It Can Be Irrational to Knowingly Choose the Best

J. DMITRI GALLOW *

In some decisions, causal decision theory (CDT) requires you to choose whichever option you predict you will choose. That is: in these decisions, CDT demands that you verify your own prediction about what you'll do. In other decisions, CDT forbids you from choosing whichever option you predict you will choose. That is, in these decisions, CDT demands that you falsify your own prediction about what you'll do.

Some heterodox causalists object to both of these kinds of prediction-sensitivity (see, in particular, Barnett, forthcoming, Gallow, 2020, and Podgorski, forthcoming). They say that what you ought to do does not depend upon what you predict you will do. Their decision theories are prediction-*in*sensitive. However, one heterodox causalist (Spencer, 2021) thinks that there's nothing objectionable about prediction-sensitivity *per se*. He says that, while you are never required to falsify your prediction, you *are* sometimes required to verify your prediction.

Spencer (forthcoming) presents an argument against prediction-insensitive causalists. He points out that their theories appear to conflict with the plausible principle

Knowingly If you know that you will choose an option, x, and you know x is better than every other option available to you, then it is permissible for you to choose x.

I agree with Spencer that prediction-insensitive causalists should reject this principle. However, I disagree insofar as he thinks that he and orthodox causalists are in a better position to accept it. In fact, both orthodox CDT and his own heterodox theory appear to contradict **Knowingly** in exactly the kinds of cases in which prediction-insensitive causalists apparently contradict it. My own view is that we should all deny the principle. We should all accept that it *can* be irrational to knowingly choose the best. But, whether you agree with me about this or not, I hope to persuade you that **Knowingly** does not favour orthodox CDT and Spencer's theory over prediction-insensitive causalism. The reason is that, for the plausible ways that Spencer or an orthodox causalist may attempt to square their theory with **Knowingly**, a prediction-insensitive causalist may just as plausibly square *their* theory with **Knowingly** in a similar way.

1 | PREDICTION-SENSITIVITY

CDT says that you should choose whichever option has the highest *utility*, U, where the utility of an option, x, is

$$U(x) \stackrel{\text{def}}{=} \sum_k C(k) \cdot D(x \wedge k).$$

Here, each k is a state of nature, C is your credence function, and D measures your desires. If you're in the state k, then CDT says that the objective value of choosing x is given by $D(x \wedge k)$, so $U(x)$ is your subjective expectation of x's objective value.

Consider the following decision.

> Amelia must choose between two envelopes labelled 'a' and 'b'. Yesterday, her fairy godfather made a prediction about which envelope she would choose. If he predicted a ('k_a'), then he put \$1 in a. If he predicted b ('k_b'), then he put \$2 in b. Amelia is certain that he predicted correctly.

Conditional on Amelia choosing a, she knows for sure that her fairy godfather predicted she would choose a. And, conditional on her choosing b, she knows for sure that he predicted she would choose b. Then, her credence in k_a will be her credence that she chooses a, and her credence in k_b will be her credence that she chooses b,

$$C(k_a) = C(a) \quad \text{and} \quad C(k_b) = C(b).$$

Assuming Amelia's desires are linear in dollars, the utility of a is her credence that she takes a, and the utility of b is two times her credence that she takes b,

$$U(a) = C(a) \quad \text{and} \quad U(b) = 2 \cdot C(b).$$

Whether $U(a)$ or $U(b)$ is higher depends upon Amelia's predictions about what she will choose. So long as she is confident enough that she'll select a/b, CDT says a/b is required. So in this decision, CDT demands that Amelia verify her prediction.

Prediction-insensitive causalists say that you should make up your mind about what to do in a way which is independent of what you predict you will do. In a choice between two options, x and y, Barnett, Gallow, and Podgorski all advise you to look at a family of *conditional* utilities, $U_y(x)$, where

$$U_y(x) \stackrel{\text{def}}{=} \sum_k C(k \mid y) \cdot D(x \wedge k)$$

is the utility x has, conditional on you selecting y. Then, $R(x, y) \stackrel{\text{def}}{=} U_x(x) - U_x(y)$ measures how much better than y you will expect x to be, conditional on you choosing x. ('R' for *ratifiability*.) If $R(x, y) > R(y, x)$, then prediction-insensitive causalists say that x is required. Spencer names this rule '**MaxRat**'. Applied to Amelia's decision, $R(a, b) = U_a(a) - U_a(b) = 1$ and $R(b, a) = U_b(b) - U_b(a) = 2$, so **MaxRat** says that b is required, no matter what Amelia predicts about how she'll choose.

In contrast, Spencer (2021) says that, in a choice between two options, if you currently expect that x is better than y, and you would *continue* to think this as you grow more confident that you will select x, then x is required. That is, if both $U(x) > U(y)$ and $U_x(x) > U_x(y)$, then x is required. Call this '**Spencer's Rule**'. Applied to Amelia's decision, if she is 80% sure that she'll choose a, then $U(a) = 0.8 > 0.4 = U(b)$ and $U_a(a) = 1 > 0 = U_a(b)$. So **Spencer's Rule** says that a is required.

2 | SPENCER'S ARGUMENT AGAINST MAXRAT

Suppose Amelia *knows* that she will take a. Because she knows that the predictions are accurate, she knows that she will take a iff a contains \$1 and b contains \$0. So, if she knows that she will take a, then she knows that a is better than b. So **Knowingly** says that it is permissible to take a. But **MaxRat** disagrees. So **MaxRat** and **Knowingly** are inconsistent. But Spencer contends that **Knowingly** is "undeniable". So Spencer concludes that we should reject **MaxRat**.

Someone might object that it's not possible for Amelia to know that she will take a, but I don't think we should. A resolute intention to choose a gives Amelia fantastic evidence that she will choose a. She may have a long history of always following through on her resolute intentions. Barring inductive scepticism, Amelia is in a position to know that she'll choose a. Nor should it matter whether this choice is rational. In general, we should acknowledge that

Irrationality is knowable It is possible to know that you'll choose x, even if x is irrational.

Some may object that it's not possible for Amelia to know that the prediction is accurate, but I don't think we should. So long as her choices are predictable, there's no reason why Amelia herself couldn't come to know that they've been accurately predicted. In general, we should acknowledge that

Predictability is knowable It is possible to know that your choice was accurately predicted.

We could grant all this but deny that Amelia can know that a is better than b. But, again, I do not think we should. We should acknowledge that

Knowledge is closed If your belief that ψ was formed by competently deducing it from beliefs, $\phi_1, \phi_2, \ldots, \phi_N$, which jointly entail ψ, then, if you know each of $\phi_1, \phi_2, \ldots, \phi_N$, you know that ψ.

Suppose that Amelia formed her belief that a is better than b by competently deducing it from her knowledge that she will take a and her knowledge that the prediction is accurate, together with her knowledge that, if she will take a and the prediction is accurate, then a is better than b. Then, **Knowledge is closed** tells us that Amelia knows that a is better than b.

Knowingly then tells us that it is permissible for Amelia to take a, which conflicts with **MaxRat**. I do not wish to reject any of these assumptions, so I accept the conclusion: **MaxRat** contradicts **Knowingly**.

3 | AN ARGUMENT AGAINST CDT AND SPENCER'S RULE

But **MaxRat** is not alone. With similar assumptions, both **Spencer's Rule** and orthodox CDT also contradict **Knowingly**. Consider the following decision:

> Casey must choose between two envelopes labelled 'c' and 'd'. Yesterday, his fairy godmother made a prediction about which he would choose. If she predicted c, ('k_c'), she put $1 in c and nothing in d. If she predicted d ('k_d'), she put $1/\epsilon$ (for some $\epsilon > 0$) in d and nothing in c.

Suppose Casey's credence that he will take c is $1 - \epsilon$, and he knows that he will take c. Assuming that Casey knows that the predictions are accurate, he knows he will take c iff c contains $1 and d contains nothing. So, if he knows he will take c, then he knows that c contains more money than d, and so he knows that c is better than d. So **Knowingly** says that it is permissible for Casey to take c.

But both orthodox CDT and **Spencer's Rule** disagree. Because Casey's credence that he will take c is $1 - \epsilon$, his credence in k_c is $1 - \epsilon$, and his credence in k_d is ϵ. So the utility of c is

$$U(c) = C(k_c) \cdot D(c \wedge k_c) + C(k_d) \cdot D(c \wedge k_d)$$
$$= (1 - \epsilon) \cdot 1 + \epsilon \cdot 0$$
$$= 1 - \epsilon$$

whereas the utility of d is

$$U(d) = C(k_c) \cdot D(d \wedge k_c) + C(k_d) \cdot D(d \wedge k_d)$$
$$= (1 - \epsilon) \cdot 0 + \epsilon \cdot (1/\epsilon)$$
$$= 1$$

So orthodox CDT says that c is impermissible, and d is required. Since the utility of d would *continue* to exceed the utility of c if you were to select d, $U_d(d) = 1/\epsilon > 0 = U_d(c)$, **Spencer's Rule** agrees. So both contradict **Knowingly**.

You may object that it's not possible for Casey to know that he will take c. I disagree. I think we are typically in a position to know that we will choose an option on the basis of our intention to do so, independent of whether that option is rational. And I do not think that this knowledge requires certainty; for small enough ϵ, having a credence of $1 - \epsilon$ that you will choose x does not preclude you from knowing that you'll choose x, so long as your evidence is good enough, and you in fact *do* choose x. But suppose you disagree, and you think that Casey cannot know he will take c. Just for illustration, suppose you think that knowledge is sensitive to the stakes of your practical situation. Then, like Fantl & McGrath (2002), you may say that you know ϕ only if it is rational for you to act as if ϕ. And you may suggest that, since it is *not* rational for Casey to act as if he will take c, he cannot know that he will take c. Let's not dwell on the fact that, for good reason, Fantl & McGrath restrict their principle so that it does not apply when

you exercise causal control over whether ϕ.[1] The important point is this: a defender of **MaxRat** could just as plausibly say that it is not rational for *Amelia* to act as if she will take *a*. Applying Fantl & McGrath's principle in the same way, they could then conclude that Amelia cannot know she will take *a*.

The point isn't that Amelia and Casey's situation is symmetric. If you accept **Spencer's Rule**, then you'll recognise an important asymmetry between Casey and Amelia: Casey's choice is irrational while Amelia's is rational. The point is rather that there is a dialectical symmetry between **Spencer's Rule** and **MaxRat**. If pragmatic encroachment allows a proponent of **Spencer's Rule** to deny that Casey knows what he'll choose, it likewise allows a proponent of **MaxRat** to deny that Amelia knows what she'll choose.

Some may object that it's not possible for Casey to know that the prediction is accurate. Again, I don't think we should. But, more importantly: it looks like any plausible reason you may have for denying that Casey can know the prediction made about *him* is accurate is a reason a proponent of **MaxRat** could use to deny that Amelia can know that the prediction made about *her* is accurate.

We could grant both that Casey knows that he will take *c* and that he knows that the prediction is accurate. Even so, we could deny that Casey is in a position to know that *c* is better than *d*, by denying that **Knowledge is closed**. But I don't see any plausible reason to deny *this* instance of closure which couldn't equally well be offered as a reason to deny the instance of closure we used in arguing that Amelia was in a position to know that *a* is better than *b*.

So it seems to me that the argument that CDT and **Spencer's Rule** contradict **Knowingly** is just as strong as Spencer's argument that **MaxRat** contradicts **Knowingly**. And it seems to me that the plausible defences available to CDT and **Spencer's Rule** are equally well available to **MaxRat**.

4 | KNOWLEDGE AND EXPECTATION

Spencer (forthcoming) does not discuss decisions like Casey's, but he does offer a reason to think that Casey cannot know that *c* is better than *d*. When discussing an unrelated objection to **Knowingly**, Spencer claims that you cannot know that *x* is better than *y* if $U(y) > U(x)$ (fn. 16). I believe the idea is this: utility is your subjective expectation of objective value. So, if $U(y) > U(x)$, then *y* is better than *x* in expectation. And if *y* is better than *x* in expectation, then you cannot know that *x* is better than *y*.

Knowledge and expectation   If your expectation of *y*'s value is greater than your expectation of *x*'s value, then you cannot know that *x* is better than *y*.

Compare: if your expectation of Emilia's height is greater than your expectation of Grant's height, then you cannot know that Grant is taller than Emilia.

Knowledge and expectation entails that Casey cannot know that *c* is better than *d*, for $U(d) > U(c)$. And it allows that Amelia can know that *a* is better than *b*, since $U(a) > U(b)$.

---

1. In their first appendix

It's important to recognise that, just because this principle applies to Casey and not to Amelia, this doesn't automatically break the dialectical symmetry between **MaxRat** and **Spencer's Rule**. Suppose you accept the principle and conclude that Casey can't know that c is better than d. It could still turn out that the best explanation for *why* Casey can't know c is better than d would tell a proponent of **MaxRat** that Amelia doesn't know a is better than b. After all, if we accept the principle, and we accept that **Knowledge is closed**, then we will have to deny that Casey knows one of the following: (1) he will choose c; (2) the prediction is accurate; and (3) if he chooses c and the prediction is accurate, then c is better than d. As I said in §3, whatever we say about *why* Casey doesn't know (1), (2), or (3), it looks like similar reasoning could allow a proponent of **MaxRat** to say something similar about Amelia.

In any event, we should reject **Knowledge and expectation**. The principle can sound more plausible than it should if we're not careful to distinguish what's expected in the everyday sense from what's expected in the sense of mathematical expectation. To appreciate the difference, suppose you're 50% sure that Emilia is 5 feet tall and 50% sure than she is 7 feet tall. Then, your (mathematical) expectation of Emilia's height is 6 feet. But, in the everyday sense, you do not expect Emilia to be 6 feet tall—after all, you're *certain* that she's not 6 feet tall! Or suppose you know for sure that Grant is 6 feet tall, and you're $1 - \epsilon$ sure that Emilia is 5 feet tall, but there's an ϵ probability that she's been zapped by a growth ray gun and is now $2/\epsilon$ feet tall. Then, your mathematical expectation of Emilia's height is greater than your mathematical expectation of Grant's height. But you do not, in the everyday sense, *expect* Emilia to be taller than Grant—after all, you're nearly *certain* that Grant is taller than Emilia! Moreover, so long as you can know that the ϵ probability event did not obtain, you can know that Grant is taller than Emilia, in spite of the fact that your expectation of Emilia's height is greater than your expectation of Grant's. And in exactly the same way, you can know that x is better than y, in spite of the fact that your expectation of y's value is greater than your expectation of x's.

5 | KNOWING

There is another noteworthy respect in which Spencer's theory and **MaxRat** are in similar dialectical positions. Spencer (forthcoming, fn 17) briefly discusses the principle

Knowing If you know that x is better than every other option available to you, then it is permissible for you to choose x.

Spencer's theory of rational choice contradicts **Knowing**. To appreciate why, consider this decision:

> Imogen must choose between two envelopes, labelled 'i' and 'j'. If her fairy godfather predicted she'd choose i, he put \$10 in j and nothing in i. If he predicted j, he put \$2 in i and \$1 in j. Imogen is certain he predicted correctly.

In Imogen's decision, $U_i(i) < U_i(j)$ and $U_j(j) < U_j(i)$. So **Spencer's Rule** is silent. But Spencer's full theory of rational choice is not. It tells Imogen she must choose j.

Roughly, in a choice between two options, x and y, if $U_x(x) < U_x(y)$ and $U_y(y) < U_y(x)$, then Spencer says: if $U_x(x) > U_y(y)$, then x is required (see Spencer, 2021, for details and caveats). In Imogen's decision, $U_i(i) = 0 < 1 = U_j(j)$. So Spencer's theory tells Imogen that she is required to take j, no matter what she predicts about what she'll do.

But now suppose Imogen knows that she will choose j. Then, she knows that her fairy godfather predicted that she'd choose j, and so knows that i contains \$2 whereas j contains only \$1. So she knows that i is better than every other option available to her. **Knowing** says that it is permissible for Imogen to choose i. But Spencer disagrees. So he rejects **Knowing**. In fn 17, Spencer (forthcoming) says that decisions like Imogen's are counterexamples to **Knowing**.

Speaking for myself: considered in the abstract, I feel the intuitive pull of **Knowing** and **Knowingly** equally. And I'm just as inclined to take Amelia's decision to be a counterexample to **Knowingly** as I am inclined to take Imogen's to be a counterexample to **Knowing**. If it is a cost of a theory that it rejects **Knowingly**, it seems to me that it should likewise be a cost that a theory rejects **Knowing**. Moreover, once **Knowing** is denied, it seems to me that there's little additional cost to denying **Knowingly** as well. So while these kinds of principles may favour orthodox CDT over both **MaxRat** and Spencer's theory, it doesn't seem to me that they favour Spencer's theory over **MaxRat**.

REFERENCES

Barnett, David James. forthcoming. "Graded Ratifiability." In *Journal of Philosophy*. [1], [2]

Fantl, Jeremy & McGrath, Matthew. 2002. "Evidence, Pragmatics, and Justification." In *Philosophical Review*, **111** (1): 67–94. [4], [5]

Gallow, J. Dmitri. 2020. "The Causal Decision Theorist's Guide to Managing the News." In *The Journal of Philosophy*, **117** (3): 117–149. [1], [2]

Podgorski, Abelard. forthcoming. "Tournament Decision Theory." In *Noûs*. [1], [2]

Spencer, Jack. 2021. "Rational Monism and Rational Pluralism." In *Philosophical Studies*, **178**: 1769–1800. [1], [3], [7]

Spencer, Jack. forthcoming. "Can It Be Irrational to Knowingly Choose the Best?" In *Australasian Journal of Philosophy*. [1], [5], [6], [7]