

PARADOXICAL MENTALITY AND ATTITUDINAL EMBEDDING

ABSTRACT

I explore evidence for the claim that attitudes taking paradoxical contents must be ‘embedded’ within other attitudes that rule out relevant, paradox-inducing worlds. I begin by clarifying this hypothesis, and explaining how evidence bearing on it can be hard won. I then argue that obstacles to holding certain circular suppositional states indirectly provide us with strong, though not dispositive evidence for the embedding hypothesis. I conclude by explaining why the embedding claim would matter: it would render paradox irrelevant to theories of rational attitude formation and rational inference and so—if logic is a theory of this kind—to choice of logic as well.

Try the following experiment. Stare at the second hand of a clock as it winds its way to the 12. Then try as hard as you can to get into the following state of mind at the precise moment the hand lands on 12: suppose, of what you are thereby actually supposing as the second hand lands on 12, that it is true. That is, as the hand lands on 12 you are to suppose something. And the thing you are to suppose is the truth of the very thing you (thereby, actually) suppose. Don’t rush the effort. Take your time and make sure you understand the instructions. Then give yourself a good chance to prepare—30 seconds of lead-up, say. Did you succeed?

Draft of July 29, 2022; please don’t cite without permission

✉: James.A.R.Shaw@gmail.com

† This paper grew out of a series of very helpful conversations with Dustin Tucker, to whom I owe a great debt. Thanks are also due to Mike Caie (who gave helpful feedback on two very different versions of this paper), Ethan Jerzak, Phil Kremer, Jon Litland, Øystein Linnebo, Ben Martin, Doug Patterson, Anders Schoubye, Pixie Shen, and audience members of the *Plurals, Predicates, and Paradox* initiative at the University of Oslo.

You did not. But *why* did you not succeed? My hypothesis is that you did not succeed because it is metaphysically impossible to fulfill this instruction. There is a class of circular mental states—the one I instructed you to have among them—that simply cannot exist. In due course, I'll say more about the class I have in mind and give my reasons for thinking they cannot be formed. But first it may help to situate those arguments. The bounds of mental logical space interest me because of their capacity to influence theories of liar-like paradox. So let me to lead into my argument by previewing its relevance to that domain.

1 PARADOXICAL PROPOSITIONS AND PARADOXICAL EMBEDDING

Eventually, I will explore evidence for the following claim.

PARADOXICAL EMBEDDING: It is metaphysically impossible to entertain (e.g., believe, suppose, etc.) a proposition p that is paradoxical at world w unless one has some other attitude that rules out w .

On this hypothesis, attitudes to paradoxical contents can only be held against a backdrop attitude that treats problematic, paradox-inducing worlds as options which are 'ruled out'—that is, no longer attitudinally live. What exactly I mean by this will hopefully become clearer shortly.

PARADOXICAL EMBEDDING would be vacuously satisfied if there were no such thing as a paradoxical proposition, or no attitudes which took them as contents. While some theorists may embrace such views,¹ I suspect they represent a minority position. In spite of the attractions of banning mental paradox outright, there are strong reasons to think paradox easily infects our attitudes and their objects.

The standard case draws on attitude ascription. Eustace passes in front of a television in a store window, and sees what he takes to be his twin displayed on screen. He surmises from his twin's contorted expression that he must be having deep, inevitably false thoughts. So it can be true to say (1).

(1) Eustace believes that the man on television is having a false thought.

¹ See PARSONS (1974), KRIPKE (1975), and GLANZBERG (2004, 2005) for important views on which paradoxical sentences, at least, fail to express any propositions.

Little does Eustace know that it is he himself on closed-circuit. Is what Eustace believes true? If it is, Eustace believes falsely. But then what he believes should be true. We appear to have paradox in the absence of language.

Or take an example of interdependence: Edmund is reminded of his classmate Lucy, and becomes confident that she is just now believing a truth about professional wrestling. In fact, Lucy speculates that Edmund is in the grip of a false belief about her whereabouts. (2) and (3) can then be truthful reports.

(2) Edmund believes that Lucy is now having a true belief.

(3) Lucy believes that Edmund is now having a false belief.

As before, we may have trouble saying who believes truly and who does not. Problems don't just arise for belief, though it will be my focus. I may hope, fear, or desire to be true or false what you hope, fear, or desire. You may do the same for me. And if things go badly our attitudes may tangle in a degenerate, liar-like way.

Many philosophers have used ascriptions like these to argue for the existence of paradoxical propositions from a few simple (often tacit) premises: that ascriptions like (1)–(3) are true, that their truth requires a propositional object, and that any propositional object of attitudes like (1)–(3) would be paradoxical at the actual world.² Though these premises are contestable,³ I think there is much recommending these simple arguments. At any rate, I'll proceed on the assumption that they succeed, since it only if they do that PARADOXICAL EMBEDDING would need non-trivial support.

On the picture that emerges, the propositional objects of ascriptions like (1)–(3) are contingently paradoxical in the sense that they are paradoxical when evaluated relative to some but not all worlds. What Eustace believes is paradoxical at Eustace's world. But the thing he believes—*that the man on television is having a false thought*—has

² See BURGE (1984, 10-2), SOAMES (1999, 193-4), FIELD (2008, 132-3), SCHROEDER (2010, 284-5).

³ There is a concern from *some* approaches in dynamic semantics that propositional objects cannot always be retrieved from true attitude ascriptions (see DEVER (2006/ms.) for some discussion). This is important because the typical examples of attitude ascription in the case for paradoxical propositions involve anaphoric expressions or quantifiers that are key targets for dynamic treatment (a fact which is closely connected to my eventual case for PARADOXICAL EMBEDDING). There is also a tradition drawing on EVANS (1982) and MCDOWELL (1984, 2005) that treats certain apparently true attitude ascriptions as involving 'mock thoughts' that would not require propositional objects. See GLANZBERG (2003, §6) for a discussion of related issues. Conversely, there are other important cases for the existence of paradoxical propositions considered independently of attitude ascription like WHITTLE (2017).

truth-conditions and can be assessed for truth and falsity relative to other worlds as well. This proposition is intuitively true at worlds where Eustace's twin is displayed on screen forming a false belief about the nature of time. This is what leads Eustace to believe the proposition in the first place.

Once we adopt this picture, there are good reasons to think propositions which are contingently paradoxical are pervasive. For example, (4) can be true whenever I get knowledge that a past occurrent belief of mine at t is false.

(4) Author knows that the occurrent thought Author has at t is not true.

The object of this knowledge (like any) must be a truth. But this object should still be contingently paradoxical in the above sense. For there are non-actual worlds where *at* t (perhaps having lost track of the time, or who I am) I occurrently think that the occurrent thought that Author has at t is not true. In having that thought, I would arguably have the same thought I actually have when I come to the knowledge in (4). If so, the knowledge *actually* reported in (4) is knowledge I have of a contingently paradoxical proposition.⁴

The fact that the propositions figuring in ascriptions like (1)–(4) are merely contingently paradoxical is important for the plausibility of PARADOXICAL EMBEDDING. While the propositions involved in such ascriptions are paradoxical relative to at least one world, they are simply true and false relative to others. PARADOXICAL EMBEDDING says that to attitudinally relate to the contingently paradoxical propositions, that relation must be facilitated by a further attitude that rules out the subset of worlds at which paradox arises for the proposition.

Claims like (1)–(3) are compatible with PARADOXICAL EMBEDDING because they concern thinkers with other beliefs that rule out actuality—the paradox-inducing world of relevance. Eustace believes his twin is on TV. As his twin is not displayed, what he thereby believes is false at actuality. Accordingly, Eustace has a belief that ‘rules out’ the actual world where paradox is induced for the proposition *that the man on television is having a false thought*. As far as Eustace's beliefs are concerned, actuality is not treated as

⁴ Note this requires slightly more than the mental analog of a point, made long ago by KRIPKE (1975), that certain true sentences could have been paradoxical in different circumstances. That is, I am not merely pointing out that I can have a thought directed at a non-paradoxical proposition which could have become paradoxical in other circumstances. Instead, I am noting that the particular true proposition expressed by the complement in (4)—one which I know—*actually* has a paradoxical status that needs accounting for in virtue of its truth-conditions at non-actual worlds.

a live possibility. Edmund believes that Lucy is having an occurrent belief about professional wrestling. As she is not, again, Edmund has a belief that rules out actuality, which is the salient paradox-inducing world for the proposition *that Lucy is now having a true belief*. And something similar is true of (4). As I come to know a past belief was false, I am also aware that the belief I had was not the very thought I am currently having.

Attitudes with paradoxical objects supplied in the literature tend to exhibit the basic structure we see in (1)–(3): an agent has some false belief which leads them to form an attitude to a proposition paradoxical at the actual world. Having the false belief rules out actuality, just as PARADOXICAL EMBEDDING would require. Still, not all cases in the literature have this structure, and it is important to know that PARADOXICAL EMBEDDING is not limited to cases of explicit, antecedently formed beliefs embedding other beliefs. Accordingly, let me make three clarifications about how PARADOXICAL EMBEDDING is supposed to work.

First, embedding and embedded attitudes may be of different types. We can form suppositions in light of certain beliefs or other suppositions. We can take on inquisitive attitudes⁵ like wonder, on the basis of things we’ve already discovered, or merely on the basis of things conjectured.⁶ And we can desire or hope what we do on the basis of what is believed or known.⁷

Second, “attitude” in PARADOXICAL EMBEDDING should be read broadly to include the possibility that the embedding is done by a relation of presupposition.⁸ Here I mean this to be a *mental* relation to certain worlds (or perhaps the absence of certain mental relations) not to be confused with the linguistic phenomenon of presupposition in which, roughly, a sentence’s use would be inappropriate unless the speaker ‘takes it for granted’ that a certain fact obtains.⁹

To give an example of the mental relation of presupposition: most citizens of France in the fifteenth century almost certainly overlooked worlds where the internet would be invented. They did not actively discount such possibilities, and so did not count as

⁵ FRIEDMAN (2013).

⁶ These inquisitive attitudes may take questions and not propositions as their objects, in which case the formulation of PARADOXICAL EMBEDDING must be adjusted in the obvious way.

⁷ Cf. MAIER (2015) and citations from the dynamic semantics literature therein going back to KARTTUNEN (1973). The fact that preferential attitudes can be held against backdrop beliefs will be important in accounting for paradox induced in preferential contexts. See JERZAK (2019) and TUCKER (forthcoming) for authors who lay emphasis on this possibility.

⁸ See STALNAKER (1984) p.88ff.

⁹ See, e.g., BEAVER & GEURTS (2014).

believing that the internet would not be invented. But neither did they treat the possibility of the internet being invented in the way they did other live candidates for the way the world might actually be. So they did not count as believing that the internet *might* still be invented either. The possibility of the invention of the internet was, in some sense, off their doxastic radars. Thinkers can count as overlooking worlds in this sense for a number of reasons: owing to issues of complexity, or lack of conceptual resources, or simply out of inattention or forgetfulness. Here, I will group these phenomena together under the heading of “presupposition” even though there may be important differences between them. I will also call presupposition an “attitude”, even though it may constitute a kind of inactivity unlike the characteristically active mental processes of belief or supposition that often play the embedding role. The important point for us is that all these ways of ‘overlooking’ worlds can leave them out of consideration in the scope one’s further attitudinal commitments. And to be remotely plausible, PARADOXICAL EMBEDDING must include this form of discounting worlds. After all, it is easy to tell a variant of a story like Eustace’s where he does not explicitly come to any beliefs about who exactly is displayed on television, but still comes to believe a proposition paradoxical at actuality. Barring further elaboration, the truth of such a story is compatible with PARADOXICAL EMBEDDING on a broad reading of “attitude”, since all the hypothesis requires is that Eustace presuppose he is not displayed.

A final qualification is that PARADOXICAL EMBEDDING does not require the embedding attitude to be temporally antecedent to the embedded attitude. This point, along with that about presupposition, gives the embedding theorist latitude in treating cases where we cannot identify a simple, active attitude ruling out actuality that precedes the acquisition of an attitude paradoxical at actuality. A case that may force the embedding theorist to rely on this latitude is that of PRIOR (1961), in which a cautious but humble individual comes to believe that at least one of her beliefs is false even though in fact all her beliefs besides this newly formed belief are true. PARADOXICAL EMBEDDING allows such a case can be coherent, though it requires the individual to presuppose, either before or at the time of forming the new belief, that not all her other beliefs besides that new belief are true. This is undoubtedly the normal way that Prior’s case would come about. The person who did not presuppose this would have to actively allow, at some time, that all their beliefs up to that time were true, and maintain this attitude while they integrate the additional belief that one of their beliefs was false. While it is not clear such a case is impossible, it is also much less clearly possible than

the simpler case that Prior certainly intended.

Qualified in the three foregoing ways, PARADOXICAL EMBEDDING is compatible with every example given in the literature cited so far of attitudes relating to paradoxical propositions. This point is of course extremely weak. These examples were not contrived with any intention of challenging PARADOXICAL EMBEDDING. Moreover, this point leaves open that we could concoct new cases with relative ease that would overturn the hypothesis. Accordingly, none of what has been said so far provides us with any reason for thinking PARADOXICAL EMBEDDING is true. It merely goes to show that we do not yet have any decisive evidence against it.

Still, we can ask: given that no example yet considered refutes the hypothesis, how could it be tested? And: what would turn on the question of whether it was true? To quickly preview, the answer to the second question is that PARADOXICAL EMBEDDING could have ramifications for the relevance of paradox to theories of *reasoning*, since reasoning is inherently a mental process. But these ramifications are not easily discerned without a clearer sense of what it would be like for the hypothesis to hold. So we should turn now to the first question above, of how we could reliably test for its truth.

2 CIRCULAR EMBEDDING

One way to test PARADOXICAL EMBEDDING can seem obvious. First, we describe some case where an agent never presupposes or believes the world is not the way it actually comes to be, and yet changes their attitudes so that they are truly described as believing a proposition paradoxical at actuality. Then we ask ourselves: does the case seem coherent?

There are two problems with this methodology. First, the combination of attitudes we must envision on this strategy is highly unusual. Such cases would be very different from those standardly produced to motivate the existence of attitudes with paradoxical contents. The variation on Prior's case recently considered should give a sense of why. Clear violations of PARADOXICAL EMBEDDING seem to require an agent to actively and consciously enter into a paradoxical attitude. If this is possible, the agent involved may have to be highly irrational or deeply confused. Because the agents of these cases are so unusual, the cases themselves are liable to appear such as well. Then we must sort feelings of unnaturalness into those which owe to incoherence, and those which owe to rational remoteness.

Second, the mental relation of presupposition is not easily detectable. There are no

easy tests to ensure that someone has not attitudinally ruled out a world by overlooking its possibility, especially provisionally, as a new attitude is formed.

So it is not that this methodology is completely unworkable. It is rather that its deliverances are inevitably murky and contestable. Intuitions of incoherence-of-agent and incoherence-of-case are liable to clash in unfruitful ways. And even when an attitude seems possible, the suspicion that an agent has subtly presupposed away a paradox-inducing possibility is always looming in the background.

It would be better to have a clearer-cut case. My aim here is to construct one that deals with the both of the aforementioned problems. The key to dealing with the first problem of irrationality is to examine whether something like PARADOXICAL EMBEDDING holds for a broader class of attitudes than those taking paradoxical objects. For many of the attitudes in this broader class there will be no clear obstacle (rational or otherwise) to their formation. We can then deal with the second problem of undetectable presuppositions by rigging the content of the attitude to force it not to rule out the worlds at which interesting circularity arises. When we do both of these things at once, our judgments will be clearer and they will also have clearer implications.

First, some definitions.

A proposition p is *significantly circular at w* just in case the truth-value of p at w significantly depends on p 's bearing a particular property at w .

An attitude is *significantly circular at w* if it is borne to content that is significantly circular at w .

Significant circularity is defined in terms of significant dependence which is best illustrated through examples. The most straightforward cases of significant dependence are present in liar-like or truth-teller-like cases. Consider a proposition which attributes non-truth to itself (at actuality). The truth-value of this proposition significantly depends on whether it bears the property of truth: on the assumption that the proposition bears the property of truth, the proposition would evaluate to falsity; but on the assumption that the proposition does not bear the property of truth, it would evaluate to truth. In this way, the truth-value of the proposition 'significantly depends' on whether the proposition bears the property of truth or not.¹⁰ Though attributions of

¹⁰ In the case of attribution of semantic properties, I prefer to model dependence relations (including 'significant' ones) by exploiting relations of supervenience that formalize the informal reasoning given here—see LEITGEB (2005) for examples of the technique. There are of course alternatives: one could

semantic properties provide key illustrations of significant dependence, I mean for the phenomenon of significant dependence to include attributions of non-semantic properties as well. A proposition which somehow attributes to itself the property of being about snakes might be an example of a proposition that is significantly circular. I also mean for significant circularity to extend to cases where the circularity is generated *indirectly*: for example by a first proposition attributing a semantic property to a second proposition, which in turn attributes a semantic property back to the first.

I qualify that p 's significant circularity at w requires the truth of p to *significantly* depend on whether it bears a property at w because I think there can be instances of a proposition's truth 'depending' on a property it bears—for example by ascribing that property to itself—but in an *insignificant* way. Typical examples of this phenomenon involve generalizations of a logical or quasi-logical kind. For example, suppose I think that all my beliefs are true or false or neither or both. I am inclined to think this particular belief is true. If so, this particular belief is an instance of the phenomenon that it is a belief about. That is, the belief in part attributes a disjunctive property to itself, and the belief must bear that disjunctive property for the generalization that the belief makes to hold. But the attribution of this disjunctive property to itself is 'insignificant.' To tell that the content of my belief was true, you needn't inspect any of my beliefs. The truth of the belief's content followed from a more general truth (that all propositions are true, or false, or neither, or both).¹¹

The distinction here between significant and insignificant dependence is imprecise, but should do for now.¹² With it in hand, we can formulate

CIRCULAR EMBEDDING: It is impossible to entertain (e.g., believe, suppose,

model dependence purely in terms of attribution of properties (cf. GAIFMAN (1992, 2000)), though this makes it harder to capture what I call *significant* dependence relations (see below). In any event, I think an intuitive understanding of these relations should suffice for the present context (see in particular n.12).

¹¹ Note that this informal characterization of the lack of significant self-dependence of certain generalizations again leans on supervenience-style reasoning (whose utility in formalizing notions of dependence was flagged in n.10): the generalizations would receive a given assignment 'no matter what' semantic properties are distributed among their bearers. There are more complex cases. For further discussion of the contrast between significant and insignificant circularity, see [omitted for blind review].

¹² In particular, the imprecision does not affect the argument to come. Previewing: roughly, I am about to argue that significantly circular mentality of a certain kind is impossible. Even if 'significant circularity' is imprecise, paradox-inducing circularity should be significantly circular on *any* precissification—so its impossibility would follow regardless of how the precissification were made.

etc.) a proposition p that is significantly circular at w without having some other attitude that rules out w .

Considering this broadened hypothesis allows us to investigate a class of circular attitudes which needn't have liar-like status. These include attitudes with truth-teller-like status, but also attitudes towards plain and simple truths.

The supposition I invited you to have at the outset of the paper falls into the class of attitudes that would be significantly circular at the actual world, without being liar-like. I asked you to suppose, as a second hand landed on 12, that the thing that you were thereby actually supposing was true. Were you to have fulfilled this instruction, you would have supposed a proposition that is significantly circular at actuality with a truth-teller-like, rather than a liar-like, status.

But there is an additional feature of this supposition that bears on the issue of presupposition which, as recently noted, is extremely tricky to handle in the context of evaluating embedding hypotheses. By inviting you to form a supposition about what you are *thereby, actually* supposing, we are hard-wiring into the content of the supposition that you not rule out features of actuality of which you are aware. In normal cases, the content of your supposition will be one of these. So to form the supposition as instructed, you typically would not be able to presuppose the content of the supposition is anything other than it actually is. That is, the content of the supposition is rigged to prevent you, when successfully forming it, from supposing or presupposing away the circularity-generating features of the supposition at actuality.

I submit that when we put these two conditions together, we cannot hold the attitude in question. And it is not merely this truth-teller-like case where problems arise. I could have asked you to suppose many other things as the second hand landed on 12. Try again as the second hand comes around: suppose of what you are thereby actually supposing as the second hand lands on 12 that it strikes fear into the hearts of accountants. Or suppose of what you are thereby actually supposing as the second-hand lands on 12 that it is the object of a supposition you make at some point today. I submit that none of these instructions is easier to follow than any of the others. They each continue to have a bewildering quality, as if one hasn't been given full directions about *what* exactly one is meant to suppose.

The resistance demands explanation. Here, I want to note two things. First, the best explanation of the resistance is that the suppositions necessarily cannot be formed. Second, the best explanation of why they necessarily cannot be formed is because CIR-

CULAR EMBEDDING holds. Let me take these points in turn.

I have not yet encountered the person who can fulfill my suppositional instructions. If this is really so, why are we encountering these obstacles? There are only two options. On the one hand, it may be simply impossible to hold these attitudes. On the other hand, it may be possible to hold them, while there is some further explanation of why attentive, reasonable, ordinary human thinkers that try to form the attitudes nonetheless typically fail.

I believe we can rule out the second option by noting that there are no good reasons why a reasonable, attentive human thinker operating in good faith couldn't entertain these suppositions if they were indeed possible. One way we could try to explain our failures is by maintaining that the formation of these attitudes is not under our control. On this view, it is not that we cannot make these suppositions *at all*, but that we cannot make them *at will*. But while I sympathize with the doxastic involuntarist, who claims that beliefs cannot be directly held at will, the same claim is manifestly false as regards suppositional states. You can suppose anything provided you have the time and the inclination. And you can do so for 'merely' pragmatic, non-epistemic reasons. If I offer you \$1,000 to suppose that the president of the United States is presently busy twiddling his thumbs, you would doubtless fulfill the instruction and pocket the money without further ado.

Note also that we cannot claim that problems arise because the supposition is overly complex. Suppositions of comparable complexity are relatively easy to hold. Here is an example that will be important for later: try as the second hand lands on 12 to suppose that you are supposing something true as the second hand lands on 12. This is obviously similar to my earlier instruction. But it is different in a crucial respect. In fulfilling my new instruction, you needn't suppose anything about what you *thereby, actually* suppose as the second hand lands on 12. And because of this, fulfilling the instruction is actually quite easy. I can fulfill it by supposing that my children will eat some okra for dinner and that I'm in my office as the second hand lands on 12, forming a supposition that my children will eat okra for dinner. That is an easy way to suppose that what I suppose as the second hand lands on 12 is a truth. It is a kind of counterfactual supposition. First of all, my children would never consent to eating okra. But more importantly, the supposition I (ur-)suppose myself to have is not any supposition I actually make. Still, for all this, my (ur-)supposition can fulfill the original instruction.

This demonstrates that complexity cannot be what is at issue. The difference in

complexity of content between suppositions which do, and do not, explicitly require respecting certain features of actuality are extremely slight—so slight, in fact, that it takes great care to distinguish them. But the gap between how hard it is to suppose them is enormous: one is trivial to suppose, and the other is something that we are unable to suppose regardless of the effort and concentration we put into the task. If complexity were at issue, we should at worst find a slight difference between the ease of making these two suppositions.

We also cannot claim that problems arise because it is somehow irrational to make the suppositions I initially instructed. As already noted, suppositions normally needn't be responsive to epistemic reasons. So they normally needn't be responsive to considerations bearing on whether what is supposed is true, or even plausible. But even if they were, we would still have no explanation of the resistance exhibited by some of the suppositions encountered. Consider the attempt to suppose of what you are thereby actually supposing as the second-hand lands on 12 that it is the object of a supposition you make at some point today. That is a supposition that is *guaranteed to be true* as soon as you suppose it. It has a kind of 'analytic' character. Still, this makes it no easier to suppose in the slightest. So it cannot be a concern for supposing what is true (or plausible, etc.) that is at issue.

Let me recap. We have been considering a would-be class of suppositions, which all involve a form of attitude that can characteristically be held at will. There is no rational obstacle to holding the suppositions. And they are not too complex. We are all attentive and reflective, and have plenty of time to prepare. Still we cannot enter into these suppositions. I think this all provides good evidence that the suppositions in question are not merely challenging, but simply impossible to hold. It would have been fine to maintain that these were merely challenging suppositions to hold, if only there were some reasonable explanation why. But none is in the offing.

Suppose for the moment that this is right. What is responsible for the suppositions' being impossible? It is clear that the circularity they involve is relevant. Of course, the circularity must be what I called a 'significant' one, since the logical generalities discussed earlier give us plenty of examples of attitudes to insignificantly circular content. But it is also mistaken to claim that problems arise *solely* because the suppositions involve significantly circular content. After all, we've seen several examples where significantly circular content is easy to entertain. All attitudes to paradoxical contents in §1, like Eustace's, are of this kind. While these attitudes are semantically perplexing, they

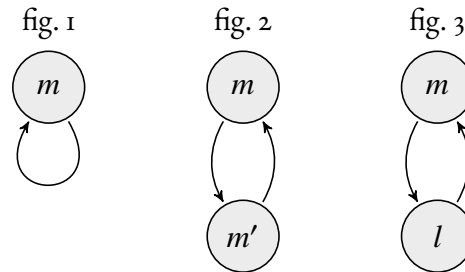
are still easily possible to hold.

What appears to be the problem is that it is not possible to entertain significantly circular content *in a particular way*. What way? Well, the common element to the cases that resist entertaining is that they required one to form suppositions about what one was ‘thereby actually’ supposing. As already noted, what this provision does is to effectively force the supposition in question not to rule out the circularity-generating features of actuality. This is, of course, precisely what is forbidden by CIRCULAR EMBEDDING.

What is more, we’ve seen that as soon as we drop this requirement, the supposition becomes easy to hold. Recall that there is no trouble for anyone to suppose as the second hand lands on 12 that they are making a true supposition as the second hand lands on 12. That’s as easy as making the supposition I noted earlier about my kids eating okra for dinner sometime after I suppose that very thing would occur. Note that the content of this latter supposition (*that one makes a true supposition as the second hand lands on 12*) can still be significantly circular at the actual world. The key difference is that the instructions for forming the supposition borne to this content permit one to rule out the actual world where the significant circularity arises.

So it is easy to suppose significantly circular contents. And it is easy to suppose in ways that don’t rule out various possibilities, including actuality. It is only when we try to force an attitude to have both of these features at once that we encounter any noteworthy resistance to attitude formation. Moreover the resistance, once encountered, seems to be insuperable. As I suggested before, the best explanation of this insuperability is that the attitudes in question simply cannot be held. And that would require something like CIRCULAR EMBEDDING to be true.

I say ‘something like’ CIRCULAR EMBEDDING, because I’ve so far only focused on cases where significant circularity arises for a mental state m through a ‘direct’ self-attribution of a property (fig. 1). But I’ve allowed significant circularity could arise indirectly: for example through property attributions to another mental state m' (fig. 2) or a bit of language l (fig. 3).



But the case for CIRCULAR EMBEDDING from mental reflexivity actually generalizes to these instances of interdependence. It is easiest to take the linguistic case (fig. 3) first.

Consider (5), where “Reader” denotes you, and t denotes a particular future time the second hand on the clock lands on 12 at which you undertake a suppositional task shortly to be elaborated.

(5) The thing actually supposed by Reader as the second hand lands on 12 at t is true.

Is it possible, as the second hand lands on 12 at t , for you to suppose what (5) actually says is true without presupposing away actuality? As before, the key problem in answering this question is managing the slippery mental relation of presupposition. In the case of reflexive supposition we got out of this problem by rigging the content of the supposition so that it, coupled with your awareness of your own mental states, ensured that you could not form the supposition at all unless you did so ‘in the right way’—without introducing illicit presuppositions. No such guarantee is yet present in this case. For example, you can fulfill the instruction to ‘suppose as the second hand lands on 12 at t that what (5) actually says is true’ in part by subtly presupposing (5) has a semantics other than it actually has, or being inattentive to the time, and so forth.

Even so, we can seemingly fix this problem. What we need to do is somehow hard-wire a suitable awareness of relevant circularity-generating features of actuality—including the content of (5), your identity, the time, and so forth—into your suppositional state, especially by integrating it into what you suppose. The relevant features of actuality are the ones that would ensure that the truth of (5) entails the truth of the very thing you thereby actually suppose in supposing (5)’s truth. In other words, to suppose (5) is true in this new way is roughly to follow this instruction: suppose, as the second hand lands on 12 at t , that a sentence labelled (5) in this paper truly says of the very supposition you are thereby actually making that it is true. Or, simplifying a little: suppose of what you thereby, actually suppose that it is true, while (5) happens to report this.

Once we get clear on what the relevant presupposition-free supposition requires, we can see that it effectively subsumes the most problematic aspects of the task of reflexive supposition. I submit that the reflexive suppositional task is not made any easier, or more intelligible, by conjoining it with another suppositional task. As such, the joint task is no more possible than the singular one. I must stress that this does not mean that there is no difference between the suppositional reflexivities and the interdependences. There is of course a substantial difference and one, when presuppositions aren't being delicately tracked, that makes a serious difference to the ease with which the two suppositions are made. The difference between the reflexivity and the interdependence is that the latter introduces *several new presuppositional footholds* to manage while generating the circularity. What we see is that when all these footholds are managed properly to bear on CIRCULAR EMBEDDING, nothing interesting changes. But the management of the presuppositions is a non-trivial task—even this 'simple' interdependence requires bearing quite a few things in mind if we are using it to test CIRCULAR EMBEDDING.

Once this point is appreciated, we can see that the case created by simple reflexivities for CIRCULAR EMBEDDING generalizes to all interdependences, including the simple mental-linguistic interdependence of fig. 3, the mental interdependence of fig. 2, and even more complex interdependences. Consider the simple mental interdependence. When presuppositions are properly managed in this case, the instruction to get into a state of mental suppositional interdependence roughly boils down to this: suppose of what you are thereby actually supposing that it is true, while supposing some other person supposes this too as part of the supposition that what they are thereby actually supposing is true. Bear in mind further that one must suppose all this 'actually' (so you must have reason to think that someone else has been given this instruction, and is executing it simultaneously). Even with a willing, attentive, logistically sophisticated partner in supposition, it is clear that the concerns from the case of simple reflexivity are multiply compounded and exacerbated. There is no interesting space to wriggle out of CIRCULAR EMBEDDING here.

Accordingly, this concludes my core case for CIRCULAR EMBEDDING. As I've been stressing, the only reliable way to test this hypothesis is to look at a highly circumscribed set of mental states directly or indirectly attributing properties to themselves. Most importantly, the only states of this kind that give clear evidence bearing on the hypothesis are those for which we can carefully guard against the encroachment of various presuppositions during their formation. But every time we formulate a case in which those

presuppositions can be properly and clearly managed, we find the same insuperable obstacles to attitude formation—a kind of bewilderment or sense of loss for what attitude we are to form—with no explanations why this would occur owing to concerns of complexity, the nature of the attitude states, or dictates of rationality. The only remaining conclusion is that such attitudes simply cannot be formed at all.

3 FROM CIRCULAR-EMBEDDING TO PARADOX-EMBEDDING

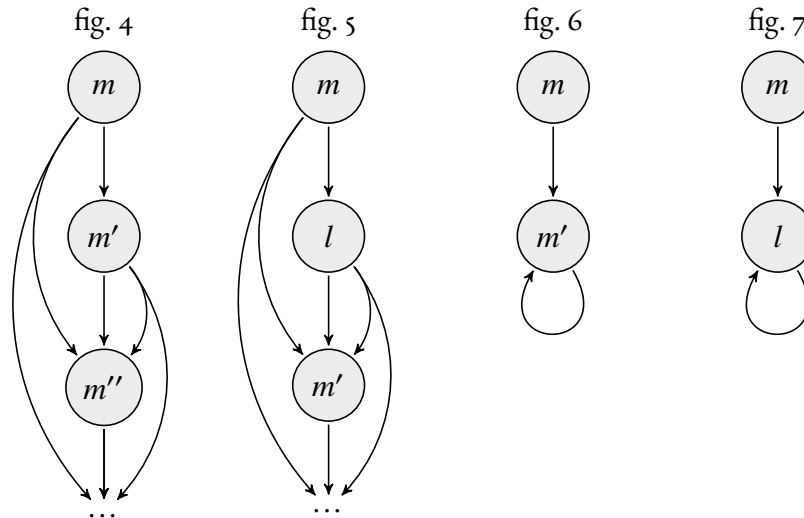
In arguing for CIRCULAR EMBEDDING, we come a long way toward a defense of PARADOXICAL EMBEDDING, since paradoxical attitudes paradigmatically involve significant circularities.

PARADOXICAL EMBEDDING: It is metaphysically impossible to entertain (e.g., believe, suppose, etc.) a proposition p that is paradoxical at world w unless one has some other attitude that rules out w .

But CIRCULAR EMBEDDING does not *entail* PARADOXICAL EMBEDDING as some instances of paradox do not arise ‘within’ circularities. Here, I want to very briefly discuss the complexities introduced by non-circular cases of paradox.

There are two broad sets of cases not covered by CIRCULAR EMBEDDING. First, there are infinite sequences of attributions, all mental (fig. 4) or partially linguistic (fig. 5), which may generate paradox.¹³ Then there are cases where one thinks *about* a circularity generated in thought or language (figs. 6 and 7).

¹³ Cf. YABLO (1985, 1993).



I will have little to say about mental paradox in the infinite setting. This is because I do not think we can get useful information about the tenability of embedding hypotheses in these contexts: there are simply too many presuppositions to manage properly. We've already seen how challenging it is to keep track of relevant presuppositions in even the simplest non-reflexive cases (as in figs. 2 and 3). I am not sure that it is even *possible* for finite agents to properly keep track of the presuppositions involved in the infinite cases. Even if we could in principle do this, it is hard to imagine that we could do so *reliably*. Accordingly, I think these cases need to be set to the side. If we get a good case for PARADOXICAL EMBEDDING independently of the infinite cases, the latter should be lumped in with the general principle. By contrast, if we have any independent evidence against PARADOXICAL EMBEDDING there will be serious concerns that any exemptions extend to the infinite case as well. But that is probably as much as we can hope to ascertain.

This leaves attitudes about independent circularities, as in figures 6 and 7, which introduce some further subtleties. The first thing to note is that if we bracket questions of embedding, mental states about independent circularities clearly exist. For example, theorists of paradox have all sorts of beliefs about the semantic status of a liar sentence like (L) (that it is both true and false, or neither, etc.).

(L) The sentence labelled “(L)” in this paper is not true.

Any such beliefs attribute semantic properties as per figure 7. With figure 6 we need to be just a little more cautious. In arguing for CIRCULAR EMBEDDING we have ruled out

certain kinds of mental states to figure as m' in that diagram. But recall that properly embedded mental states can play that role, since these can still be paradoxical at actuality. Eustace's embedded belief about the man on television is an example. Whatever one believes about paradox generally one will believe about the status of Eustace's belief at actuality—and that will generate the attribution structure in figure 7.

Could exploring the nature of these mental states cast doubt on PARADOXICAL EMBEDDING? We actually encounter two serious obstacles here.

The first is an exacerbated version of familiar concerns about presupposition. If the mental states alluded to are to raise troubles, they must not be embedded in attitudes that rule out paradox-generating worlds. But now that we have a bearer of paradox with an independent existence and set of semantic properties, we have a new set of presuppositions to guard against: those which give this vehicle of paradox semantic properties other than it actually has. After all, any such presupposition could rule out worlds at which paradox arises. But since many aspects of a vehicle of paradox are controversial, we can't obviously use our trick of trying to hard-wire the 'correct' semantics into an attitude's content. There will be no agreement on what that correct semantics is.

This concern is compounded by a second, perhaps more serious problem. Namely, it is controversial whether *any* mental state m figuring in relations as per figures 6 and 7 could be paradoxical to begin with. There are pervasive and varied formal resources that can confine paradox within circularities and prevent it from 'bubbling up' to claims *about* the status of those circularities. Roughly, this can be accomplished by any theory with something like a tiered hierarchical structure. Such hierarchical approaches are traditionally associated with Tarski, though they can in principle be tacked on to many 'rival' formal theories of paradox. The approach needn't be stipulative however: the use of principles that would naturally give rise to separation in cases like figures 6 and 7 are actively embraced by many contextualists like GLANZBERG (2004, 2005). And in fact (traditional) contextualism is not required for such views as demonstrated by the work of GAIFMAN (1992, 2000).

What these theories show us is that, at least as a purely formal or technical matter, paradox never need arise for the mental states m in figures 6 and 7. Views about independently generated paradox can always in principle simply be right or wrong. And this point interacts with the problem of presupposition: *even if* we find evidence for the existence of mental states suitably guarded from improper presupposition, why think this counts against PARADOXICAL EMBEDDING rather than against the para-

doxicality of the attitudes? Not only would confining paradox minimize its presence, but it would result in a simpler theory (given CIRCULAR EMBEDDING). In particular, paradox-confining theories would not require a disjunctive explanation of when embedding is required for the presence of mental paradox, as any alternative theory would.¹⁴

What the foregoing interacting problems tell us is that we will have no uncontroversial evidence against PARADOXICAL EMBEDDING from attitudes about independent circularities. Merely searching for this evidence requires the rejection of substantive theories of paradox on which independent circularities never generate paradox. Obtaining evidence in that context will still require a very delicate management of presuppositions. And even with those caveats, there will be the challenge of disentangling evidential force against PARADOXICAL EMBEDDING from that against the background theory that allows paradox to arise for independent circularities to begin with.

That is hardly the end of matters, as these obstacles certainly do not represent a dispositive case for PARADOXICAL EMBEDDING. Still, the obstacles are enough to shift argumentative burdens squarely to the camp of opposition. And since any further exploration of these issues would require detailed examination of particular theories of paradox, I will leave matters there.

Since this concludes what I will have to say in support of PARADOXICAL EMBEDDING, let me take a moment to review what has been argued so far. First, I've argued that we have clear evidence for PARADOXICAL EMBEDDING from CIRCULAR EMBEDDING. This argument however leaves open that certain non-circular contexts could provide a limited space for unembedded paradoxical attitudes to arise. Unusual features of these non-circular contexts reveal, however, that we will not be able to get clear counterevidence in these domains. In fact, some theories confine paradox almost exclusively to the non-circular setting, in which case the evidence from CIRCULAR EMBEDDING is really

¹⁴ Note these obstacles reveal why, if PARADOXICAL EMBEDDING holds at all, there will be no threat to familiar processes of *hypothetical reasoning* with the semantic properties of vehicles of paradox, which is often how we establish their paradoxical character (just as I did in Eustace's case). Perhaps hypothesizing properties of a vehicle of paradox is possible without any embedding at all since the hypothesizing itself is not paradoxical. Or if it could be paradoxical, it should at least be possible against the backdrop of certain presuppositions, which would not obviously affect the utility of the hypothetical reasoning to reveal paradox. The only place where embedding hypotheses could raise real troubles for hypothetical reasoning is with respect to instances of mentality illustrated in figs. 1–3 of §2. But if PARADOXICAL EMBEDDING holds there is no problem here either, because unlike in the case of Eustace there are *no* attitude states here to reason about in the first place, and so no reasoning 'about them' to safeguard either.

all we will have to go on. We should recognize of course that such theories are controversial, and can be denied. But we've also uncovered a series of interrelated obstacles to getting useful evidence against PARADOXICAL EMBEDDING even within the context of such denials.

4 RAMIFICATIONS FOR REASONING AND LOGIC

My ambition has been to shift argumentative burdens away from defenders of PARADOXICAL EMBEDDING onto its would-be detractors. This is a bit unusual as an ambition because, as noted in §1, there are no actual explicit detractors of yet. The question of PARADOXICAL EMBEDDING really hasn't yet arisen in the literature on paradox. This of course raises the important question of why the hypothesis would even matter.

PARADOXICAL EMBEDDING matters because it radically constrains the roles that paradox can play in theories of rational reasoning and attitude formation. This is in turn because processes of reasoning and attitude formation are inherently mental processes. And according to PARADOXICAL EMBEDDING, such mental processes are shielded from paradox in limited ways. Let me say why this shielding would matter by describing how I understand the relevance of truth-value assignments to theories of rational reasoning and attitude formation.

In a theory of rationality we evaluate one or more attitudes, or transitions between them, as rationally permissible or not, or as good or bad, and so on. Obviously the content of these attitudes can play some part in these evaluations. When the content of some attitude is truth-conditional, its truth conditions—the truth-values of the content relative to particular possibilities—can contribute to an explanation of what makes the attitude appropriate or inappropriate, good or bad, and so on.

Here is a simple, and admittedly simplified, example. Suppose I deductively infer that it is raining from my otherwise rational belief that there is precipitation. This is not a good deductive inference. One way of explaining why it is problematic is to say that there is a possibility where there is precipitation without rain—say, a possibility where it is merely snowing. There is a world at which the content of my concluding attitude is false but at which the content of my basing attitude was true. Arguably, a deductive inference cannot be good or rational if the concluding attitude rules out more worlds than the attitudes on which it is based. Perhaps the explanation of what makes the inference go bad should be slightly more nuanced than this. But the basic idea is near universal: that the truth-conditions of attitudes can play an important role

in determining whether they are rational or not, and since truth-conditions are nothing more than truth-value allotments relative to worlds, these too can make a pertinent explanatory contribution.

This makes it seem like our theories of paradox should have an important role to play in a theory of reasoning. Paradox creates some kind of unusual status at the level of truth-value assignments (e.g., a gap, or a glut, etc.) that may ‘project’ in odd ways with connectives. So it is reasonable to suppose that this status will play a role in our theory of reasoning, and that this role could have disruptive significance.

But if PARADOXICAL EMBEDDING were true, not only would the role of paradox be diminished in a theory of rationality, but it would be obviated entirely. This is because if PARADOXICAL EMBEDDING is true the rationality of an agent’s attitude state vis-a-vis a particular world is never determined by their attitude toward a proposition which has paradoxical status at that world, but (if anything) by a relevant embedding attitude. In other words, for any attitude taking a contingently paradoxical content as object, the role of the content’s paradoxical status at various worlds is not rationally relevant to the assessment of that attitude. And that means the role of paradox is irrelevant, full stop.

How so? Consider an inference where paradox surfaces. Suppose Eustace has the true belief that his brother generally has false beliefs. And suppose he ‘directly’ deductively infers from this the conclusion that the man displayed on screen is having a false thought. This is an inference which has as its conclusion an attitude borne to a content that is paradoxical at actuality. Couldn’t this status at actuality play some role in the explanation of whether the inference is good, or the attitude rational to hold, just as the status of certain propositions at snowy worlds did above?

It cannot. In fact, if PARADOXICAL EMBEDDING is true, the ‘inference’ just given is either under-described or simply impossible. This is because PARADOXICAL EMBEDDING entails that it is impossible for Eustace to arrive at his final belief unless it is properly embedded. And the basing attitude in this case cannot do the embedding. If Eustace ever comes to believe that the man displayed on screen is having a false thought, this is necessarily facilitated by some further belief or presupposition that rules out the actual world. Perhaps this attitude was an additional tacit basing belief, or intermediate conclusion in a chain of reasoning. Or perhaps it was a presupposition Eustace held before, or as, he came to his concluding belief. Either way, it is *this* attitude that settles Eustace’s attitudinal relations to the actual world. For if PARADOXICAL EMBEDDING is

true, the embedded, paradoxical content simply cannot do this, as a matter of metaphysical necessity. Accordingly whether or not Eustace's stance vis-a-vis the actual world is rational/good/advisable/etc. is determined by whether or not this embedding attitude is rational/good/advisable/etc. Since it is metaphysically impossible for the embedded attitude to settle Eustace's attitudinal relation to actuality, no account of the rational status of Eustace's attitudes which bears on his attitudinal relation to actuality will make mention of the content that is paradoxical at actuality.

Of course, this does not mean the embedded attitude is immune from rational assessment. Eustace could be blamed because his concluding attitude rules out worlds where his brother is on TV but uncharacteristically believes truly, since this is not secured by his basing belief that his brother *generally* has false beliefs. As such the concluding embedded attitude could be irrational, or poorly formed, and so on. But the explanation why, if it adverts to the truth-conditions at all, will advert only to those of the embedded content at worlds where paradox does not arise.

The points made in this example are in no way particular to the case. If PARADOXICAL EMBEDDING holds, the points generalize to all attitudes. When any agent has an attitude taking content that is paradoxical at some worlds, PARADOXICAL EMBEDDING requires as a matter of metaphysical necessity that this agent's attitudinal relations to those worlds be settled by other attitudinal means. Any way of evaluating embedded attitudes, whether on their own, as part of broader mental states, or at the basing or conclusory attitudes in a transition, will not advert to information about their relation to worlds at which their contents are paradoxical. Any assessment of the attitudes of an agent at those worlds will only draw on features of the attitudes which do the embedding.

I should concede that there is of course *some*, highly constrained role that a theory of paradox will have in a theory of rational attitudes. This is the role that a theory of anything—for example entomology—might have. There is, I presume, a correct theory of paradox. And there can, I hope, be good evidence for it. Accordingly, there may be rational or irrational views to hold *about* paradox, and in particular about the linguistic and mental vehicles of paradoxical content. Agents may be blamed for forming the wrong views about paradox, especially given their evidence, the way they may be blamed for having the wrong views about the mating habits of dung beetles. When I say that paradox has no real relevance to a theory of reasoning, what I mean is that the theory of paradox, much like the theory of the mating habits of dung beetles, does not contribute

anything to our understanding of general patterns of rational attitude formation or rational attitude transitions.

Here is one way of seeing the force of this point. There is a ‘broadly classical’ picture of good deductive inference, according to which bad inferences involve transitioning to conclusory attitudes that ‘rule out’ more possibilities than basing attitudes. It is the picture I adverted to above in my explanation of the badness of my inference about precipitation. I don’t want to stake too much on the term ‘classical.’ The important point is that this theory of good deductive inference is framed in entirely binary terms: an attitude state rules out some worlds, allows others. And the requirements enforced on good deductive inference in terms of that binary distinction are incredibly straightforward. On them, the hallmark of a problematic deductive inference is a transition to new attitudes that rule out more alternatives than those with which one began.

If PARADOXICAL EMBEDDING is true, the existence of paradox in no way complicates this view of good deductive inference. Insofar as we are concerned with deductive inference, then all thinkers—from Eustace, to Lucy and Edmund, to Prior’s humble thinker and beyond—can be treated with the simple resources of the formulation: in deductive inference, rule out no more worlds than those with which one began. The binary formulation suffices, because even if paradox introduces distinctions that push us into a ternary framework, the third statuses play no role in saying what one should or should not do in deductive inference. If PARADOXICAL EMBEDDING is true, they cannot.

What holds of reasoning or attitude formation will also hold of logic, if logic ultimately concerns itself with something like rational processes of attitude formation. Actually, something very similar to this point was made early on in KRIPKE (1975). Kripke makes use of a three-valued Strong Kleene scheme for interpreting languages containing a self-applicable truth predicate. In this context, Kripke interprets the third value as marking failure to express a proposition. And in a notorious footnote he has the following to say:

I have been amazed to hear my use of the Kleene valuation compared occasionally to the proposals of those who favor abandoning standard logic “for quantum mechanics” or positing extra truth values beyond truth and falsity, etc. Such a reaction surprised me as much as it would presumably surprise Kleene, who intended (as I do here) to write a work of standard mathematical results, provable in conventional mathematics. “Un-

defined” is not an extra truth value, anymore than—in Kleene’s book—*u* is an extra *number* in sec. 63. Nor should it be said that “classical logic” does not generally hold, any more than (in Kleene) the use of partially defined functions invalidates the commutative law of addition. *If* certain sentences express propositions, any tautological truth function of them expresses a true proposition. Of course formulas, even with the forms of tautologies, which have components that do not express propositions may have truth functions that do not express propositions either... Mere conventions for handling terms that do not designate numbers should not be called changes in arithmetic; conventions for handling sentences that do not express propositions are not in any philosophically significant sense “changes in logic.” The term ‘three-valued logic’, occasionally used here, should not mislead. All our considerations can be formalized in a classical metalanguage. (KRIPKE, 1975, 700-1, n.18)

Some have expressed perplexity at Kripke’s remarks. But they strike me as entirely correct given two assumptions. The first is the assumption Kripke explicitly makes that paradoxical sentences fail to express propositions. The second assumption, which Kripke does not flag, would be that logic’s concern with sentences is only instrumental, and that the real target of logical investigation is relations on the contents expressed by such sentences insofar as they bear on something like reasoning or rational attitude formation. In the context of this second assumption, sentences that fail to express propositions are just linguistic excrescences that fail to capture features of the target phenomenon of study. The situation would be, just as Kripke suggests, analogous to the number-theoretic case, where expressions that fail to denote numbers do not require a ‘change in number theory.’ Introducing linguistic tools that fail to express the propositions that are the ultimate targets of inquiry in a theory of rational attitude formation does not, of itself, require any changes in the core of that theory. What it requires is a new degree of attentiveness in one’s linguistic bookkeeping.

For this reason, just as Kripke suggests, if sentences fail to express propositions, then there would be no ‘significant’ change in logic through the transition to a trivalent scheme. There is no pressure induced by this transition to alter the ground-level account of what makes for good reasoning or attitude formation. For example: necessarily, any belief borne to a classical propositional validity will continue to be true, even after the shift to trivalence. The theory does not require us to give up even a single belief

in classical validities that we antecedently viewed as rational. There are of course sentences with the form of classical validities which are not true—but these correspond to no possible mental state. There is accordingly complexity introduced into our theory in the shift to trivalence, but it is complexity that necessarily does not bear on any theory of rational attitude formation.

Having said this on Kripke's behalf, I cannot fully accept his conclusion. The analogy to the number theoretic case would be apt *if* paradoxical sentences failed to express propositions altogether. But I am moved by the reasons given in §1 to think that paradoxical propositions do exist. And if so, I see no reason why sentences should be precluded from expressing such propositions.

Still, what is interesting is that if PARADOXICAL EMBEDDING is true, the lessons of Kripke's footnote may effectively hold anyway. This is perhaps surprising because the argument for PARADOXICAL EMBEDDING did not (like Kripke) make use of any assumptions about what paradoxical status involved. I did not assume paradoxical sentences failed to express propositions, or expressed propositions with gaps, or gluts, or anything else. Rather, the argument looked to the *effects* on attitudes that a theory broadly in the spirit of Kripke's remarks would have required, and found a limited version of those effects to be independently substantiated. The effects found were 'limited' in the sense that we hardly found that there were no possible objects of attitudes with paradoxical statuses. If anything, we found the opposite. But we did find that even though such objects exist, there was noteworthy evidence that paradoxical statuses could not play any explanatory role in the theory of rational attitude formation. The result for logic (if it ultimately concerns itself with such rational attitude formation) is thus broadly as Kripke states: paradox complicates our bookkeeping while we investigate the target phenomenon of reasoning, but paradox does not pressure us to change our ground-level theory of the target phenomenon itself. Reasoning itself could be 'fundamentally classical', even if the logic we use to model that reasoning is significantly more complex.

Fundamentally, PARADOXICAL EMBEDDING bears on theories of reasoning or attitude formation. I want to stress that it only further bears on logic given the additional assumption that logic is a theory of this kind. Though I hold such a view,¹⁵ I will not presume it here. I merely want to note that *other* philosophers often seem to presume it, and make substantive use of that presumption in discussions of paradox.

¹⁵ [Citation omitted]

Just to give one example, consider a standard form of criticism of the Kripkean framework for paradox. Feferman, for example, says that a problem for Kripke's view is that "...*nothing like sustained ordinary reasoning can be carried out in* [Strong Kleene Logic]."¹⁶ Field echoes these remarks in his own criticism of the Kripkean framework for lacking an adequate conditional. In the Strong Kleene setting, he claims, "... [t]he lack of a[n adequate] conditional (and also of a biconditional) cripples ordinary reasoning."¹⁷ Indeed, this claim provides one of the main motivations for Field's attempt to augment Kripke's framework with a 'suitable' conditional.

Neither of these authors (to my knowledge) comments on Kripke's claim that ordinary, broadly classical reasoning is in fact left unchanged in the passage to the Strong Kleene framework, suitably interpreted.¹⁸ In any event, if PARADOXICAL EMBEDDING holds, Field's and Feferman's criticisms are unwarranted. There is no reason to think that there is any pitfall in processes of reasoning for a thinker making use of a concept of truth that was not antecedently present before the introduction of that concept. The only obstacles are for the *theorist of reasoning* in using the linguistic vehicles of paradox to keep track of the relevant rational relations. Strikingly, it is not merely that Field and Feferman's criticisms are unfair as attacks on Kripke, or the Strong Kleene scheme in particular. Rather, the point is that PARADOXICAL EMBEDDING would summarily render *all considerations of reasoning simply irrelevant to paradox*. What counts as a 'good' form of reasoning is an issue that could not immediately speak for or against *any* particular logical treatment of paradox.

Feferman and Field aren't alone. Criticisms like those they raise are found throughout the literature on paradox: that we need this or that change, or this or that connective, to help facilitate or recapture good reasoning in the presence of paradox. PARADOXICAL EMBEDDING would provide a challenge to the tenability of any such claims. And since claims like this drive the direction of a noteworthy portion of the work on paradox—work aiming to 'safeguard' good reasoning—a fair amount turns on whether PARADOXICAL EMBEDDING is true.

So, to summarize: PARADOXICAL EMBEDDING is not important because philosophers have objected to it. In fact, I know of no philosopher who has explicitly offered reasons either in favor of the hypothesis or against it. PARADOXICAL EMBEDDING is

¹⁶ FEFERMAN (1984, 95), emphasis in the original.

¹⁷ FIELD (2008, 73).

¹⁸ Though, in fairness, Field does argue against the 'propositional expression failure' view of paradox that undergirds Kripke's remarks—see n.2.

rather important because it is frequently presupposed false, insofar as it is presupposed that a theory of paradox will bear on a general theory of rational attitude formation or attitude transitions. If PARADOXICAL EMBEDDING holds, a theory of paradox can have no such bearing. This paper has primarily aimed to bring this issue to light, and to provide some evidence which appears to shift dialectical burdens onto those who would like to safeguard the relevance of theories of paradox to theories of rational attitude formation. But I have conceded that this case is hardly dispositive. Hopefully, given the stakes at play, there is more evidence bearing on the issue yet to be uncovered.

REFERENCES

- BEAVER, DAVID I. & BART GEURTS. 2014. "Presupposition." In *The Stanford Encyclopedia of Philosophy*, EDWARD N. ZALTA, editor. Metaphysics Research Lab, Stanford University, winter 2014 edn. [5]
- BURGE, TYLER. 1984. "Epistemic Paradox." *Journal of Philosophy*, vol. 81 (1): 5–29. [3]
- DEVER, JOSH. 2006/ms. "Living the Life Aquatic." [3]
- EVANS, GARETH. 1982. *The Varieties of Reference*. 137. Oxford University Press. [3]
- FEFERMAN, SOLOMON. 1984. "Toward Useful Type-Free Theories. I." *Journal of Symbolic Logic*, vol. 49 (1): 75–111. [25]
- FIELD, HARTRY H. 2008. *Saving Truth From Paradox*. Oxford University Press. [3], [25]
- FRIEDMAN, JANE. 2013. "Question-Directed Attitudes." *Philosophical Perspectives*, vol. 27 (1): 145–174. [5]
- GAIFMAN, HAIM. 1992. "Pointers to Truth." *The Journal of Philosophy*, vol. 89: 223–61. [8], [18]
- . 2000. "Pointers to Propositions'." In *Circularity, Definition and Truth*, A. CHAPUIS & A. GUPTA, editors, 79–121. [8], [18]
- GLANZBERG, MICHAEL. 2003. "Against Truth-Value Gaps." In *Liars and Heaps*, J. C. BEALL, editor, 151–94. Oxford University Press. [3]
- . 2004. "A Contextual-Hierarchical Approach to Truth and the Liar Paradox." *Journal of Philosophical Logic*, vol. 33: 27–88. [2], [18]
- . 2005. "Truth, Reflection, and Hierarchies." *Synthese*, vol. 142: 289–315. [2], [18]
- JERZAK, ETHAN. 2019. "Paradoxical Desires." *Proceedings of the Aristotelian Society*, vol. 119 (3): 335–355. doi:10.1093/arisoc/a0z003. [5]
- KARTTUNEN, LAURI. 1973. "Presuppositions of compound sentences." *Linguistic Inquiry*, vol. 4 (2): 167–193. [5]

- KRIPKE, SAUL A. 1975. "Outline of a Theory of Truth." *Journal of Philosophy*, vol. 72 (19): 690–716. [2], [4], [23]
- LEITGEB, HANNES. 2005. "What Truth Depends On." *Journal of Philosophical Logic*, vol. 34 (2): 155–192. doi:10.1007/s10992-004-3758-3. [8]
- MAIER, EMAR. 2015. "Parasitic Attitudes." *Linguistics and Philosophy*, vol. 38 (3): 205–236. doi:10.1007/s10988-015-9174-z. [5]
- MCDOWELL, JOHN. 1984. "De Re Senses." *Philosophical Quarterly*, vol. 34 (136): 283–294. [3]
- . 2005. "Evans's Frege." In *Thought, Reference, and Experience: Themes From the Philosophy of Gareth Evans*, JOSÉ LUIS BERMÚDEZ, editor. Clarendon Press. [3]
- PARSONS, CHARLES. 1974. "The Liar Paradox." *Journal of Philosophical Logic*, vol. 3: 381–412. [2]
- PRIOR, A. N. 1961. "On a Family of Paradoxes." *Notre Dame Journal of Formal Logic*, vol. 2 (1): 16–32. [6]
- SCHROEDER, MARK. 2010. "How to Be an Expressivist About Truth." In *New Waves in Truth*, CORY D. WRIGHT & NIKOLAJ J. L. L. PEDERSEN, editors, 282–298. Palgrave Macmillan. [3]
- SOAMES, SCOTT. 1999. *Understanding Truth*. Oxford University Press USA. [3]
- STALNAKER, ROBERT. 1984. *Inquiry*. Bradford Books, MIT Press. [5]
- TUCKER, DUSTIN. forthcoming. "Paradoxes and the Limits of Theorizing About Propositional Attitudes." *Synthese*, 1–20. doi:10.1007/s11229-018-01902-2. [5]
- WHITTLE, BRUNO. 2017. "Self-Referential Propositions." *Synthese*, vol. 194 (12): 5023–5037. doi:10.1007/s11229-016-1191-0. [3]
- YABLO, STEPHEN. 1985. "Truth and Reflection." *Journal of Philosophical Logic*, vol. 14 (3): 297–349. doi:10.1007/BF00249368. [16]
- . 1993. "Paradox Without Self-Reference." *Analysis*, vol. 53 (4): 251. doi:10.1093/analys/53.4.251. [16]