

Representation and Rationality

Ray Buchanan & Sinan Dogramaci

Abstract: David Lewis (1974, 1994/1999) proposed to reduce the facts about mental representation to facts about sensory evidence, dispositions to act, and rationality. Recently, Robert Williams (2020) and Adam Pautz (forthcoming) have taken up and developed Lewis's project in sophisticated and novel ways. In this paper, we aim to present, clarify, and ultimately object to the core thesis that they all build their own views around. The different sophisticated developments and defenses notwithstanding, we think the core thesis is vulnerable. We pose a dilemma by considering the two sides of a current epistemological controversy over the relation between evidence and rational belief: permissivism vs. uniqueness. As we argue, the prospects for the Lewisian project look dim when either supposition is clearly made.

Key words: reduction of mind, David Lewis, rationality, permissivism, uniqueness

Word Count: 5890

Introduction

Karl believes this, he desires that, and more generally he has various mental states that carry representational contents. In virtue of what is this so? How do the non-representational facts ground such facts about mental representation?

David Lewis (1974) made an ambitious proposal: the correct assignment of beliefs and desires to Karl holds in virtue of the fact that that assignment maximizes Karl's *rationality* given the evidence of his senses and given his behavior. (Lewis's proposal offers one clear version of earlier ideas from Quine, Davidson, Dennett, Grandy, and others.) Robert Williams (2020) and Adam Pautz (forthcoming) have each taken up Lewis's proposal and developed and defended it in their own sophisticated ways. Though their views are different, they both build on Lewis's core idea that our representational attitudes are grounded in facts about rationality, and their work plausibly shows us what the strongest case for the Lewisian project looks like.

But we will argue that these attempts to answer a big question in philosophy of mind have buried important issues from contemporary epistemology, and when we re-frame the mind problem while highlighting these epistemological issues, it becomes clear that rationality cannot play such a load-bearing role in the theory of mental representation.

The View

The version of the view that we will target our attack on is just what we've already said. Let's give the view an official indentation and label, and then we'll provide some clarificatory commentary.

Rationality-Max: An assignment of beliefs and desires to Karl is correct in virtue of the fact that it best jointly maximizes (1) his epistemic rationality given the evidence of his senses, and (2) his practical rationality given his dispositions to act.

For the main endorsements of the view, see Lewis (1974, pp.336-7), Williams (2020, ch.1), and Pautz (forthcoming, section 3.2). We'll call these advocates of *Rationality-Max* the "Lewisians". (Page references are to these three works when not marked otherwise.)

We of course use Lewis's character Karl as a placeholder for anyone.

We follow Williams and Pautz who said we should appeal to Karl's "dispositions to act" rather than just to his actual actions. (Arguably Lewis intended this too---see p.331---though he is less clear on this.) This means the Lewisians can consider not only facts about Karl's actual actions but also some facts about his counterfactual actions, actions that he would take. But while the Lewisians may appeal to *some* facts about actions Karl would take in other possible worlds, they may not appeal to *all* facts about the possible actions he takes.¹ They cannot appeal, for example, to the counterfactual that Karl would take this or that action if he were stipulated to have these or those very different attitudes than he actually has. The theory aims to give the grounds of all the belief/desires facts, which of course are not themselves belief/desire facts. So, the appeal to dispositions to act is helpful to the Lewisian but only by letting us consider, for

¹ We don't find this clarification explicitly acknowledged anywhere by any of the Lewisians, but we think they must agree.

example, how Karl would act in nearby possibilities like ones where, say, he sees an ice cream in front of him, even though he actually never comes near any ice cream. (If Karl were to see an ice cream, then of course he would form some beliefs different from his actual ones, including a belief that there is ice cream here---that's all perfectly fine. What we prohibit is the Lewisian including any belief/desire facts---including the fact that Karl would believe there is an ice cream here---in the full grounds of his having his attitudes.) With this clarification in place, we'll henceforth refer to Karl's "actions" intending thereby to refer to the appropriate range of all of his actual and some of his counterfactual actions.

By epistemic rationality, we mean the traditional condition on knowledge, as in the JTB theory. (We take "rational belief" and "justified belief" to be synonyms.) By practical rationality, we mean the feature that Karl's attitudes and actions have (or we can say it's a feature just those actions have, or a feature Karl himself has) when the actions would promote the fulfillment of Karl's desires if his beliefs were true. This is what some people may call "instrumental" rationality. (In a credal framework, some kind of expected utility maximizing norm may generalize this kind of notion of practical rationality. Here we'll usually keep the presentation in the simpler belief/desire form. Our objections easily generalize to expected utility decision theories, as we'll point out below.)

The Lewisians sometimes suggest building more into practical rationality. For example, Williams (pp.26-8 and ch.4) proposes that rationality demands that Karl's attitudes and actions be *moral*. Below, after we introduce our arguments, we will consider the important question of whether the Lewisians can overcome them by strengthening the demands of practical rationality beyond our instrumentalist gloss, whether by adding such moral demands or others.

Rationality-Max says intentionality is grounded in a fact about how two things are *jointly* maximized. How do we measure this? Sometimes we'll have a choice between one assignment of attitudes that gives Karl a high score on epistemic rationality at the expense of a low score on practical rationality, and another assignment that does the reverse. What decides which assignment jointly maximizes his epistemic and his practical rationality, i.e. which assignment has the highest "overall" score? Lewis, Williams, and Pautz don't tell us how to measure Karl's overall rationality score. (Eriksson & Hajek, 2007, p.201, complained that Lewis didn't say how

to do this. For Williams's most relevant comments, see pp.27, 100.) But we won't complain about this. We would be happy to grant, as the Lewisians sometimes suggest, that it would be indeterminate which of two assignments is true of Karl when neither one beats the other on *both* epistemic and practical rationality maximization. (Lewis, pp.342-3, suggests indeterminacy arises when this happens. And Pautz, near the end of section 3.2, top of p.25 in ms., says that determinate attitude attributions require the most rational assignment to be unique.)

Rationality-Max aims to reduce (give the grounds of) only belief and desire. The Lewisians, and we too, work with this convenient restriction. Our objections to this view will undermine any more ambitious developments of the project that aim to reduce additional representational states, for example fearing or wondering.

The Lewisians want to ground the facts about how language represents in facts about mental representation, which *Rationality-Max* in turn grounds in facts about evidence and action. However, some sophisticated attitudes might require Karl to have a language to help ground them, such as attitudes about sophisticated mathematical, scientific or social theories. (Pautz in particular, secs. 3.3-3.4, strongly recommends this way of executing the program.) Let's bracket whatever those sophisticated attitudes would be and focus our target just on the sufficiently unsophisticated "prelinguistic" attitudes that the Lewisians agree can be fully grounded just by facts about evidence and actions. Let *Rationality-Max* be understood to concern just these prelinguistic attitudes, wherever the Lewisians want to draw the line around them. (Pautz's intention is that they are any beliefs and desires that a life-long Robinson Crusoe could have. These are attitudes that concern local environment facts that our sensory evidence can give us reasons to believe, and even some somewhat spatiotemporally distant facts that a Crusoe can easily get inductive reasons to believe, such as the colors of certain as yet unexamined emeralds.)

With this restriction to the prelinguistic in place, as well as our caveat that practical rationality might extend beyond instrumental rationality, *Rationality-Max* states a core thesis of the Lewisian project that all our interlocutors, Lewis, Williams, and Pautz, agree on. They do each diverge from the others when it comes to elaborating the further features of the complete theory of the grounds of mental representation. (The most significant point of divergence may be that they all treat the "evidence" facts differently: Lewis understands them as purely physical

facts, Williams aims to reduce them to teleological tracking facts via Karen Neander's work (2017), and Pautz treats them as irreducible phenomenally conscious representational facts.) But despite these differences in how to further elaborate the program, they all agree on *Rationality-Max*.

Rationality-Max, even if it succeeds, only reduces the belief and desire facts to evidence and action facts. If the facts about Karl's evidence and his actions are themselves representational facts, or if they are grounded in representational facts, then the proposal does not give us a completely reductive theory of the representational to the non-representational. We won't complain about this character of the proposal. (Pautz and Williams both accept it.) It would still be tremendous progress on the larger program (of reducing the representational to the non-representational) if we could reduce the facts about what we believe and desire to facts about what is most rational given our evidence and actions. But we will argue that even this project doesn't succeed.

Dilemma: Permissivism or Uniqueness?

We will frame our critique of *Rationality-Max* by setting up a dilemma that uses the two sides of a major debate in contemporary epistemology, the debate over *uniqueness*.

Uniqueness says that, for any person, given their total evidence, there is exactly one set of attitudes that that evidence justifies (or makes rational---again we understand "justified" and "rational" as synonyms). On the popular view that people have credences, uniqueness says that, for any person, given their total evidence, there is a unique credence in each proposition that is the rational credence for that person. (We have here stated the so-called "intrapersonal" version of uniqueness. The "interpersonal" version moves some quantifiers around: it claims that, given a total body of evidence, there is a unique credence that any person with that total evidence must have. Since the "intrapersonal" version is entailed by the "interpersonal" version of uniqueness, our objections to the former will apply to both.)

Opponents of uniqueness endorse *permissivism* and are called permissivists. (Advocates of uniqueness are called impermissivists.) Most philosophers find permissivism to have far more face value appeal than uniqueness. None of our three Lewisians explicitly mentions the

philosophical controversy over uniqueness and permissivism, but at one point Williams (p.27, bullet point 3) does explicitly say that, in general, multiple attitudes will be rationally permissible.

Let's now examine how defensible *Rationality-Max* looks when we assume either side of the debate over uniqueness.

The Permissivist Horn of the Dilemma

Suppose then, for the first horn of our dilemma, that we are permissivists. As mentioned, Williams seems to endorse permissivism. We offer two arguments that, on the assumption of permissivism, *Rationality-Max* fails to adequately pin down Karl's attitudes.

Let's use an example to make it easy to present our arguments. Suppose that Karl's total evidence puts him in a permissive case regarding the belief that the next emerald he finds in that quarry over there will be green, just like the couple of emeralds he's already found and observed. We'll treat this as a paradigm permissive case. (Schoenfield 2019 endorses permissivism specifically about this kind of simple enumerative inductive example, and as mentioned Pautz explicitly classifies it as a prelinguistic belief.) Karl is permitted to adopt the belief that this next emerald will be green, and he is permitted to refrain from adopting the belief. We assume that what is, in this way, permitted for Karl is also *possible* for Karl. This mild permitted-implies-can assumption will let us move from our normative premises about permissivism to our conclusions in the metaphysics of mind. There are two different belief states it is possible for Karl to have here. The indeterminate assignment of both states, or of neither, is not a witness to these *two distinct* possibilities---Karl can believe, and, as an alternative possibility, Karl can refrain from believing. We can imagine this as two possible Karls. Bold Karl-1 proceeds to make the prediction that the next emerald will be green, a prediction which the more cautious Karl-2 is unwilling to make on the evidence they share.

We ask the Lewisians: what grounds this difference in the two possible Karls' beliefs? It cannot be a difference in evidence, because we said they share all their evidence.² So, according

² In an inductive case like this one, the Karls' beliefs also share their logical forms, so we don't have to worry about Williams's views (ch.2) about how differences in inferential dispositions help assign different contents to our (mental concepts for) logical operators.

to *Rationality-Max*, the grounds must involve a difference in how they do or would act. It must be that Karl-1's actions are practically rational for a believer that the next emerald will be green, but these actions are not (equally) practically rational for an agnostic. And conversely, it must be that Karl-2's actions are ones that are practically rational for an agnostic about the next emerald's being green, but not for a believer.

Our first argument against *Rationality-Max* is this: it is intuitively *possible* for the two Karls to have all the same dispositions to act, even though they have different attitudes. As far as the next emerald goes, they might, for example, share a disposition to seek it out, even though these actions are associated with different beliefs and desires. Karl-1, the believer, might go searching for the emerald because he is motivated by his love of green emeralds. And agnostic Karl-2 might be equally disposed to search for the emerald, but one part of Karl-2's motivation is his love of blue emeralds, and he thinks there's a chance that this emerald will be blue. And, we suggest, all their other dispositions to act can safely be stipulated to be the same. Both Karls might equally desire both colors, indeed they might share all the same desires, but different particular desires are motivating their parallel behaviors. (And remember a point we made when clarifying *Rationality-Max* earlier: we cannot appeal to actions Karl would take if he believed green emeralds are worth a lot of money, and blue emeralds are not, or anything like that.) This contradicts *Rationality-Max*, which says that a difference in attitudes requires some difference in evidence or action.

Here is our second argument, which is more general and more ambitious, and doesn't rest on an intuition about a specific possible case. Consider again the question: when Karl's evidence permits either of two belief states (believing the emerald will be green, or harboring doubt it will be green), how can his actions finish the job of determining what he actually believes? What's required, according to *Rationality-Max*, is that Karl's actions settle the question by making him more practically rational if he has one of these beliefs rather than the other. But, we claim, his actions cannot do this. And the reason they cannot do this is that, while permissivism about epistemic rationality is a popular but controversial position, a corresponding kind of permissivism about practical rationality is uncontroversially true. However he is disposed to act, perfectly practically rational Karl could have (nearly) any beliefs (we discuss possible exceptions

shortly). Practical rationality issues only a structural norm on Karl's attitudes: his beliefs *and* desires must holistically fit together in a way that makes him and his actions practically rational (Lewis 1974, p.337; Pautz p.20 in ms). What practical rationality rules out, given Karl's actions, is only certain belief-desire combinations---different "packages" of attitudes. And, since Karl's evidence didn't constrain what he desires at all, he can be perfectly practically rational with either one of his epistemically permitted beliefs. All he needs to do is pair his belief with the right desires. Our first argument gave an illustration of how the recipe goes here: consult Karl's actions (he seeks out the emerald), and attribute the right desires (he wants a green / blue emerald) that will fit with whichever belief he has (he believes the next emerald will be green / might be blue) so as to rationalize his actions. So, when epistemic and practical rationality are both permissive in this way, Karl's evidence and actions will fail to determine his attitudes.

Let's pause to point out that the problems we've raised look just as bad for the Lewisians if they use a credal framework. We've first presented things using the full belief framework for ease of presentation, but like our Lewisian opponents, we think credences are fundamental to a realistic theory of mind. Suppose Karl is in a permissive case that permits him to give either one of two different levels of confidence to the proposition that the next emerald will be green. (He is permitted to be pretty sure it will be green, and he is also permitted to be not quite so sure.) Karl also has his dispositions to act. (He seeks out the emerald.) Given those actions, we can find a utility function that allows us to interpret him as perfectly practically rational. (And we can even plausibly assign the same utility functions to Karl-1 and Karl-2 in the way that we indicated in our first argument. In Karl-1 the high utility of a green emerald is the only "motivating" utility for his emerald seeking behavior since Karl-1 has high confidence that the emerald will be green. In Karl-2, since there is a chance the emerald will be blue and a chance it will be green, his action is based on both the high utility of a blue emerald and the high utility of a green emerald.) As in the full belief model, Karl's evidence and actions don't suffice to pin down which one of two different possible, and perfectly rational, sets of attitudes he actually has.³

³ Some overly ambitious decision theorists think that a so-called representation theorem can be used to assign a unique set of rational (and by 'rational' these theorists usually just mean probabilistically coherent) credences to Karl given his actions, but no representation theorem also assigns Karl unique utilities, so this wouldn't help the Lewisians. In any case, none of our Lewisians ever appeals to or even mentions a representation theorem, and this

How can the Lewisian object to our arguments here? We've argued that, when epistemic rationality and practical rationality are both permissive, Karl's evidence and actions fail to determine a unique set of maximally rational attitudes among the options that are rationally permitted, and thus possible, for him. And the problem of underdetermination is severe, because Karl's evidence and the norms of epistemic rationality impose no constraints on Karl's desires, and Karl's actions and the norms of practical rationality impose a structural constraint that only rules out various combinations of beliefs and desires—no particular beliefs or particular desires are ruled out.

The Lewisian can reply by arguing that the demands of rationality are tighter than we first characterized them to be, even though they are still permissive (by hypothesis for the epistemic, and uncontroversially for the practical to at least some extent). Pautz and Williams each have views that suggest how this might go.

Pautz (end of sec. 3.1 and start of sec. 3.2, p.19 in ms) says that once Karl has the sensation of pain, it is rational for him to desire to not have it. Pautz is saying that Karl's sensory *evidence* puts some constraint on what it is *practically* rational for him to *desire*, an interesting suggestion that crosses the usual wires, since usually evidence hooks up to epistemic rationality and action hooks up to practical rationality. Pautz is happy to ground the representational in the phenomenal, but others won't be so willing to compromise the prospect of reducing the representational to the physical, including our other Lewisians, Lewis and Williams. And many philosophers just won't accept Pautz's claim that pain makes it rational to desire avoidance: many philosophers would instead say that a pain with no desire for avoidance is not irrational, but rather is simply impossible. But, in any case, even if the Lewisian can appeal to Pautz's link between pain and rational desire, we don't think our arguments against *Rationality-Max* and permissivism can be avoided, for three reasons.

First, Pautz's link is more plausible if it is a permissive one: given his pain experiences, perfectly rational Karl could have different strengths of desire to avoid pain. But then we are back to the situation where Karl's evidence permits different attitudes, and his actions permit

kind of ambitious metaphysical application of representation theorems has been very thoroughly debunked already—see Meacham and Weisberg (2011) and Hattiangati and Stefansson (forthcoming).

different attitudes, and we still don't see how the Lewisian can extract a unique set of maximally rational attitudes for Karl from the slack constraints imposed by these two anchors.

Second, we could run our argument by considering creatures that never experience pain. Suppose Karl never experiences pain. These sorts of pain-free cases would still be enough to raise a problem for the Lewisian project, which aims to explain how any creature is assigned attitudes. Could the Lewisian get any help by appealing to the actions Karl would be disposed to take if he *were* to experience pain? On its own, that would not be enough to make any difference to Karl's behavior with emeralds. Karl would need to also somehow differentially associate pain with, say green emeralds as opposed to blue ones, and then we worry it's starting to look like the Lewisian is just appealing to worlds where Karl is specified to have certain beliefs very different from his actual ones, which we've emphasized the Lewisian may not do.

Third, and perhaps most importantly, even if our last point is wrong and the Lewisian could somehow safely insist that Karl would act differently if he were to learn, say, that green emeralds cause pain but blue ones don't, we want to insist that this is not a plausible position about what grounds Karl's actual attitudes. We want to insist that it's just not plausible that the facts about Karl's possible desire to avoid (sensory) pain play any role at all in grounding his attitudes about completely unrelated matters, such as his attitudes about the colors of emeralds. If we stipulate, as we can, that Karl actually has no reason at all to suspect that emeralds, whatever their color, pose any risk of causing (sensory) pain, then Karl's attitudes about emeralds are in no part grounded in any facts about pain.

To be clear, we don't believe Pautz would ever suggest that the irrationality of desiring pain helps ground Karl's attitudes about emeralds. We just want to examine all the remotely plausible options for tightening up the demands of rationality so that we can see if they can overcome our objections. This one doesn't.

Another option for tightening the demands of rationality is Williams's view (p.26, ch.4) that perfectly rational Karl must obey morality in certain ways. But again, we would give the same lines of reply we gave for pain: in a case like Karl's beliefs and desires about some colorful emeralds, morality plays no plausible role in grounding those attitudes. We chose to feature a

case with colors and emeralds partly because it so clearly has nothing to do with morality and we can suppose there's no prospect of pain involved.

To avoid our arguments, what the Lewisians need is for there to be some demands on the rationality of Karl's desires for green or blue emeralds, and these demands are not merely structural but substantive---they hold regardless of what Karl believes or does. But we think this is absurd, and none of the Lewisians make such an unwise suggestion. Is there any desire concerning green things or gemstones that we can know any perfectly rational person must have? Is a green emerald something like Anscombe's (1958, p.70) famous example of a saucer of mud, something that we can by default take it any rational person does not (or does?) have a basic desire for? However plausible (or utterly implausible, we happen to think) such a claim about the basic desirability of a saucer of mud may be, such a claim about the basic desirability or undesirability of a green emerald is flatly implausible. The desirability of a shade of color or a gemstone is completely arbitrary and can vary from rational person to rational person. Anyone's desires concerning a green emerald are just too potentially variable for them to be grounded in a brute fact about what any rational person desires.

In the end, nothing in Karl's evidence or actions seems to ground the fact that he actually has one or another of the different but equally rational sets of beliefs and desires that he could possibly have. We conclude that permissivism is not a good option for advocates of *Rationality-Max*. We think this conclusion should be intuitive, at least in retrospect. If the demands of rationality are slack, then they are ill-equipped to pin down exactly what our attitudes are.

The Uniqueness Horn of the Dilemma

What if we supposed uniqueness then? If permissivism leaves the demands of rationality too slack to pin down Karl's attitudes in the ways Lewisians hope to, then that of course suggests endorsing uniqueness might have been the wiser option all along. But, as we'll now argue, *Rationality-Max* doesn't end up looking very plausible on this horn of the dilemma either.

The assumption of uniqueness only concerns epistemic rationality. It says that, given his evidence, there is a unique set of beliefs that would be maximally epistemically rational for Karl.

So, if we cared only about assigning him the maximally epistemically rational attitudes, we would know exactly what beliefs to assign to Karl, given his evidence. But of course, *Rationality-Max* has us also care about assigning him desires and making him practically rational. How then should we assign desires to Karl? The odd thing, and the problem now, is this: having assigned him maximally epistemically rational beliefs, we can now also assign Karl a set of desires that make him come out maximally practically rational as well. We've already fixed his beliefs, so what we'll do now is just look at his actions, and we will assign to him exactly those desires that make his actions perfectly practically rational. These will be whatever desires those actions most effectively promote (the fulfillment of), supposing the truth of the beliefs we've already assigned to him. We have the absurd result that, on the uniqueness horn of the dilemma, Karl gets to be perfectly rational.

We're relying on just this premise to take us to the claim that we can assign Karl desires that make him perfectly practically (as well as epistemically) rational: if you have (perfectly rational) beliefs *B*, and your actions are *A*, then there will be at least some set of rational desires *D* you could have that would make you (you, your attitudes, your actions) perfectly rational, epistemically and practically.

Now, we know Williams will object to this premise for cases where, given *B* and *A*, we must assign Karl desires *D* that are immoral. So let us bracket all such cases and see if the premise holds otherwise. Let's only consider morally neutral cases where the premise will be unobjectionable, such as when Karl is only dealing with and only thinking about things like green emeralds. We can even restrict our argument to creatures that never form any attitudes toward such moral matters. And Pautz, we know, will object to the premise for cases where Karl is assigned a desire for pain. As we discussed earlier, Pautz says Karl's having an experience of pain gives him a reason to desire that it go away, but, again, let us accommodate it by restricting our objection to the range of cases where Karl has, and would have, no reason to think his actions pose any risk of pain, again for example when he's just dealing with emeralds. We can even again restrict our argument to creatures that never experience pain. (And remember, the Lewisians cannot appeal to what Karl is disposed to do when we stipulate that he has entirely different beliefs.)

This set of rational desires D that we assign to Karl would seem to be a unique set of desires, but even if two different sets of desires were equally rational given Karl's actions, we would only get the result that he indeterminately has either set. As we said earlier, we don't want to complain about indeterminacy here. Our complaint about *Rationality-Max*, on this horn of the dilemma, is that (morally fraught and painful cases aside) it assigns to Karl a (determinate or not) set of attitudes to Karl that is perfectly rational both epistemically and practically. There's no gap left between what attitudes are rational for Karl and what attitudes he really has, and that's of course absurd.

To resist our objection here, the advocate of *Rationality-Max* must try to insist that, even in morally neutral and painless cases, we still go wrong with our premise that says, again, given any set of rational beliefs B and set of actions A , there is some set of rational desires D that makes the actions in A rational. The Lewisians have an available position in logical space here but, as it was on the permissivist horn, it's a costly way to escape our objection. Lewis and Williams do indeed point toward this line of resistance when they suggest that some *non-moral* desires are substantively irrational. These are desires that are irrational no matter what you experience or believe and no matter how you act. They both give the same example (Lewis 1994/1999 p.320; Williams 2020, pp.16, 25, 28), which we already mentioned above, that famous example of an (allegedly irrational) basic desire for a saucer of mud. But it seems plausible to us---and we think many people will see it this way too---that either there are *no* desires that are in this way substantively irrational (every non-moral preference is as personal as one's preference concerning green gemstones), or else there are *extremely few* (after all it's hard to come up with many compelling examples). If there are *none*, then the result is still that Karl is perfectly rational. And if there are only *a few*, then we can say that Karl is irrational only insofar as we must attribute these bad desires to him, but this seems to be only a small, local ding to his overall rationality score, and he will still come out very close to perfectly rational. And there will be little pressure to attribute very different desires or beliefs to him, since doing so will not give him a higher score on *both* practical and epistemic rationality---it would be another indeterminate case. So, in order to make it possible for Karl to be even moderately irrational as we humans always are, or to make Karl as deeply irrational as we sometimes are, the Lewisian is

committed to thinking there is a pretty large range of things that it is substantively irrational to desire, that is, there are lots of desires that it is irrational to have no matter what else you may think or do. The famous example of the saucer of mud isn't enough to make this a very attractive option, we think. Our own view is that there are no such substantively irrational desires, but at the very least we'll be happy to have shown that the viability of the Lewisian project is tied to claiming otherwise.

After bringing up the possibility that a basic desire for the saucer of mud is irrational, Williams (pp.28-9) says that he himself doesn't know, and remains agnostic, on the question of whether such a desire really is irrational. He leaves it to the theorists working on first-order normativity to figure this out, and so *Rationality-Max* must wait for and depend on a very favorable---and we think unlikely---outcome from this independent work in another field of philosophy.

Conclusion

What is the real connection between rationality and representation? We think the connection is a contingent one. It is possible for you to have the beliefs and desires that you have even when your evidence and actions are compatible with your having very different beliefs and desires that would be far more rational, both epistemically and practically.

If not along the Lewisian lines, then how does the reduction of the representational to the non-representational go? Or should we lose hope in physicalism and turn to a primitivist or a dualist theory of representation? No, we remain optimistic about physicalism. One possibility we find appealing is that, while any representational fact is grounded in some physical or functional facts, it is an *a posteriori* matter which such facts do the grounding (as suggested by Horwich 1995, 2005 and Soames 1998). We learn, through a mix of introspection of our own mental states and a partly simulation-based method of mind-reading, how to reliably attribute beliefs and desires to ourselves and to others, and we also thereby learn about the contingent connections between representation and rationality. (On the role of simulation in mind-reading, see Barlassina and Gordon 2017.) But as for finding the physical grounds of Karl's attitudes, it is as opaque to us as the physical grounds of the mental states we first introspect in ourselves. We leave this

brief positive suggestion here just for the sake of offering something that might keep others from losing hope for physicalism.

Acknowledgements: For help with this paper, we are very grateful to a reading group at UT Austin, Daniel Drucker, Miriam Schoenfield, and Gary Ostertag and especially Adam Pautz and an anonymous referee. The two authors contributed equally to this paper and would like to thank each other.

Bibliography

- Anscombe, Elizabeth, (1957), *Intention*, Harvard University Press
- Barlassina, Luca and Robert Gordon, (2017), “Folk Psychology as Mental Simulation”, in E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/archives/sum2017/entries/folkpsych-simulation>
- Eriksson, Lina and Alan Hajek, (2007), “What Are Degrees of Belief?”, in B. Fitelson (Ed.), *Studia Logica: An International Journal for Symbolic Logic*, Vol.86, Special Issue: *Formal Epistemology*, (pp.183-213). doi: [10.1007/s11225-007-9059-4](https://doi.org/10.1007/s11225-007-9059-4)
- Hattiangati, Anandi, and H. Orri Stefansson, (forthcoming), “Radical Interpretation and Decision Theory”, *Synthese*. doi: [10.1007/s11229-021-03078-8](https://doi.org/10.1007/s11229-021-03078-8)
- Horwich, Paul, (1995), “Meaning, Use, and Truth”, *Mind*, Vol.104, (pp.355-368). doi: [10.1093/mind/104.414.355](https://doi.org/10.1093/mind/104.414.355)
- -----, (2005), *Reflections on Meaning*, Oxford University Press
- Lewis, David, (1974), “Radical Interpretation”, *Synthese*, Vol.23, (pp.331-44). doi: [10.1007/BF00484599](https://doi.org/10.1007/BF00484599)
- -----, (1994/1999), “Reduction of Mind”, in S. Guttenplan (Ed.), *A Companion to Philosophy of Mind*, Blackwell, page references to reprint in Lewis’s *Papers in Metaphysics and Epistemology*, (pp.291-324), Cambridge University Press
- Meacham, Christopher, and Jonathan Weisberg, (2011), “Representation Theorems and the Foundations of Decision Theory”, *Australasian Journal of Philosophy*, Vol.89, (pp.641-663). doi: [10.1080/00048402.2010.510529](https://doi.org/10.1080/00048402.2010.510529)
- Neander, Karen, (2017), *A Mark of the Mental*, MIT Press

- Pautz, Adam, (forthcoming), “Consciousness Meets Lewisian Interpretation Theory: A Multistage Account of Intentionality”, in U. Kriegel (Ed.), *Oxford Studies in Philosophy of Mind*, ms. online at <https://philpapers.org/rec/PAUCML>
- Schoenfield, Miriam, (2019), “Permissivism and the Value of Rationality”, *Philosophy and Phenomenological Research*, Vol.99, (pp.286-297). doi: [10.1111/phpr.12490](https://doi.org/10.1111/phpr.12490)
- Soames, Scott, (1998), “Skepticism about Meaning”, in A. Kazmi (Ed.), *Canadian Journal of Philosophy*, Supplementary Vol.23., (pp.211-249)
- Williams, Robert, (2020), *The Metaphysics of Representation*, Oxford University Press