

The Euthyphro Challenge in Metasemantics

Bar Luzon

Forthcoming in *Philosophy and Phenomenological Research*

Abstract: This paper argues that functionalist metasemantic views, such as Conceptual Role Semantics and Interpretivism, face a Euthyphro challenge. The challenge, put roughly, is this: functionalist metasemantic views reverse the order of explanation. According to such views, representational mental states have the contents that they do partly because they play certain roles in our mental lives. According to an intuitive picture of the roles that representational mental states play in our mental lives, however, these states play the roles they do partly because they have the contents that they do. Since explanations—even partial ones—are asymmetric (call this principle ‘the Euthyphro Principle’), we can’t accept both this intuitive picture and a functionalist metasemantics. Hence, we should reject all versions of the latter. This paper can be seen as an attempt to better articulate an intuition that has been suggested in various forms in the literature. The paper proceeds as follows. First, it shows that functionalist metasemantic views imply certain ‘partly because’ claims. Second, it argues for the Euthyphro Principle. Third, it defends the intuitive picture of representational mental states and shows that it implies ‘partly because’ claims that conflict with those made by functionalist metasemantic views. Finally, it considers ways of relaxing the intuitive picture and finds them all wanting.

Ella is standing in her office and looking through the window. She can see the rain falling down, and she decides that she had better take the train back home rather than walk, as she didn’t bring an umbrella with her. For a moment, she can’t remember what day it is—is it Tuesday or Wednesday? If it’s Tuesday, then the weather forecast she read this morning was plainly wrong—it predicted a sunny day. If it’s Wednesday, the forecast did fine—it said that there was a good chance of rain. So she concludes that it’s either Wednesday or she needs to find a new weather forecast to rely on. Ella then looks down at her desk and sees the handout she used in her lecture today, realizing it must be Wednesday—she teaches Monday and Wednesday, but not Tuesday. This realization causes her to infer that there’s a talk tonight, for there is one every Wednesday. She checks out the schedule to see who’s the speaker and what’s the subject. Ella sees that it’s a talk about metasemantics, and is disposed to conclude that she’s going to disagree with everything said in it. Reading the abstract leads her to believe that the speaker’s view is a version of Conceptual Role Semantics, prompting a decision to finally write down that Euthyphro challenge paper she’s been toying with for about a year now.

This story you just read must have a familiar ring to it—not because you have faced this particular series of quite mundane events yourself, but because you are familiar with the kind of mental activity described. Ella gathers information both from the world and from her own memory, forms beliefs on the basis of this information (that it is raining), and goes about forming further beliefs (that today is either Wednesday or...) and making decisions (to finally write this paper), as well as performing actions (checking the schedule). The ‘going about’, here, importantly depends upon the ‘gathering’: Ella first comes to form certain beliefs¹ about what the world is like, and only on the basis of those she forms other beliefs, makes decisions, or performs actions. These beliefs, decisions and actions are *explained* by her initial beliefs (and desires), and by further beliefs she forms on the basis of those initial beliefs.

Why did Ella infer that it’s either Wednesday or she needs to find a new weather forecast to rely on? Because she believed that it was raining, that the forecast she read predicted rain on Wednesday but not on Tuesday, and that it was either Tuesday or Wednesday. Why did Ella look at the schedule for the Wednesday talks? Because she came to believe it was Wednesday, believed there is a talk every Wednesday, and desired to know what this talk was about. Notice that Ella’s beliefs not only explain her other attitudes and her actions, but they also *rationalize* them, i.e., explain why they’re rational. Presumably, Ella’s belief that it was Wednesday was rational. Why? Because it was formed on the basis of her belief that it was either Tuesday or Wednesday, that she taught that morning, and that she teaches on Wednesday but not Tuesday.

To generalize this very natural description of what goes on in our example, we can say that representational mental states such as beliefs have two important roles: (1) they represent the world as being a certain way and (2) they help explain and rationalize our actions, as well as other mental states. Call the first role ‘the semantic role’, and the second role ‘the explanatory role’. Those two roles are tightly connected: it is in virtue of fulfilling the former role that representational mental states fulfill the latter. It’s because our representational mental states tell us what the world is like (and what it must be like in order for our goals to be met) that we act in the world—particularly in a way that makes sense. Put another way: it’s because of the semantic profiles of our representational mental states that we navigate our way in the world in the successful way we do.

Now, you may wonder why I took you on this long ride. Isn’t the description of events I just gave an obvious one, and aren’t the conclusions I drew from it trivial? Obviously, one forms beliefs about the world, and then uses them to change one’s attitudes (form new ones, get rid of old ones) and act. This is the intuitive picture of how representational mental states work,

¹ And desires, but those are left mostly implied in the passage above.

and I think that this intuitive picture is correct. The intuitive picture, however, it is not uncontentious. In this paper, I argue that some prominent metasemantic views are incompatible with it. That is, I argue that if some views about what makes it the case that our representational mental states have the contents that they have are correct, then the intuitive picture just sketched must be mistaken. Since this picture shouldn't be given up, those views must be.

The views I have in mind are *functionalist* metasemantic views. A metasemantic view is functionalist, in my sense, if it takes the (explanatory) role of a representational mental state to (at least) partly determine the content of the state. In this paper, I'll focus on two functionalist views: Conceptual Role Semantics (henceforth 'CRS') and Interpretivism.² According to CRS, a representational mental state has the content that it does in virtue of its role in our mental lives. Hence, according to CRS, the content of a mental state is partly determined by what that mental state brings about (or is disposed to bring about).³ According to Interpretivism, a representational mental state has the content that it does in virtue of the fact that the interpretation that best rationalizes the subject's dispositions to act in light of the subject's experiences maps the relevant mental state to that content. Hence, according to Interpretivism, the content of a mental state is partly determined by the dispositions that are usually taken to be brought about by that very mental state.

The challenge, put roughly, is this: we should reject functionalist metasemantic views because they reverse the order of explanations. Explanations (even partial ones) are asymmetric. Call this 'the Euthyphro Principle'.⁴ From this principle it follows that if it is the case that some representational mental state has its content partly because it plays a certain role in our mental lives, it cannot play that role partly because it has the content that it does. Applied to our example, this means that if Ella has the belief that it is Wednesday partly because this belief causes her to infer that there's a talk that day, it cannot be that she is caused to infer that there's a talk that day partly because she believes that it is Wednesday. Similarly, if Ella has

² Functionalist metasemantic views in the relevant sense include at least those of Lewis (1974), Harman (1987), Block (1986), Peacocke (1992), Wedgwood (2001), Schwitzgebel (2002), Greenberg (2005), Williams (2020), Chalmers (2021), Pautz (2021), and maybe also Dickie (2015).

³ Notice that CRS need not be committed to the claim that content is determined by *everything* the mental state that has this content brings about. One can think, for example, that only some implications of the relevant mental states are to be among those that determine their contents—in particular, only some of the implications that have to do with other representational mental states. This distinction doesn't affect the main point of the paper; one can just substitute the examples I'm using with one's favorite example of content-determining implications and run the argument *mutatis mutandis*.

⁴ Interestingly, Bridges (2006) argues that Informational Semantics is susceptible to a different Euthyphro challenge. The core idea of Informational Semantics, according to Bridges, is that "the contents of items in the head are constituted by the nomic (i.e., lawful) relations these items bear to elements in the surrounding world" (p. 523). Bridges' challenge to this view relies on the following principle: "[i]f it follows from a true constitutive claim that all F's are g's, then it's not an implemented law that all F's are g's" (p. 530). This principle is quite different from the one I draw from the Platonic dialogue, and it's one I am not convinced we should endorse.

the belief that it is Wednesday partly because she's rational in inferring that there's a talk that day, it cannot be that she's rational in inferring thusly partly because she believes that it is Wednesday. So, if CRS or Interpretivism are correct, then the intuitive picture is mistaken. Since the latter isn't, the former must be.

We can put my Euthyphro-style argument as follows:

1. If CRS or Interpretivism are true, then representational mental states have the contents that they do partly because they play certain explanatory roles.
2. Representational mental states play these explanatory roles partly because they have the contents that they do.
3. The Euthyphro Principle: For all facts A and B , if A partly because B , then it is not the case that B partly because A .

Therefore,

4. CRS and Interpretivism are false.

Notice that I use the slightly technical operator 'partly because' rather than the more colloquial 'because'. The purpose of this is to emphasize that for the relevant claims to be true, the explanation need not be full. ' A partly because B ' is true whenever the fact that B holds is one of the facts in virtue of which the fact that A holds, or one of the facts that explain why A holds. Alternatively, ' A partly because B ' is true whenever the fact that B plays some role in bringing about the fact that A holds, where that 'bringing about' expresses an ontic explanatory relation, such as grounding or causation.

Notice further that although the paper focuses on CRS and Interpretivism, the challenge generalizes. It applies to every view that takes something like CRS or Interpretivism to be a part of the metasemantic story even if not the whole of it. This means that this challenge, if taken seriously, should make us reconsider a big chunk of our metasemantic project.

The source of this challenge can be traced back to the origins of functionalism. Functionalism—originally meant to give an encompassing theory of the mind, but today most popular with respect to the representational features (attitude and content) of representational mental states—is a cousin of behaviorism.⁵ One problem with behaviorism is its failure to account for the causal and explanatory role of mental states. If mental states are to be identified with actions and dispositions to act, how can having those mental states cause those very actions and explain those very dispositions to act? Behaviorism makes mental states token-epiphenomenal with respect to our actions and dispositions to act: having a token mental state cannot be that which brings about a token action/disposition if they are the exact same state. Functionalism is supposed to improve on behaviorism by allowing mental states

⁵ See, for example, Lewis (1966) and Block (1978).

“to be something real and so to be the effects of their occasions and the causes of their manifestations, as common opinion supposes them to be” (Lewis 1966, p. 21).

However, functionalism is too close to its behaviorist origins to be of help here. Although it doesn't face the exact same problem behaviorism faced—mental states are, under functionalism, distinct from actions and dispositions to act—it still faces a fairly similar problem. The problem is this: functionalism is so close to its origins that it allows mental states to cause actions and explain dispositions to act only at the cost of not allowing them to cause anything in virtue of having their mental properties. That is, mental states end up being type-epiphenomenal with respect to our actions and dispositions to act: the states cause the relevant actions and explain the relevant dispositions to act but not in virtue of their mental properties.⁶ Functionalism lets mental states be distinct from actions and dispositions to act, but those are still *determined* by actions and dispositions to act, which makes it the case that the latter cannot be themselves determined—causally or otherwise—by the former. Applied to our case, my accusation amounts to the following claim: functionalist metasemantic theories allow representational mental states to cause actions and dispositions to act only at the cost of not allowing them to cause anything in virtue of their representational properties.⁷

And indeed, functionalism has been accused by many of robbing mental states of their causal efficacy. In particular, it's been accused of robbing content of its causal efficacy. The two best-known challenges are the causal exclusion challenge, originating in Jaegwon Kim's (1989), and the 'dormitive virtue' challenge, originating in Ned Block's (1989) and argued for at length in Robert Cummins' (1992). The most important difference between the Euthyphro challenge and these challenges is this: the other challenges argue that two explanatory relations *pointing in the same direction* are incompatible (or redundant). For example, they rely on the claim that x cannot explain y both causally and conceptually.⁸ The Euthyphro challenge, however,

⁶ This epiphenomenalism is one introduced by McLaughlin (1989).

⁷ Interpretivism isn't a functionalist view in the more traditional sense, but it does seem to have its roots in behaviorism as well. See, for example, Williams (2020)'s account of the history of Interpretivism as starting with Quine (1960).

⁸ According to the exclusion challenge, functional properties cannot have causal powers because their causal efficacy is excluded by the causal efficacy of their realizers. Since the effect is 'already' caused by the realizer, there's no causal work left for the functional property to do. According to the dormitive virtue challenge, there's a conceptual/logical/metaphysically necessary connection between functional properties and the relevant effects, and so there's no need to posit a causal connection as well. Or, in Cummins' version, since CRS identifies the content of a mental representation with its consequences, the only kind of explanations of behavior that appeals to content has the form of 'E resulted because the occurrence of C has the consequence that E results' (where what comes after the 'because' is supposed to be the definition of content according to CRS), which are "evidently empty" (Cummins 1992, 116).

relies on the idea that two explanatory relations *pointing in opposite directions* are incompatible, i.e., that nothing can both explain and be explained by the same thing.⁹

Euthyphro-style challenges to CRS have been articulated by Fred Dretske (1981), Jerry Fodor (2004) and Paul Boghossian (2014). Like the challenge developed here, these objections push the thought that CRS is incompatible with content having causal efficacy, because it implies that content itself is explained by what it is supposed to cause. I see my challenge as a better version of these arguments, one which applies to a much wider range of views (including Interpretivism and other functionalist views) and uses more plausible premises.¹⁰ My challenge shows that functionalism—of all varieties—is, at its core, incompatible with the explanatory (not just causal) efficacy of content. And it does that by appealing only to a very general and highly plausible feature of explanations: asymmetry.

The paper proceeds as follows. In section I, I present Interpretivism and (what I view as) the most plausible version of CRS. I argue that both theories imply certain ‘partly because’ claims, thereby securing premise 1 of the argument presented in page 4. In section II, I present and motivate the Euthyphro Principle. In section III, I make the intuitive picture presented above more precise. In particular, I claim that the intuitive picture implies opposing ‘partly because’ claims, thereby securing premise 2 of the argument. I then derive a conflict between those claims and the ones implied by CRS and Interpretivism using the Euthyphro Principle. In section IV, I consider several answers to this challenge. Section V concludes by considering the use of the Euthyphro challenge beyond metasemantics.

I Functionalist Metasemantic Views

Metasemantic theories purport to tell us in virtue of what representational mental states have the contents that they do. They purport to tell us, for example, in virtue of what Ella believes

⁹ Other philosophers have noticed that popular views in different domains in philosophy reverse the order of explanation in the way pointed out here. See Kripke (1982) and Strawson (2010).

¹⁰ Dretske’s challenge relies on the claim that the consequences of representational mental states to which what he calls ‘consequentialist’ metasemantic view appeals can only explain content if they are themselves contentful, and that they can only be contentful if they derive their contents from the representational mental states that cause them. Regardless of the merits of this claim, this is a further assumption that goes beyond what is needed, according to the Euthyphro challenge, to derive the incompatibility of functionalism with the explanatory efficacy of content.

Fodor’s challenge to CRS relies on a version of CRS according to which a possession of a concept C is dependent upon grasping some rules of inference, and argues that such grasping must itself depend on the individual already possessing C. As noted by Harman and Greenberg (2005), CRS need not commit itself to anything about grasping a rule, as is needed for Fodor’s argument to work—a disposition to infer according to the rule is all that the most plausible versions of CRS presuppose.

Boghossian raises his challenge as a short corollary of his view of inference. Accordingly, the challenge is hostage to the particular view developed in that paper.

that today is Wednesday. *Functionalist* metasemantic theories—the kind of metasemantic views we’re interested in in this paper—are ones that take representational mental states to have the contents that they do partly in virtue of facts about what we usually take to be their consequences, i.e. the kinds of things these states bring about, are disposed to bring about, rationalize, etc. Recall that representational mental states have two important roles: a semantic role (to tell us what the world is like) and an explanatory role (to cause, ground and rationalize our other mental states, actions and dispositions to act). Functionalist metasemantic views take the explanatory role of representational mental states to be prior to their semantic role. In this way, they can explain the more mysterious features of representational mental states—their aboutness or intentionality—with more familiar terms—their relations to actions and dispositions to act.

There are many ways to develop the idea behind functionalist metasemantics. Here, I’m going to focus on two: Interpretivism, and what I take to be the most plausible version of CRS. This version of CRS is one focused on the descriptive dispositions of the representational mental states themselves. Other versions of CRS might focus on other features of the relevant states, such as their actual effects (the actions and dispositions of the subject) and their actually causing those effects, or normative variants of those features. I focus on the former for three reasons. First, it’s suggested by the works of proponents of CRS, especially when they try to explicate their picture. Second, it fits better with the program in the philosophy of psychology investigating the causal and explanatory role of mental representations and their content. Finally, and most importantly for our purposes, this version is the hardest to reject using the Euthyphro challenge. Explanatory claims about the descriptive dispositions of states are the ones that have the least intuitive appeal, making it harder to establish that the explanatory claims needed for my arguments are indeed true. Therefore, the paper focuses on the inadequacy of the dispositional version of CRS. I consider the generalization of the argument to other versions of CRS at the end of section III.

In this section, I argue that Interpretivism and the chosen version of CRS each imply a certain ‘partly because’ claim about our subject Ella.

Starting with Interpretivism, the view is committed to the following claim:

(INT) Ella believes that today is Wednesday partly because she is disposed to check her schedule in circumstances C.¹¹

¹¹ Notice that not only proponents of Interpretivism might be committed to this ‘partly because’ claim, but plausibly also proponents of simpler versions of CRS. Of the latter, many might like to talk about the dispositions of the subject *to use a particular mental symbol* (see, for example, Greenberg 2005). My example of a disposition to check the schedule should be read as a mere example. Whatever the proponent of CRS takes as the facts that explain why Ella believes that today is Wednesday would be susceptible to the Euthyphro challenge.

Here's why. According to Interpretivism, a representational mental state has the content that it does in virtue of the fact that the correct interpretation of the relevant subject maps the relevant representational mental state to that content. Taking an interpretation to be a mapping between a subject's internal states and contents (or alternatively as a mapping from internal states to ordered pair of attitudes and contents), the correct interpretation (i.e. the interpretation that determines the contents of the subject's representational mental states) is the one that best rationalizes the subject's dispositions to act in light of the subject's experiences.¹² Going back to our example, Ella believes that today is Wednesday because the interpretation that best rationalizes her dispositions to act given her experiences maps an internal state of hers to the pair <belief, the proposition that today is Wednesday>. One of her dispositions that is best rationalized by this interpretation is her disposition to check her schedule. Hence Ella believes as she does partly because she has the relevant disposition, which is exactly what (INT) says.¹³

As for CRS, the view is often characterized by its proponents in the very terms I used to characterize functionalist metasemantic theories more generally. That is, according to CRS, a representational mental state has the content that it does in virtue of its role in our mental lives.¹⁴ 'Role' can be understood in a few different ways, but the most common way to flesh out the notion is by understanding it as dispositional and descriptive. That is, the role of a representational mental state is what it is disposed to be caused by and what it is disposed to cause.¹⁵ In what follows, I focus only on the dispositional *effects* of the relevant representational mental state, namely, what they are disposed to cause or are rationally apt to cause.

If one accepts CRS in its dispositional and descriptive form, one takes the role of representational mental states to be *explanatorily prior* to what they mean. See, for example, Mark Greenberg and Gilbert Harman (2006):

Conceptual role semantics (CRS) is the view that the meanings of expressions of a language (or other symbol system) or the contents of mental states are determined or explained by the role of the expressions or mental states in thinking. The theory can be taken to be applicable to language in the ordinary sense, to mental representations, conceived of either as symbols in a 'language of thought' or as mental states such as beliefs, or to certain other sorts of symbol systems. CRS rejects the competing

¹² See Lewis (1974), Williams (2020), Pautz (2021).

¹³ Notice that the Interpretivist is not committed to speaking in terms of internal states being mapped to ordered pairs of attitudes and contents; she can instead talk of mappings of *subjects at times* to pairs of attitudes and contents. This difference doesn't matter much for our purposes.

¹⁴ Some philosophers (for example, Block 1986) distinguish between 'narrow' and 'broad' content, and among those, some only take the former to be determined by role. However, this debate isn't relevant for our purposes here, as even those who think that role determines narrow content would think that it *partly* determines broad content (together with, say, causal relations to the environment, or some such), in which case they should still accept the 'partly because' claims to follow, and are therefore part of my target in this paper.

¹⁵ See, for example, Block (1986) and Chalmers (2021). The view has a normative counterpart. See, for example, Wedgwood (2001) and Greenberg (2005).

idea that thoughts have intrinsic content that is prior to the use of concepts in thought. According to CRS, meaning and content derive from use, not the other way round.

Note that some proponents of CRS use the notion of ‘identification’¹⁶ rather than notions that more easily suggest explanatory priority, such as ‘explanation’ or ‘determination’. That is, some proponents of CRS would like to say that the fact that a subject has a belief that *p* is *identical* to the fact that the belief plays a certain role in the subject’s psychology. See, for example, Ned Block (1998):

According to Conceptual Role Semantics ("CRS"), the meaning of a representation is the role of that representation in the cognitive life of the agent, e.g. in perception, thought and decision-making. It is an extension of the well known "use" theory of meaning, according to which the meaning of a word is its use in communication and more generally, in social interaction. CRS supplements external use by including the role of a symbol inside a computer or a brain. The uses appealed to are not just actual, but also counterfactual: not only what effects a thought does have, but what effects it would have had if stimuli or other states had differed.¹⁷

I think that this difference is mostly irrelevant for our purposes, for three reasons. First, proponents of CRS who use the identification locution also describe their views in terms that suggest explanatory priority. For example, Block, in the very same entry, explains that CRS is engaged in the project of metaphysical (rather than linguistic) semantics, which is ‘one of investigating the fundamental nature of meaning, especially what it is about a person that gives their words or thoughts whatever meanings they have in the first place’. It’s natural to assume that the facts that give thoughts their meanings are facts that determine their meanings. Second, identifications plausibly imply claims of explanatory priority. Consider: if to be a vixen just is to be a female fox, this plausibly implies that Vixi is a vixen partly in virtue of the fact that she’s a fox and partly in virtue of the fact that she’s female.¹⁸ That is, the fact that Vixi is a fox and the fact that she is female are explanatorily prior to the fact that she is a vixen. Finally, the challenge of this paper can be modified to apply to a view given purely in terms of identification, as I explain in section III. For these reasons, I stick to talking in terms of explanatory priority throughout.

So proponents of CRS take the role of representational mental states to be explanatorily prior to their contents. Given that we’re interested in ‘role’ in its dispositional version, this means that this version of CRS is committed to the following claim:

¹⁶ Where ‘identifications’ are claims of the form ‘to be F is to be G’. See Dorr (2016).

¹⁷ Notice that ‘representation’ can mean either *concept* or *state*, and indeed some proponents of CRS want to speak about the determination of the content of concepts as prior to that of states. This difference doesn’t matter much for our purposes, as everything said in this paper can be translated to concept talk.

¹⁸ Dorr (2016) considers this maneuver as a way to avoid some seemingly counterintuitive implications of his characterization of identifications.

(CRS) Ella believes that today is Wednesday partly because she is in a state that is disposed to cause her to check the schedule in circumstances C.

For our purposes, the main difference between Interpretivism and the relevant version of CRS is this: while the Interpretivist appeals to *subject-level effects* to ground content, the proponent of our chosen version of CRS appeals to *state-level effects*. That is, according to (INT), it's something about the dispositions *of the subject* that explains why she believes that today is Wednesday, while according to (CRS), it's something about the dispositions *of the state* the subject is in that explains that. While this makes the argument against the Interpretivist somewhat more straightforward, both views, I claim, are susceptible to the Euthyphro challenge, and for very similar reasons. In the next section, I precisify the Euthyphro Principle on which the challenge relies and argue for its truth.

II The Euthyphro Principle

“Is the pious loved by the gods because it is pious, or is it pious because it is loved by the gods?” (*Euthyphro*, 10a)

One natural lesson to draw from Plato's *Euthyphro* is this: explanatory relations, those captured by 'because' claims, are asymmetric. Consider Euthyphro's claim that the pious acts are pious because the gods love them. While arguing against this proposal (or, rather, bringing Euthyphro to see that this claim can't possibly be true), Socrates proposes a competing 'because' claim, according to which the gods love the pious acts because they are pious. This claim is presented as a strict alternative: Euthyphro cannot accept at once both his original proposal and Socrates' competing claim. Thus, when Euthyphro comes to realize he indeed accepts Socrates' claim, he drops his original proposal.

That these two claims cannot be true together doesn't seem to be the case due to a special feature of the particular claims made in the context of the dialogue. Rather, the two speakers seem to be relying on a more general principle, one mentioned already in the introduction to this paper:

The Euthyphro Principle (EP): For all facts A and B , if A partly because B , then it is not the case that B partly because A .

The original dialogue uses the operator 'because' rather than 'partly because', which may suggest that the considerations raised by the dialogue support only the weaker claim that *full* explanations are asymmetric. In what follows, I motivate the claim that those considerations can be modified to support the stronger claim, namely, that *partial* explanations are asymmetric.

To see why, consider the relation holding between one god's love and the pious acts if Euthyphro's original proposal was indeed true. Let's take Zeus as our example. If it is true that the pious acts are pious fully because the gods love them, then it seems to follow that, given that Zeus is one of those gods, the pious acts are pious *partly because* Zeus loves them.¹⁹ This is because the fact that Zeus loves those acts is one of the facts that help make it the case that those acts are pious—one of the facts in virtue of which the latter fact holds.

Furthermore, it's reasonable to think that Socrates' use of 'because' in fact means *partly because*. Presumably, each god that loves the pious acts doesn't love them just because they're pious, but also because they know them to be so, or believe them to be so, or something similar. Hence, the 'because' in Socrates' claim is better understood as meaning the same as 'partly because': an act's being pious doesn't *fully* explain the gods' love for it. Applied to the case of Zeus, it follows that Zeus loves the pious acts partly because they're pious.

Now, consider the two claims we just arrived at: that the pious acts are pious *partly because* Zeus loves them, and that Zeus loves the pious acts partly because they're pious. I claim that just like our two original claims, these two claims can't be true together. Moreover, they cannot be true at the same time for what seems to be basically the same reason: explanations, whether full or partial, are asymmetric. And so I conclude that the real principle guiding our intuitions in the Euthyphro case is indeed EP.

Note the generality of EP. It's quite plausible to think that the two 'partly because' connectives are backed by different explanatory relations: Euthyphro's proposal sounds like a grounding claim, while Socrates' sounds like a causal claim. According to EP, there couldn't be two ontic explanatory relations running in opposite directions between the same two facts. It's not only that grounding itself is asymmetric, and causation itself is asymmetric, and whatever other ontic explanatory relations are out there are themselves each asymmetric; rather, it's also that the disjunction of all the ontic explanatory relations is asymmetric. This thought fits well with a more general picture of ontic explanatory relations, according to which they all belong to a more general kind, and in virtue of belonging to this kind they must obey some common principles.²⁰

To summarize the lesson we've just learned from Plato's Euthyphro: the disjunction of ontic explanatory relations (i.e. those captured by 'because' claims) is asymmetric. This lesson is

¹⁹ If we take the relation captured by Euthyphro's 'because' claim to be that of grounding—which seems to be the most natural reading—then this claim just follows from some general facts about the transitivity of grounding and the grounding of generalizations by their instances.

²⁰ Notice that someone who, like Loewer (2012), wants to claim that each explanatory relation is asymmetric, but the disjunctions of explanatory relations isn't, will need to explain what drives our intuitions in the original Euthyphro case.

further supported by other theoretical considerations. First, if we think that the operator ‘partly because’ is irreflexive and transitive, then its asymmetry follows for free.²¹

Second, when we consider explanations more abstractly, the following seems to be a natural starting point: if facts are related by the explanatory relation signified by ‘partly because’, then one of the facts is explanatorily prior to the other. Put another way, it’s in virtue of the explanans already being the case that the explanandum holds, as the explanandum is determined by the explanans and depends on it. Hence, the relation between those two facts must be asymmetric. Furthermore, EP can be seen as encapsulating a kind of ‘no bootstrapping’ constraint on reality: if fact A and B are codetermined in the way forbidden by EP—that is, if each of them requires the holding of the other—it’s unclear how either of them can begin to exist. Compare: if I went back in time using a time machine and gave my earlier self the plans to build that very machine, how do the plans enter into existence? Now, one might be tempted to give up EP and the rationale just presented in very weird worlds: maybe in a world where time travel is possible, we would have to allow for some facts to just ‘pop into existence’, as it were, and for an event to cause that which causes it. Be that as it may, we have no reason to think that our world is weird in this way, and so these cases don’t provide a reason to think that EP is false in our world and worlds like ours, and definitely not that it fails occasionally in our world. And if one wanted to avoid the Euthyphro challenge by rejecting EP, one would have to take it to fail quite often.²²

In the next section, I use EP to argue that Interpretivism and the dispositional version of CRS conflict with the intuitive picture roughly described in the introduction. These two theories, I claim, should therefore be rejected.

III The Euthyphro Challenge

Consider what is involved in taking the intuitive picture presented in the beginning of this paper seriously. Recall, the picture says that representational mental states have two roles: a semantic role and an explanatory role. The semantic role is that of informing us of what the world is like and what it could or should be. The explanatory role is that of explaining—causing, grounding and rationalizing, at the very least—other mental states as well actions and dispositions to act. Furthermore, the picture makes a claim about how these two roles interact: it’s in virtue of fulfilling the former role that representational mental states have the ability to

²¹ Although philosophers have had some doubts about each of these characteristics (See, for example, Schaffer (2012)), those doubts seem to be relevant only to very particular kind of cases, and so shouldn’t affect what one thinks about their application here.

²² For a treatment of *prima facie* counterexamples to the asymmetry of ‘because’, see Schnieder (2015).

fulfill the latter role. That is, the explanatory role of representational mental states depends on their semantic role.

In our opening story, Ella looks at the schedule of talks given on Wednesday. Why does she do that? Because she believes that it's Wednesday and therefore that there will be a talk that evening, and she desires to know what it will be about. If it wasn't for the relevant semantic features of her belief that it's Wednesday, say, the belief wouldn't have caused our subject to look at the schedule. Furthermore, checking the schedule is a rational thing for Ella to do. Why is that? Again, whatever story we tell must invoke the information the subject has about the world together with her desires. If she didn't believe that there will be a talk that night, or if she didn't want to know what the talk will be about, then she wouldn't have a reason to check out the schedule, and her action (if she would have still performed it) wouldn't be rational.

Notice further that we don't only think of the semantic role of representational mental states as explaining their actual consequences—what actions, dispositions to act and other mental states they in fact cause, ground or rationalize—but also as explaining their power for bringing about these consequences. It's in virtue of having certain semantic features that representational mental states have the causal and rational powers that they have, where those should be understood as powers to cause and powers to rationalize, respectively. If Ella's belief didn't have the content that *today is Wednesday*, the belief wouldn't be disposed to cause her to check the schedule, whether or not she in fact checks the schedule. The same goes for the rational powers of that same belief: it's in virtue of its content that the belief can rationalize checking the schedule, whether that action is in fact performed and is in fact thus rationalized.

This way of conceiving of the connection between the semantic and the explanatory roles of representational mental states runs deep. It's not clear what other picture of how we act and reason we could have that would respect our claim to rationality (but more on that in the next section). And it's not true only of acting and reasoning in everyday contexts—the same is true for more sophisticated kinds of reasoning, such as scientific and mathematical reasoning. When a biologist decides to run another experiment, he does so on the basis of the information he has about how previous experiments resulted, information 'stored' in his beliefs. When a mathematician infers that some theorem is true on the basis of some axioms, she does so because she believes that the latter are true, and sees that the theorem follows from them given inference rules she accepts. In both cases, facts about the semantic features of the subject's beliefs are part of the explanation of the consequences of those beliefs.²³

²³ This intuitive picture is advocated, to different degrees, by Dretske (1981) and Fodor (1987, 1989). Dretske urged us to accept the claim that 'it is the *content* of our beliefs, *what* we believe, that shapes our behavior (what

If what I said above is true, then what the intuitive picture of representational mental states commits us to is a list of ‘partly because’ claims of the following sort:

(Subject Because) Ella is disposed to check her schedule in circumstances C partly because she believes that today is Wednesday.

(Subject Rational Because) Ella is rational in being disposed to check her schedule in circumstances C partly because she believes that today is Wednesday.

(State Because) Ella is in a state that is disposed to cause her to check the schedule in circumstances C partly because she believes that today is Wednesday.

(State Rational Because) Ella is in a state that is rationally apt to cause her to check the schedule in circumstances C partly because she believes that today is Wednesday.²⁴

Now, there’s an important difference between the first two claims and the last two claims. Obviously, the former are more intuitive than the latter; we more generally are more committed to claims about what explains the dispositions of *subjects* rather than those of *states*. Nevertheless, I think we have a strong reason to accept both (State Because) and (State Rational Because), even if we don’t have direct intuitions about them. The reason is this: the robust intuitions that lead us to accept (Subject Because) and (Subject Rational Because) are explained by the more general picture, and that picture implies (State Because) and (State Rational Because) as well. The intuitive picture, which explains and generalizes our intuitions about the subject’s actions and dispositions to act, is not just one according to which representational mental states bring about actions, etc.; it is also a picture on which content plays a role in explaining why those mental states have the relevant powers. If talk of a representational mental state’s dispositions is understood in terms of that state’s having the causal powers that it does, then (State Because) and (State Rational Because) are very much a part of our intuitive picture of the role of representational mental states in our lives.

The intuitive picture hasn’t seemed mandatory to everyone. There have been many attempts to understand the explanatory role of content in a less demanding way than it suggests. One well-known way of developing this line is due to Ernest LePore and Barry Loewer (1987). Another is due to Frank Jackson and Philip Pettit (1990). According to LePore and Loewer,

we do), and we want this fact reflected in the causal efficacy of semantic structures. [...] To qualify as a cognitive structure, therefore, an internal state must not only *have* a semantic content, it must *be* this content that defines the structure’s causal influence on output” (1981, pp. 199-200). Fodor famously claimed that “if it isn’t literally true that my wanting is causally responsible for my reaching, and my itching is causally responsible for my scratching, and my believing is causally responsible for my saying... if none of that is literally true, then practically everything I believe about anything is false and it’s the end of the world” (1987, p. 77).

²⁴ (Subject Rational Because) and (State Rational Because) don’t directly conflict with Interpretivism and the version of CRS the paper focused on, but fleshing them out explicitly will prove useful at the end of this section, as well as in the next section.

content can play an explanatory role in a counterfactual fashion—if Ella didn’t believe that today is Wednesday, she wouldn’t have been disposed to check her calendar. So the content of Ella’s beliefs plays an explanatory role, although this role doesn’t vindicate the intuitive picture. According to Jackson and Pettit, content can play an explanatory role by providing “program explanations”—the fact that Ella believes that today is Wednesday doesn’t produce her disposition to check her calendar, but it indicates that there is an underlying fact that does produce this disposition. Here, again, we have an attempt to provide content with an explanatory role that doesn’t vindicate the intuitive picture.

I think that these attempts to conceive of the explanatory role of content in less demanding ways fail. The problem is that they make playing an explanatory role too easy to come by. To see this, consider a world where all and only Christmas songs are sung at a pitch high enough to break glass. Let’s further assume that this regularity is due to some robust social convention. Both accounts imply that in that world the fact that something is a Christmas song plays an explanatory role vis-à-vis the breaking of glasses: if the song wasn’t a Christmas song, the glass wouldn’t have been broken (other songs are sung at pitches that just don’t lead to the breaking of glasses), and the fact that the song is a Christmas song does indicate that there is an underlying fact that produces the breaking of the glass.

Given the results of the previous sections, it’s easy to derive a conflict between CRS and Interpretationism, on the one hand, and the intuitive picture, on the other, using EP. Recall that Interpretivism and CRS imply the following ‘partly because’ claims:

(INT) Ella believes that today is Wednesday partly because she is disposed to check her schedule in circumstances C.

(CRS) Ella believes that today is Wednesday partly because she is in a state that is disposed to cause her to check the schedule in circumstances C.

We can now present more fleshed out versions of the argument given in the introduction, one for Interpretivism and one for CRS:

The Euthyphro Argument against Interpretivism

1. If Interpretivism is true, then (INT) Ella believes that today is Wednesday partly because she is in a state that is disposed to cause her to check the schedule in circumstances C.
2. (Subject Because) Ella is disposed to check her schedule in circumstances C partly because she believes that today is Wednesday.
3. EP: For all facts *A* and *B*, if *A* partly because *B*, then it is not the case that *B* partly because *A*.

Hence,

4. Interpretivism is false.

The Euthyphro Argument against CRS

1. If CRS is true, then (CRS) Ella believes that today is Wednesday partly because she is disposed to check her schedule in circumstances C.
2. (State Because) Ella is in a state that is disposed to cause her to check the schedule in circumstances C partly because she believes that today is Wednesday.
3. EP: For all facts A and B , if A partly because B , then it is not the case that B partly because A .

Hence,

4. CRS is false.

This is the Euthyphro challenge to CRS and Interpretivism. According to the challenge, these views switch the order of explanation: they claim that it's in virtue of fulfilling their explanatory role that representational mental states fulfill their semantic role, when it's in fact the other way around. Note that the challenge applies not only to theories that take CRS or Interpretivism to be the whole story about how content is determined, but also to theories that have a functionalist component. Provided that a theory takes facts about the contents of representational mental states to be partly determined by the role of those states in our mental lives, it conflicts with the claim that the determination holds in the other direction.

I have thus far limited my attention to a particular version of CRS. I now want to briefly mention how the argument can be generalized to target other versions of CRS. Simple versions of CRS that take content to be explained by the actions or dispositions it in fact produces face pretty much the same problem as Interpretivism. Versions of CRS that take content to be explained by the fact that the state (of which it is the content) causes certain actions or dispositions will conflict both with the intuitive claim that the state causes what it does in virtue of its content and with (State Because), given some plausible assumptions about the connection between dispositions to cause and causing.²⁵ Normative versions of CRS that

²⁵ Usually what explains an object having a disposition to cause something will be part of what explains why some event involving that object caused that something. Now, it could be that certain dispositions are sometimes 'blocked' by other dispositions—maybe an object O has two different dispositions that lead to the same result, D_1 and D_2 , but D_2 only kicks in when D_1 doesn't, and it so happens that D_1 always does. So whatever feature in virtue of which O has D_2 is never one in virtue of which events involving O bring about the common effect of D_1 and D_2 . But even if this is sometimes the case, there's no reason to think this happens in every case of mental causation.

take the role of content to be explicated in terms of what it rationalizes or is rationally apt to cause²⁶ will conflict with either (Subject Rational Because) or (State Rational Because).

Finally, in section I, I mentioned the identification version of CRS, and claimed that even if one rejects the claim that the identification version of CRS implies any ‘partly because’ claim, we may still run a version of my argument against the view. The argument goes as follows: the fact that explanations are asymmetric entails the fact that they are also irreflexive. That is, for all facts A and B , if A partly because B , then it’s not the case that A is identical to B . It follows that identifications of the form ‘for Ella to believe that Today is Wednesday is for Ella to be in a state that is disposed to cause her to check her schedule in circumstances C ’ are also in conflict with the kind of explanatory priority claims implied by the intuitive pictures. The intuitive picture would imply a ‘partly because’ claim of the form ‘Ella is in a state disposed to cause her to check the schedule in circumstances C partly because she believes that today is Wednesday’, and that, together with the irreflexivity of explanations, would entail the falsity of the identification made by the proponent of the version of CRS under consideration. In a sense, the failure of this version of CRS is even more similar to the failure of behaviorism, as it identifies content with what it is supposed to explain.

In conclusion, Interpretationism and CRS conflict with the intuitive picture of the role of representational mental states in our lives. In the next section, I consider what the proponents of CRS and Interpretivism can say in reply.

IV Ways Out?

The most natural way out of the Euthyphro challenge involves denying at least some part of the intuitive picture. In what follows, I first argue that there’s a part of the intuitive picture that nobody should deny—namely, the rationalizing explanatory role of content. Second, I consider a rejection of the descriptive explanatory role of content, and argue that this part of the intuitive picture is implied by its rationalizing part. Third, I argue that a more circumscribed rejection of some descriptive and rationalizing claims about content won’t work either. Finally, I claim that an important motivation for CRS conflicts with giving up on the descriptive part of the intuitive picture.

Let’s start. I think that nobody should reject the rationalizing explanatory role of content as specified by the intuitive picture. There are two reasons for this. First, if the correct interpretation is the one that *best rationalizes* the subject’s dispositions to act in light of their

²⁶ See, for example, Wedgwood (2001) and Greenberg (2005).

experiences, then the Interpretivist is committed to the rationalizing role of content. Content, according to the Interpretivist, is the thing in virtue of which representational mental states are apt to rationalize the subject's dispositions. Second (and more importantly), it's not clear what it would be for a transition to be rational if content cannot be appealed to.²⁷ Think of what is left to explain rationality if content is not to be invoked: facts about the states' neuronal and syntactic features. The former is a non-starter: what one's neuronal state is like is not, all by itself, a fact that explains why some mental state or action makes sense or is reasonable. As for the latter, we can talk about validity in a way that appeals to form only. However, even then the form is understood in terms of the contents of the connectives, and any number of syntactic systems can be devised such that reasoning in accordance with them would be a rational disaster. Just because something can be formalized does not yet mean that it should be used. It's only by connecting the formalism to content that we have a reason to use some syntactic system rather than another, that moving from one syntactic object to another makes sense.²⁸

If this is correct, everyone should accept the following claims of the intuitive picture:

(Subject Rational Because) Ella is rational in being disposed to check her schedule in circumstances C partly because she believes that today is Wednesday.

(State Rational Because) Ella is in a state that is rationally apt to cause her to check the schedule in circumstances C partly because she believes that today is Wednesday.

We're now moving to the second part of this section. Notice that (Subject Rational Because) and (State Rational Because) each imply respectively:

(Subject Because) Ella is disposed to check her schedule in circumstances C partly because she believes that today is Wednesday.

(State Because) Ella is in a state that is disposed to cause her to check the schedule in circumstances C partly because she believes that today is Wednesday.

This is because rationalization implies explanation: for a feature to have rational powers, it must have explanatory powers in the metaphysical sense as well, and for particular transitions to be rationalized by some feature, that feature must play an explanatory role in bringing those transitions about. In arguing for this claim, I'm going to talk in terms of causal efficacy (rather than explanatory efficacy). I'll do this for two reasons. First, it's easier to work with a concrete

²⁷ Chalmers (2021) makes the same point.

²⁸ The claim isn't that there couldn't be any level of explanation of rational mental transitions that doesn't mention content. Anyone who gives an account of content in content-free terms must agree that there are such levels of explanation. What's important is that there's a level of explanation—maybe even a direct level of explanation—of rational mental transitions that appeal to content.

example of an ontic explanatory relation. Second, it is plausible that many of the relevant ontic explanatory relations holding in the ‘partly because’ claims of the intuitive picture are causal ones.

Why should we think that rationalization implies causation?²⁹ Consider the following version of our case: Ella has the belief that today is Wednesday just as before, and is disposed to check her schedule just as before. This time, however, her belief wasn’t what caused her to check the schedule. Instead, the minute she formed the belief that today is Wednesday, a lightning bolt hit her head, leading her (all by itself) to go check her schedule. In such a case, we wouldn’t accept (Subject Rational Because), as the relevant causal connection is missing. Although Ella might have a good reason to be disposed to act as she does, she is not disposed *for that reason* unless that reason causes her to be disposed in the way she is. If this is correct, then (Subject Rational Because) implies (Subject Because).³⁰

Notice that if what is said above is correct, the Euthyphro challenge is internal to Interpretivism. Recall that Interpretivism is committed to:

(INT) Ella believes that today is Wednesday partly because she is disposed to check her schedule in circumstances C.

But it is also committed to (Subject Rational Because). Here’s why. Interpretivism is committed to the correct interpretation rationalizing dispositions to act—this is what makes it the correct interpretation. Now, it’s reasonable to assume that the extent to which an interpretation rationalizes the subject’s entire set of dispositions to act depends on the extent to which particular assignments of contents rationalize particular dispositions to act. That is, other things being equal, an interpretation is more rationalizing if particular transitions in the subject’s mind are better rationalized. For example, other things being equal, an interpretation that maps the subject to the pair <belief, the proposition that today is Wednesday> is more rationalizing than alternative interpretations partly because believing that today is Wednesday would better rationalize the subject checking the schedule when she wants to know what the topic of that day’s talk is. This kind of assumption—one which moves from the rationalizing force of an interpretation to the rationalizing force of particular representational mental states—is what Williams calls a ‘localizing assumption’. Williams himself uses similar assumptions to derive particular plausible verdicts about content in his (2020). Since (Subject Rational Because) implies (Subject Because), Interpretivism implies (given EP) two conflicting claims: (INT) and (Subject Because).

²⁹ The argument that follows is a variation of the one found in Davidson (1963).

³⁰ The argument would work also for the actions caused by the belief, not just the dispositions brought about by it.

Next, (State Rational Because) implies (State Because). More generally, for some feature to explain rational powers (i.e. the powers to rationalize), it must also explain causal powers (i.e. the powers to cause). To see this, consider two subjects, Gabriel and Garcia. Gabriel is a normal subject whose C-fibers firing realizes pain. Garcia is Gabriel's zombie counterpart: he is a physical, functional and behavioral duplicate of Gabriel, but he has no pain experiences. Now, assume that both Gabriel and Garcia are in the state of C-fibers firing, and that for both of them, this state is disposed to cause them to take pain killers. Since they're functional duplicates, were either of them to take pain killers, the story of why they took pain killers would have been identical. Furthermore, since they're functional duplicates, the following is also true of both: it is not the case that their state of C-fibers firing is disposed to cause them to take pain killers partly because it realizes pain. Since Garcia doesn't experience any pain, it is not the case that his state of C-fibers firing is rationally apt to cause him to take pain killers partly because it realizes pain. Since Gabriel and Garcia share every relevant causal feature in common, it seems like the same must be true of Gabriel as well. The mere presence of pain, as the lightning bolt case showed, isn't enough to make it rationally relevant. Since the only difference between Gabriel and Garcia is that of whether pain is present or not, it cannot be that in one of them pain explains the subject's state's rational powers and in the other it doesn't.

Although this case is quite outlandish, the point is pretty simple and intuitive, and it can be supported by further, less outlandish cases. If what potentially makes it rational for me to believe that the ball in front of my is red is the fact that it is red, I cannot be rational in my belief (nor can my disposition to form the belief be rationality-conferring) unless I'm causally sensitive to the redness of the ball. It wouldn't be enough to be causally sensitive to anything else if I'm not also sensitive to the fact that the ball is red—not even to the fact that the ball is scarlet, which presumably grounds the fact that the ball is red. If, for example, I don't know that scarlet is a shade of red, and my perceptual system doesn't classify it as such, then my sensitivity to the fact that the ball is scarlet does nothing to rationalize my belief. So rationality demands quite a strong connection between the rationalizing feature and the rationalized feature—the latter must be causally (or, more generally, explanatorily) sensitive to the former, and not merely to nearby features, close as they may be to the rationalizing feature.

Of course, this is not a knock-down argument, and my opponent can deny the intuitions needed for the cases above. Be that as it may, if one gives up on the intuitive picture, one is forced to provide a new, more complicated story of how rationalization works—one that would presumably conflict with the literature about acting for a reason and inferential justification, all of which assume the causal (or explanatory) efficacy of content.³¹ I am

³¹ See, for example, Boghossian (2014) and Wright (2014).

skeptical that such a project could be carried out. So holding onto functionalism, while not impossible, is more costly than one might have initially supposed.

Moving forward to the third part of the section: can my opponent reject just part of the descriptive explanatory role that content plays, together with the associated rationalizing role? One rather circumscribed way of doing it is the one suggested by footnote 2: functionalists need not (and do not) claim that *everything* that is intuitively taken to be a consequent of representational mental states plays a role in determining the contents of those states. Consider, for example, a proponent of descriptive CRS that takes only certain dispositions to have other beliefs as partly determining content but doesn't take any dispositions to act as determining content. According to this reply, content doesn't play the entire explanatory role presented in section III, because content will not play an explanatory role vis-à-vis the dispositions to have other beliefs that partly determines it. But that's not too worrying, because content does play an important explanatory role—it can still explain and rationalize our dispositions to act.

I think that this approach faces the following problem. Presumably, this proposal restricts the facts about implications that help determine content to some core ones. Otherwise, we won't get sufficiently determinate content. And if this is the case, then giving up those 'partly because' claims would be quite costly—for these are among the most important implications of representational mental states that we wanted content to explain to begin with. Sure, there are other implications we'd like to explain as well (like dispositions to act, for example), and so this does mark an improvement in the functionalist position. Nonetheless, we're still left with many core implications that are barred from being explained by content.

Suppose you tell me that Vixi is a vixen, and I form the belief that Vixi is a vixen on the basis of your testimony. I also form the disposition to infer that Vixi is female. If part of what determines the fact that my belief has the content that it does is that I'm disposed to infer that Vixi is female, then we cannot explain my disposition to infer that Vixi is female using the fact that I believe that Vixi is a vixen. But this seems like something that we'd very much like content to be able to help us with—explaining why I form the inferential disposition that I do—whether or not I can use content to explain actions, e.g., why I say 'Vixi is a vixen'.

Finally, giving up on part of the intuitive picture seems to me to conflict with one of the main motivations for adopting a view like CRS. Following Lewis (1966),³² we may think of CRS as making progress over its behaviorist ancestor by letting representational mental states be

³² Lewis discusses phenomenal mental states, but similar points apply to representational mental states.

causes of actions and dispositions to act.³³ That is, one reason why we should be functionalists rather than behaviorists about representational mental states is that only according to the former can representational mental states be causes. Instead of identifying representational mental states with some actions and dispositions to act, we take them to be those which cause those actions and dispositions to act. However, for representational mental states to really be the causes of those actions and dispositions to act, the relevant ‘partly because’ claims made by the intuitive picture must be true. If my belief that there’s beer in the fridge plays a causal role in making me disposed to go to the kitchen, then it must be true that I’m disposed to go to the kitchen *partly because* I believe that there’s beer in the fridge. Thus, if functionalists are correct in claiming that their view is one on which representational mental states do play a causal role (and is therefore preferable to behaviorism), functionalism must imply the relevant ‘partly because’ claims of the intuitive picture. Otherwise, representational mental states don’t play any causal role according to functionalism; at most, some correlates of them do. And as Lewis himself agrees, this is just a sophisticated version of epiphenomenalism that runs afoul of the original motivation for becoming a functionalist.³⁴

This sketch of an argument depends on what our best reason to accept CRS is. If one wants to accept the view solely for its extensional adequacy, then this argument doesn’t carry much force. The same goes if one wants to accept CRS for radically different reasons than the one I alluded to. Theorists inclined to accept CRS on the basis of the reasons highlighted above, however, cannot give up on the ‘partly because’ claims of the intuitive picture.

V Conclusion

The Euthyphro challenge posits constraints on what a good metasemantic theory must look like. In particular, it tells us that such a theory must respect both the intuitive picture of representational mental states and the asymmetry of partial explanations. My preferred reaction to this challenge would be to adopt a metasemantic theory that meets both of these constraints—preferably, one that helps us see *why*, given that content is determined the way the theory suggests, content is suitable for playing the explanatory roles that the intuitive picture says that it does. Other reactions are possible, of course—one can instead deny one of the constraints, costly as this denial might be. Regardless of how one reacts to the

³³ One may reasonably think that the connection between beliefs and dispositions to act is that of grounding rather than causation. For our purposes, this difference doesn’t matter much—everything I say about causation can be translated into grounding-talk. I choose to speak in terms of causation both for the sake of simplicity and because the motivation for functionalism is traditionally tied to causation.

³⁴ Applied to functionalism in metasemantics in particular, Block (1989) argues that we should accept CRS because it is the only view that can allow content to be causally efficacious, although he recognizes another problem for the causal efficacy of content if CRS is true, which is discussed in fn. 6.

Euthyphro challenge, one important lesson of this paper is this: functionalist metasemantic views are committed to quite strong claims either in the (extended) philosophy of mind or in metaphysics. In particular, functionalist metasemantic views cannot retain the explanatory efficacy of content without some serious revisions to our notion of explanation or our notion of rationalization. This is not an obvious implication, and it is one that should be taken seriously when deciding whether to opt for a view of this kind.

This challenge generalizes beyond metasemantics. In fact, the Euthyphro challenge applies to quite a few popular metaphysical theories in the literature. To start, it applies to functionalism about the *attitudes* involved in representational mental states no less than it applies to their contents. Put succinctly, the challenge is this: it's part of our intuitive picture of representational mental states that attitudes are efficacious vis-a-vis the consequences of those mental states. For example, it's partly because the subject *believes* that it's Wednesday, rather than suppose or disbelieves, that she infers as she does. The challenge also applies to functionalism about phenomenal character and dispositionalism about color. If we take it that I'm disposed to scream because of the pain, then it cannot be that I'm in pain because I'm so disposed.³⁵ Similarly, if I'm disposed to say that some ball is red because it is red, the ball being red cannot itself be explained by my disposition to say that it is.³⁶

The application of the Euthyphro challenge isn't restricted to philosophy of mind either. This is clear from the fact that the challenge originates from a discussion in metaethics. But even if history hadn't progressed in the way it did, we would still have good reason to consider the application of the challenge in many other domains. The reason is simple: the Euthyphro challenge exploits a highly general feature of explanations. The challenge applies, then, in any context where an entity is being explained in terms of what is pre-theoretically considered to be its consequences.³⁷

Acknowledgements

This paper has benefited from the help of many people. Special thanks to David Chalmers, Cian Dorr and Jorge Ferreira for many cycles of comments and discussions. I also thank Carl Christian Abrahamsen, Cristina Ballarini, Tez Clark, Verónica Gómez Sánchez, Noga Gratvol, Ayala Colette Haddad, Adam Lovett, Adam Pautz, Ali Rezaei, Richard Roth, Michael Strevens, Itamar Weinshtock Saadon, Robbie Williams, Crispin Wright and Patrick Wu, as well as participants and audiences at NYU's Washington Square Circle, NYU's work in progress seminar, the GAP.11 and the Eastern APA 2023.

³⁵ The challenge has been applied to second-order desire theories of pain by Jacobson (2019).

³⁶ A version of the dormitive virtue argument is given by Johnston (1998) against dispositional views of color.

³⁷ To give just one example, a known argument against Humeanism about laws of nature—according to which laws of nature cannot be grounded in particular matters of fact since they scientifically explain the same matters of fact—can be easily presented as a Euthyphro challenge. For one well-known version of the challenge, see Lange (2013).

References

- Block, Ned (1978). "Troubles with Functionalism." *Minnesota Studies in the Philosophy of Science* 9:261–325.
- (1986). "Advertisement for a Semantics for Psychology." *Midwest Studies in Philosophy* 10 (1):615–678.
- (1989). "Can the Mind Change the World?" In George S. Boolos (ed.), *Meaning and Method: Essays in Honor of Hilary Putnam*, pp. 137–170.
- (1998). "Semantics, Conceptual Role." *Routledge Encyclopedia of Philosophy*.
- Boghossian, Paul (2014). "What Is Inference?" *Philosophical Studies* 169 (1):1–18.
- Bridges, Jason (2006). "Does Informational Semantics Commit Euthyphro's Fallacy?" *Noûs* 40 (3):522–547.
- Chalmers, David J. (2021). "Inferentialism, Australian style." *Proceedings and Addresses of the American Philosophical Association*, pp. 184–214.
- Cummins, Robert (1992). "Conceptual Role Semantics and the Explanatory Role of Content." *Philosophical Studies* 65 (1-2):103–127.
- Davidson, Donald (1963). "Actions, Reasons, and Causes." *Journal of Philosophy* 60 (23):685–700.
- Dickie, Imogen (2015). *Fixing Reference*. Oxford University Press.
- Dorr, Cian (2016). "To Be F Is To Be G." *Philosophical Perspectives* 30 (1):39–134.
- Dretske, Fred I. (1981). *Knowledge and the Flow of Information*. MIT Press.
- Fodor, Jerry A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. MIT Press.
- (1989). "Make Mind Matter More." *Philosophical Topics* 17 (1):59–79.
- (2004). "Having Concepts: A Brief Refutation of the Twentieth Century." *Mind and Language* 19 (1):29–47.
- Greenberg, Mark (2005). "A New Map of Theories of Mental Content: Constitutive Accounts and Normative Theories." *Philosophical Issues* 15 (1):299–320.
- Greenberg, Mark and Harman, Gilbert (2006). "Conceptual Role Semantics." In Ernest LePore and Barry C. Smith (ed.), *The Oxford Handbook of Philosophy of Language*, pp. 295–322.
- Harman, Gilbert (1987). "(Nonsolipsistic) Conceptual Role Semantics." In Ernest LePore (ed.), *New Directions in Semantics*, pp. 55–81.
- Jackson, Frank and Philip Pettit (1990). "Program Explanation: A General Perspective." *Analysis* 50 (2):107–117.
- Jacobson, Hilla (2019). "Not Only a Messenger: Towards an Attitudinal-Representational Theory of Pain." *Philosophy and Phenomenological Research* 99 (2):382–408.
- Johnston, Mark (1998). "Are Manifest Qualities Response-Dependent?" *The Monist* 81 (1):3–43.
- Kim, Jaegwon (1989). "The Myth of Non-Reductive Materialism." *Proceedings and Addresses of the American Philosophical Association* 63 (3):31–47.
- Kripke, Saul (1982). *Wittgenstein on Rules and Private Language*. Harvard University Press.
- Lange, Marc (2013). "Grounding, Scientific Explanation, and Humean Laws." *Philosophical Studies* 164 (1): 255–261.
- LePore, Ernest and Barry Loewer (1987). "Mind Matters." *Journal of Philosophy* 84:630–642.
- Lewis, David K. (1966). "An Argument for the Identity Theory." *Journal of Philosophy* 63 (1):17–25.
- (1974). "Radical Interpretation." *Synthese* 27:331–344.
- Loewer, Barry (2012). "Two Accounts of Laws and Time." *Philosophical Studies*, 160 (1): 115–137.
- McLaughlin, Brian P. (1989). "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical." *Philosophical Perspectives* 3:109–135.
- Pautz, Adam (2021). "Consciousness Meets Lewisian Interpretation Theory: A Multistage Account of Intentionality." In Uriah Kriegel (ed.), *Oxford Studies in Philosophy of Mind*, pp. 263–314.
- Peacocke, Christopher (1992). *A Study of Concepts*. MIT Press.
- Plato (1981). "Euthyphro." In G. M. A. Grube (tr.), *Plato: Five Dialogues*, pp. 5–22.
- Quine, Willard van Orman (1960). *Word and Object*. MIT Press.
- Schaffer, Jonathan (2012). "Grounding, Transitivity, and Contrastivity." In Fabrice Correia & Benjamin Schnieder (eds.), *Metaphysical Grounding: Understanding the Structure of Reality*, pp. 122–138.
- Schnieder, Benjamin (2015). "The Asymmetry of 'Because'." *Grazer Philosophische Studien* 91:131–164.
- Schwitzgebel, Eric (2002). "A Phenomenal, Dispositional Account of Belief." *Noûs* 36 (2):249–275.
- Strawson, Galen (2010). "The Depth(s) of the Twentieth Century." *Analysis* 70 (4): 607.
- Wedgwood, Ralph (2001). "Conceptual Role Semantics for Moral Terms." *Philosophical Review* 110 (1):1–30.
- Williams, J. Robert G. (2020). *The Metaphysics of Representation*. Oxford University Press.
- Wright, Crispin (2014). "Comment on Paul Boghossian, 'What is inference'." *Philosophical Studies* 169 (1):27–37.