Nozick 1993 and MacAskill 2016.

²⁶For evidence that our intuitions generally (about, for example, causation and intentionality) are influenced by moral factors see Knobe 2003; Knobe and Fraser 2008.

6 The Smoker's Challenge

So far, so good. Unfortunately, at this point a further problem for indeterminacy pluralism arises. To get to this, consider *the Smoking Lesion*, a prominent example of what is often called a medical Newcomb's Problem:²⁷

Imagine that smoking does not itself increase a person's cancer risk but that smokers are nevertheless more likely to get cancer than non-smokers are because of a brain lesion that both makes a person more likely to smoke and more likely to develop cancer. Cathy would strongly prefer not to develop cancer but enjoys smoking. Should Cathy smoke?

In *Newcomb's Problem*, there is significant disagreement about which decision it would be rational to make. However, in *the Smoking Lesion* there is almost unanimous agreement that Cathy ought to smoke. After all, she enjoys smoking and smoking doesn't increase her cancer risk. In the light of this, Cathy may as well get the pleasure of smoking, given that doing so doesn't come with any disadvantage.

Now this decision is permissible from the causal perspective: smoking doesn't cause Cathy to be more likely to get cancer but it does cause her some enjoyment and so smoking has a better causal impact than not smoking, which doesn't cause the same enjoyment.²⁸ On the other hand, at least on the face of it, not smoking is labelled as rationally required on the evidential perspective: smoking provides Cathy with evidence that she has the lesion, and so is more likely to get cancer, while not smoking provides Cathy with evidence that she lacks the lesion, and so is less likely to get cancer. Not

²⁷This case has been discussed for decades but my presentation follows Egan 2007.

²⁸See Briggs (2010: 7–8) for formal details.

smoking provides better evidence than smoking, so not smoking is rationally required on the evidential perspective.²⁹

As a consequence of this a problem arises. After all, not smoking appears to be outright impermissible here but indeterminacy pluralism instead takes this decision to be indeterminately permissible, as this decision is permissible on one sharpening of permissibility (the evidential perspective) but not the other (the causal perspective). Consequently, indeterminacy pluralism seems to get this case wrong and so we have a challenge to this theory.

Here there are two possible responses. The first appeals to existing defences of EDT, which attempt to show either that this theory does not endorse smoking in *the Smoking Lesion* or that it endorses smoking only in strange versions of the case where our intuitions are no longer reliable. Such defences come in a number of forms but one of the simplest, and most prominent, is the tickle defence (for a more detailed defence of EDT here, see Ahmed 2014). According to the tickle defence, in realistic versions of *the Smoking Lesion*, the lesion will influence Cathy's decision by influencing her beliefs or her desires (at least in versions of the case where Cathy can be said to still be deciding freely). Further, it is argued that Cathy will typically be able to detect this influence (that is, Cathy will feel some sort of *tickle* if she has the lesion). If so, however, then she will already know whether or not she has the smoking lesion, from reflection on whether or not she feels the tickle. As such, smoking will not provide further evidence that she has the lesion. Then, as deciding to smoke still provides her with evidence that she will get the pleasure of smoking, this decision will now provide the most desirable evidence. So the tickle defence concludes that EDT endorses the

²⁹See Briggs (2010: 6–7) for formal details. Note that by pointing to the tickle defence (cf. Eells 1981) it could be argued that EDT, and hence the evidential perspective, endorses smoking. I am dubious that this defence truly succeeds but if it does then the challenge to indeterminacy pluralism collapses immediately. So much the better!

requirement to smoke in *the Smoking Lesion*.

Now, if one accepts this argument outright then the problem considered in this section is resolved, as indeterminacy pluralism now labels smoking as determinately permissible. Still, even if one doubts that the tickle defence *always* does the desired job, it must be accepted that it often does. So, at worst, the evidential perspective endorses not smoking only in versions of *the Smoking Lesion* where tickles do not arise. Further, many such cases will either be unrealistic or will involve Cathy's free will being impinged upon. Consequently, the indeterminacy pluralist can now bite the bullet at little cost: they only need to accept the indeterminate permissibility of smoking in a constrained class of cases, many of which will be unrealistic.

Still, another response is also available. After all, a second option would be to reject the claim that the evidential perspective corresponds to a sharpening of the indeterminate notion of permissibility and instead accept a version of indeterminacy pluralism on which one sharpening corresponds to the causal perspective and another sharpening corresponds to some distinct perspective (that is, a perspective that is neither the causal nor the evidential perspective).

For example, perhaps a second sharpening is captured by a theory along the lines of Spohn 2012 and Easwaran MS. Glossing over the details, these theories are variants on CDT that focus not on the causal impact of the agent's decisions but rather on the causal impact of the agent's dispositions to decide. So according to this version of indeterminacy pluralism, what is indeterminate is not how we should evaluate decisions but rather *what* it is that decision theory should be evaluating (decisions or dispositions).

The disposition theory will endorse one-boxing in versions of *Newcomb's Problem* where the prediction of the agent's behaviour is made based on

the agent's psychological dispositions. As such, the latest version of indeterminacy pluralism continues to deliver the result that one and two-boxing are indeterminately permissible in many versions of *Newcomb's Problem*. However, the disposition theory labels smoking as required in *the Smoking Lesion*, as being disposed to smoke leaves the agent best off, regardless of whether or not she has the lesion. So the latest version of the indeterminacy theory avoids the challenge of *the Smoking Lesion*. As such, moving to this version of the indeterminacy theory presents a second potential line of response here.³⁰

I conclude that not only do the four arguments above support indeterminacy pluralism but that this theory can survive objections based on *the Psychopath Button* and *the Smoking Lesion*. Indeterminacy pluralism is a promising theory of choice.

7 Conclusions

In this paper, I have defended three main claims. First, we should accept pluralism. That is, we should accept that both the causal and evidential perspectives play a role in an adequate account of rational choice. Second, we should accept indeterminacy pluralism. That is, we should spell out pluralism by accepting that our concept of rational permissibility admits of indeterminacy and that the evidential and causal perspectives correspond to different sharpenings of this concept. Finally, we should supplement indeterminacy pluralism by an appeal to the context-sensitivity of our intuitions, to explain why our intuitions sometimes hew particularly close to either the causal or evidential perspective.

³⁰This account will plausibly not deliver the desired result in *the Psychopath Button*. So there is more work to be done to make this account fully adequate.

Further, as well as establishing these three claims, the discussion in this paper provides a lens through which the broader debate over *Newcomb's Problem* can be viewed. In particular, this debate once focused almost exclusively on EDT and CDT but recent years have seen an explosion in the number of interesting decision-theoretic views being defended. For example, Wedgwood (2011), MacAskill (2016) and Spencer and Wells (MS) all present hybrid pluralist versions of decision theory. Meanwhile, Price (2012) defends a view on which CDT and EDT turn out to be one and the same theory, when appropriately construed. Finally, Briggs (2010) has developed an impossibility proof, which uses both causal and evidential principles to attempt to demonstrate that no adequate theory of rational choice can be developed.

At the moment, all of these results, and a number of others, have been presented as disparate competitors and, as a result, the decision theoretic debate has lost the clarity it once had, with possible options multiplying. However, pluralism can bring some coherence back into this debate, as pluralist intuitions underpin all of this recent work. Consequently, focus on pluralism reveals not a disparate mess of views but rather reveals that a new movement in decision theory has developed in recent years and that it is gaining pace.

In light of all of the above, I conclude: by ignoring pluralism, we're doing rational choice theory wrong; we're doing rational choice theory wrong and to do it right we need to take seriously both the causal and the evidential perspectives. This paper, then, is a call to the decision-theoretic barricades and on our uniforms engraved these words: 'decision theory', and below, 'plurality, indeterminacy, contextuality'.

Adam Bales

References

- Ahmed, A. (2014) *Evidence, Decision and Causality*. Cambridge: Cambridge University Press.
- Anand, P. (1990) 'Two Types of Utility: An Experimental Investigation Into the Prevalence of Causal and Evidential Utility Maximisation', *Greek Economic Review*, 12/1: 58–74.
- Arntzenius, F. (2008) 'No Regrets, or: Edith Piaf Revamps Decision Theory', *Erkenntnis*, 68/2: 277–97.
- Briggs, R.A. (2010) 'Decision-Theoretic Paradoxes as Voting Paradoxes', *Philosophical Review*, 119/1: 1–30.
- Dougherty, T. (2014) 'Vague Value', *Philosophy and Phenomenological Research*, 89/2: 352–72.
- Easwaran, K. (MS) 'Towards a Classification of Newcomb-Style Problems'.
- Eells, E. (1981) 'Causality, Utility, and Decision', *Synthese*, 48/2: 295–329.
- Egan, A. (2007) 'Some Counterexamples to Causal Decision Theory', *The Philosophical Review*, 116/1: 93–114.
- Horgan, T. (1981) 'Counterfactuals and Newcomb's Problem', *Journal of Philosophy*, 78/6: 331–56.
- (1985) 'Newcomb's Problem: a Stalemate', in R. Campbell and L. Souden (ed.) *Paradoxes of Rationality and Cooperation: Prisoner's Dilemma and Newcomb's Problem*, 223–33. Vancouver: University of British Columbia Press.

- Horwich, P. (1985) ‘Decision Theory in Light of Newcomb’s Problem’, *Philosophy of Science* 52/3: 431–50.
- Joyce, J. M. (1999) *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.
- (2012) ‘Regret and Instability in Causal Decision Theory’, *Synthese* 187/1: 123–45.
- Knobe, J. (2003) ‘Intentional Action in Folk Psychology: An Experimental Investigation’, *Philosophical Psychology* 16/2: 309–325.
- Knobe, J. and B. Fraser (2008) ‘Causal Judgment and Moral Judgment: Two Experiments’, in W. Sinnott-Armstrong (ed.) *Moral Psychology, Vol. 2*, 441–8. Cambridge: MIT Press.
- Lewis, D. (1973) *Counterfactuals*. Oxford: Blackwell Publishers.
- (1979) ‘Counterfactual Dependence and Time’s Arrow’, *Noûs* 13/4: 455–76.
- (1981) ‘Causal Decision Theory’, *Australasian Journal of Philosophy* 59/1: 5–30.
- MacAskill, W. (2016) ‘Smokers, Psychos, and Decision-Theoretic Uncertainty’, *Journal of Philosophy* 113/9: 425–45.
- MacCrimmon, K. R. and S. Larsson (1979) ‘Utility Theory: Axioms Versus Paradoxes’, in M. Allais and O. Hagen (ed.) *Expected Utility Hypotheses and the Allais Paradox*, 333–409. Dordrecht: Reidel.
- Nozick, R. (1969) ‘Newcomb’s Problem and Two Principles of Choice’, in N. Rescher (ed.) *Essays in Honor of Carl G. Hempel*, 114–46. Dordrecht: Reidel.

- (1993) *The Nature of Rationality*. Princeton: Princeton University Press.
- Price, H. (2012) ‘Causation, Chance, and the Rational Significance of Supernatural Evidence’, *Philosophical Review* 121/4: 483–538.
- Rinard, S. (2015) ‘A Decision Theory for Imprecise Probabilities’, *Philosophers’ Imprint* 15/7: 1–16.
- Schoenfield, M. (2015) ‘Moral Vagueness Is Ontic Vagueness’, *Ethics* 126/2: 257–82.
- Spencer, J. and I. Wells (MS) ‘Rational Monism, Rational Pluralism, and the Metaethical Foundations of Causal Decision Theory’.
- Spohn, W. (2012) ‘Reversing 30 Years of Discussion: Why Causal Decision Theorists Should One-Box’, *Synthese* 187/1: 95–122.
- Wedgwood, R. (2011) ‘Gandalf’s Solution to the Newcomb Problem’, *Synthese* 190/14: 1–33.
- Williams, J. R. G. (2012) ‘Indeterminacy and Normative Silence’, *Analysis* 72/2: 217–25.
- (2014) ‘Decision-Making Under Indeterminacy’, *Philosophers’ Imprint* 14/4: 1–34.
- (2016) ‘Indeterminacy, Angst and Conflicting Values’, *Ratio* 29/4: 412–33.