

SAVING MENTAL FICTIONALISM FROM COGNITIVE COLLAPSE

Meg Wallace

Abstract: Mental fictionalism maintains that: (1) folk psychology is a false theory, but (2) we should nonetheless keep using it, because it is useful, convenient, or otherwise beneficial to do so. We should (or do) treat folk psychology as a useful fiction—false, but valuable. Yet some argue that mental fictionalism is incoherent: if a mental fictionalist *rejects* folk psychology then she cannot appeal to fictions in an effort to keep folk psychological discourse around, because fictions *presuppose* the legitimacy of folk psychology. Call this the Argument from Cognitive Collapse. In this paper, I defend several different mental fictionalist views against cognitive collapse.

1 Fictionalism, Mental Fictionalism, and Cognitive Collapse

An individual who denies (or doubts) the existence of some entity, *X*, will generally refrain from adopting an *X*-committing discourse, *D*.¹ If you think that dragons do not exist, for example, you will not utter claims to the contrary. But not always. You are discussing *Game of Thrones* with a friend and utter (1):

(1) There are three dragons on Khaleesi's shoulders.

Your utterance is acceptable—perhaps even true—even though you do not think dragons (or Khaleesi) exist. This case is not exceptional; we often utter (and accept) such statements when talking about fictions or fictional characters.

Some philosophers have taken the above phenomenon to be instructive, for it shows that we are able to divorce the (seeming) ontological commitments of our utterances from what we in fact believe exists. This suggests that a similar phenomenon occurs in less obviously fictional cases—for example, in our conversations about numbers, unobservable entities, and possible worlds.

¹ An *X*-committing discourse involves at least one statement (in its logical form) that existentially quantifies over *X*s. For simplicity, I assume that the grammatical form and logical form in (1) and (2) do not come apart.

Suppose, for example, that you are a reductive physicalist who thinks that there is no room in a purely material universe for moral properties of the sort required for true moral statements. Yet you may nevertheless utter (2):

(2) There are moral obligations.

One way to reconcile the acceptability of (2) with an ontology that does not include moral properties or obligations is to claim that an utterance of (2) is similar to an utterance of (1) (Nolan et al. 2005). Both seemingly involve ontologically suspicious entities—dragons and moral obligations, respectively—yet both are nonetheless acceptable or true. So *however it is* that we divorce the seeming ontological commitments of an utterance of (1) from our actual commitments, one might argue, *that* is what is going on with (2).

Fictionalists about *Xs* maintain that one may endorse an *X*-committing discourse, *D*, even if one is an eliminativist (or agnostic) about *Xs*. This may be for several reasons. The fictionalist may (i) deny that certain sentences in *D* are *really X*-committing; (ii) deny that uttering sentences in *D* is really *asserting* anything about *Xs*, and is only *pretending* to assert or *quasi-asserting* sentences in *D*; and (iii) grant that certain sentences in *D* are *X*-committing but nonetheless accepts and uses such sentences because, say, in certain contexts usefulness trumps truth, and so on.²

Philosophers have been fictionalists about fictional characters and content (Lewis 1978; Brock 2002), possible worlds (Rosen 1990, 1993, 1995; Nolan 1997 and 2011), mathematical objects (Field 1980 and 1989), scientific unobservable entities (van Fraassen 1980), moral truths (Nolan et al. 2005), composite objects (Dorr and Rosen 2002), and truth (Burgess and Burgess 2011; Woodbridge 2005). Many find fictionalism attractive because it allows us to continue talking as if there are *Xs*, even though we do not think that *Xs* exist. Why would one want to keep talking about *Xs* if *Xs* do not exist? There may be lots of reasons: for entertainment, because it is useful, it is aesthetically valuable, it may be more expressive to do so, or it may be just psychologically and sociologically convenient.

Motivating moral fictionalism, Nolan, Restall, and West explain:

One obvious advantage of fictionalism over eliminativism is psychological convenience. . . . Eliminativism about moral discourse would force great and wide ranging changes to our patterns of speech and thought on much the same scale as would eliminativism about folk psychological concepts of the sort famously proposed by the Churchlands. Ceasing to talk of goodness and badness, rightness and wrongness, duties, justice, and obligations would be much like ceasing

² See, for example, Yablo's distinction between Instrumentalism, Meta-fictionalism, Object-fictionalism, and Figuralism (2001).

to talk of people having beliefs, desires, and emotions: possible, perhaps, but not an easy thing to do and certainly not a consequence to embrace lightly. We think that this is no small motivation for seeking a way of retaining the talk, without committing the error. (2005, 326)

Indeed. Doing away with moral discourse—while possible—is as unrealistic and implausible as *doing away with folk psychological discourse*. Yet if such considerations provide strong motivation to explore *moral* fictionalism, they may likewise provide strong motivation to explore *mental* fictionalism.

A *mental fictionalist* is at heart a mental eliminativist, or mental agnostic.³ She maintains that common sense folk psychology is somehow untrue. She may claim that the entities folk psychology quantifies over—belief states, desires, sensations, qualitative feels, propositional attitudes, etc.—either don't exist, or she refrains from committing to their existence, but also refrains from committing to the denial of their existence. Yet she also maintains that folk psychology is incredibly useful, that ordinary discourse is riddled with it, and as such, should not (or could not) be abandoned. Mental fictionalists aim to find a way to keep folk psychological talk while rejecting the folk psychological walk.^{4,5}

However, unlike other fictionalist views, the mental fictionalist may appear particularly problematic. To see this, consider *eliminative materialism*: the thesis that there are no folk psychological mental states, desires, or propositional attitudes.⁶ And consider what we intuitively take fictions to be: stories that we have certain *propositional attitudes* toward. We *pretend* that so-and-so is the case, and act *as if* it is true. We use our imagination and engage in pretense, all of which may require beliefs and mental engagement—the very things that an eliminative materialist is an eliminativist about. If fictions and pretense require folk psychological states and activities, then any eliminative materialist who denies folk psychology

³ I say 'mental eliminativist' instead of 'eliminative materialist' because the latter is often associated with the particular form of mental eliminativism endorsed by the Churchlands—1981; 1986, for example. I take mental eliminativism to be a broader category, of which eliminative materialism is a type. See section 2.1.

⁴ Mental fictionalism is relatively new in the philosophical literature, and not all proposals for mental fictionalism allow for agnosticism about the relevant entities. See Wallace Unpublished, Joyce 2013, Sprevak 2013, Hutto 2013, Parent 2013, Marton and Janos 2013, for discussion.

⁵ It should be noted that fictionalism need not assume eliminativism or agnosticism—one may be a realist and still be a fictionalist. See Kalderon 2005, for example. However, because the argument from cognitive collapse (the focus of this paper) only gains traction when eliminativism is assumed, I will only focus on mental fictionalisms that assume mental eliminativism.

⁶ "Eliminative materialism is the thesis that our common sense conception of psychological phenomena constitutes a radically false theory, a theory so fundamentally defective that both the principles and the ontology of that theory will eventually be displaced, rather than smoothly reduced, by completed neuroscience" (Churchland 1981, 67).

denies fictions and pretense as well. But if a *mental fictionalist* is an eliminative materialist at heart, then she must deny propositional attitudes toward fictions as well, which seemingly directly undermines her own view.⁷

In other words, fictionalists about *X* generally take our intuitive understanding of fictions and apply this to some theoretical discourse. But our intuitive understanding of fictions, one might argue, assumes the legitimacy of certain folk psychological concepts. So one cannot *reject* FP and simultaneously appeal to fictions to *save* FP appearances—for fictions assume FP. Hence, mental fictionalism is incoherent. Call this the *argument from cognitive collapse*.⁸

What follows is a defense of mental fictionalism (MF) against arguments from cognitive collapse. I begin by canvassing varieties of mental eliminativism. Because fictionalism is often motivated by arguments for eliminativism, there will be at least as many types of mental fictionalism as there are types of mental eliminativism. In addition, there are a variety of ways to interpret *fictions*, which, combined with one or another of the many mental eliminativist positions, makes for a variety of possible mental fictionalist views. Some of these views, I readily admit, have little plausibility. But others are not so easily dismissed, and have notable theoretical advantages. I will analyze some of the more promising mental fictionalism positions, and suggest ways to avoid cognitive collapse.

2 Varieties of Mental Fictionalism

Many categorize fictionalism as having an *ontological* thesis and a *linguistic* thesis.⁹ The ontological thesis is that entities of a certain kind in a given discourse do not exist.¹⁰ The linguistic thesis is a claim about how to treat the relevant discourse—i.e., how to keep it in use without admitting widespread error from its users. For the fictionalist, this involves saying what it is for something to be fictionally true, or true in the fiction. Since both the ontological and linguistic theses will be relevant to understanding mental fictionalism, I discuss each in turn.

⁷ I take ‘folk psychological states and activities’ to be those states and activities that feature in everyday psychological discourse, as in Lewis’s platitude sense of folk psychology.

⁸ This is similar to Baker’s ‘cognitive suicide’ argument against eliminative materialism in Baker 2004. Joyce (2013) calls this same problem “fictionalist suicide.” See Parent 2013 and Sprevak 2013 for variations of this argument against mental fictionalism, and Joyce 2013 for discussion of two ways to address the argument. The present paper is a continuation of some of the ideas put forward in Joyce 2013, which (Joyce acknowledges) is an expansion of Wallace Unpublished. This paper differs from Joyce 2013 and Wallace Unpublished because of the emphasis on the varieties of mental eliminativism, which gives rise to a variety of unique responses to worries of cognitive collapse.

⁹ See Nolan et al. 2005, Eklund 2009, Parent 2013, et. al., for example.

¹⁰ Or that they *may* not exist, and we thus should—for whatever epistemic reasons—be agnostic about whether they do or not exist.

2.1 Mental Eliminativism

A mental eliminativist denies the ontological commitments uniquely posited by folk psychology. She is committed to (NegEx):

(NegEx): Some of the entities posited by folk psychology (FP) do not exist.¹¹

Importantly, however, a commitment to (NegEx) leaves open exactly *which* entities posited by folk psychology do not exist and *why*. As we will see below, specifying these details will result in distinct eliminativist positions.

Feyerabend (1963), for example, maintains that many folk psychological terms presuppose a dualist ontology. Notions such as ‘belief’ and ‘desire,’ he claims, are essentially non-physical in character. There is no hope for mental/physical reduction because mental terms inherently presuppose a non-physicalism. Such philosophers accept the following as a reason to reject FP:

(DualOnt): Some of the entities posited by folk psychology presuppose a dualist ontology.

We might extend Feyerabend’s reasoning as a version of Mackie’s argument from queerness against moral properties. Mackie (1977) maintains that our commonsense moral notions assume that moral facts or properties have a normative force that is not found elsewhere. So, moral facts and properties, if they exist, would be very strange entities—so strange that this should give us reservation about admitting them into our ontology. Similarly, we might think that, for all folk psychology says, mental states or properties would have to be very unusual or ‘queer’ sorts of things. So, one might argue, the entities and properties required for the truth of folk psychology are just as odd as those required for the truth of our moral statements, in which case both FP and moral discourse should be jettisoned.¹²

Quine (1960) and parts of Churchland (1981), in contrast, do not claim that folk psychological terms require a non-physicalist ontology. But they argue that folk psychological notions are too sloppy or vague to pick out anything real. Put another way: purely scientific and physical explanations are *more accurate* than folk psychological ones. Suppose, for example,

¹¹ This is a variation of Parent’s (Elim) in Parent 2013.

¹² Interestingly, one move available in the moral case that is not (obviously) available for this particular kind of mental eliminativism is non-cognitivism. Non-cognitivists about moral discourse aim to retain moral discourse but avoid its presumed problematic ontology by insisting that moral claims are non-propositional, or non-truth-apt. So assertions in the moral discourse are not ontologically committing, despite surface grammar appearances to the contrary. Instead, a moral claim is (say) an expression of an attitude such as ‘yay helping granny across the street’ or ‘boo kicking puppies.’ This strategy will not work with folk psychology, since the attitudes needed for an expressivist explanation are the very attitudes the (above kind of) mental eliminativist is suspicious of—i.e., *expressing attitudes* is a paradigmatic folk psychological notion, and of the very sort that (certain) mental eliminativists are eliminativists about.

I ask how a certain light fixture works. One answer may be “Flip the switch, and the light turns on.” But supposing that I was looking for a more informative explanation—one that might allow me to install a light fixture in the future—then this oversimplified explanation is less accurate than a more detailed one involving information about currents and wires. Analogously, one might argue that folk psychology is a less accurate or detailed description of mental activity than, say, neuroscience.¹³

If so, you might endorse:

(Vague): Some of the entities posited by folk psychology are too vague, sloppy, indeterminate, or inaccurate to pick out anything real.

Alternatively, one might reject specific posits of folk psychology. Rorty (1970) and Dennett (1978; 1988), for example, argue that mental *sensations* or *qualia* do not exist, respectively. Rorty argues that folk psychological talk of *sensations* should go the way of talk of demons: “‘sensation’ might lose its reporting role as well as its explanatory role, just as ‘demon’ had lost both its roles, and that both of these roles might be taken over by reference to brain-processes” (1970, 112).

Dennett (1988) argues that our folk psychological notions of *qualia* are so mistaken:

My claim, then, is not just that the various technical or theoretical concepts of qualia are vague or equivocal, but that the source concept, the ‘pre-theoretical’ notion of which the former are presumed to be refinements, is so thoroughly confused that, even if we undertook to salvage some ‘lowest common denominator’ from the theoretician’s proposals, any acceptable version would have to be so radically unlike the ill-formed notions that are commonly appealed to that it would be tactically obtuse . . . to cling to the term. Far better, tactically, to say that there are no qualia at all. (382–383)

¹³ “FP is *at best* a highly superficial theory, a partial and unpenetrating gloss on a deeper and more complex reality” (Churchland 1981, 74). And, “A serious advance in our appreciation of cognitive virtue would thus seem to *require* that we go beyond FP, that we transcend the poverty of FP’s conception of rationality by transcending its propositional kinematics entirely, by developing a deeper and more general kinematics of cognitive activity, and by distinguishing within this new framework which of the kinematically possible modes of activity are to be valued and encouraged (as more efficient, reliable, productive, or whatever). Eliminative materialism thus does not imply the end of our normative concerns. It implies only that they will have to be reconstituted at a more revealing level of understanding, the level that a matured science will provide” (84).

Such philosophers may endorse:

(NoQualia): Private sensations and qualitative feels—as posited by folk psychology—do not exist.

Importantly, someone who accepts (NoQualia) may grant that human beings have beliefs and desires, propositional attitudes, or other posits of folk psychology, provided that such concepts do not presuppose the qualitative, conscious experiences that Rorty and Dennett deny. That is, accepting (NoQualia) is consistent with the view that human beings understand and communicate with one another propositionally.¹⁴

Yet another reason to be a mental eliminativist may be for somewhat more abstract, theoretical reasons, much like the ones Paul Churchland (1981) gives, which may leave one agnostic about which entities in FP are problematic. One could admit that folk psychology is rife with explanatory failures and infertility and is unable to integrate with other, advancing scientific theories, yet also acknowledge ignorance about which details, exactly, FP gets wrong. Thus, one might endorse:

(BadTheory): Folk psychology is a stagnant, infertile theory that has no hope for integrating with other advancing theories.

Alternatively, one might endorse a view closer to eliminative materialism and accept:

(NoUS): There is no unified state of the brain to correspond with common sense notions of belief, desire, etc., as posited by folk psychology.

An endorser of (NoUS) maintains that it is folk psychology's presumed *unified states* (US) of the brain—belief or desire states—that are problematic. Defenders of connectionist networks may hold this. Ramsey et al. (1990), for example, claim that connectionist networks

do not correlate in any systematic way with the functionally discrete, semantically interpretable states posited by folk psychology and by more traditional cognitive models. Since information is encoded in a highly distributed manner, with each connection weight and bias embodying information salient to many propositions, and information regarding any given proposition scattered throughout the network, the system lacks functionally distinct, identifiable sub-structures that are semantically interpretable as representations of individual propositions. (514)

¹⁴ Marton and Janos, for example, maintain that “there is a *difference in metaphysical category* between our conscious experiences and the kind of propositional attitudes such as belief” (2013, 631). I am not endorsing this claim, but I am allowing that it is a theoretical possibility.

If FP presupposes discrete, unified states of the brain, and neuroscience's best theory of cognitive activity is a connections model, then FP is radically mistaken and should be rejected.¹⁵

One might also endorse a stronger claim such as:

(NoProp): There is no such thing as propositional content—hence, there are no propositional attitudes or states, as FP claims.¹⁶

Or even stronger still:

(NoRep): Some of the entities posited by folk psychology do not exist because there is no such thing as intentionality, content, or representation *tout court*.¹⁷

(DualOnt)-(NoRep) are by no means exhaustive; they are just some of the ways to be a mental eliminativist. Mark Sprevak (2013), for example, proposes (but does not endorse) Neural Representation Fictionalism, which is fictionalism about the mental representations *as posited by cognitive science*. This does not explicitly entail fictionalism about *folk psychology*. Alternatively, one might deny the unified self (*a la* Hume), which would undermine most of FP. Such a move isn't motivated by a rejection of representation across the board, but by a rejection of the notion of a unified self—the entity or thing which, according to FP, *has* mental states, beliefs, and desires. However, I hope the above variations are enough for the purposes of this paper.

2.2 Fictions and Fictional Content

Fictionalists are committed to fictions, fictional talk, or some kind of distinction between the literal and non-literal. But what *are* fictions, fictional talk, and the distinction between literal and non-literal?

According to Walton (1990), fictions are works of representational art that we actively and mentally engage in. We use our imagination, which is somehow guided by objects that are intended (by the author or artist) to represent certain scenarios, and allows us to engage in games of make-believe. If Walton is correct, then it is easy to see how someone might think that this undermines mental fictionalism. If one endorsed (NoRep), for example, and thought that there is no intentionality, representation, or propositional content across the board, then it would be inconsistent to also maintain that we should be fictionalists about the mental—where fictions are understood as stories that *represent* the world in some way, are *about* thus-and-so, and involve us *imagining that* such-and-such is the case. If a mental fictionalist eliminates all representation and propositional content,

¹⁵ (NoUS) may be a more detailed version of (Vague).

¹⁶ Churchland is often interpreted as endorsing (Vague), (BadTheory), (NoUS), and (NoProp).

¹⁷ Boghossian (1990), for example, argues that certain mental eliminativists are committed to something like (NoRep).

then she cannot also appeal to a *representative* fiction—with propositional content!—to save folk psychology.

And it is not just fictions that interest the metaphysical fictionalist, but also story-telling, pretense, and acting as if. Lewis claims:

Storytelling is pretense. The story teller purports to be telling the truth about matters whereof he has knowledge. He purports to be talking about characters who are known to him, and whom he refers to, typically, by means of their ordinary proper names. But if his story is a fiction, he is not really doing any of these things. (1978, 40)

Yablo explains:

The fictionalist holds that we “make as if” we are asserting that *S* and/or believing that *S* and/or receiving the news that *S*. Our reason for making as if we are doing these things . . . is that it serves some larger purpose. Making as if *S* enables us simplifies [sic] our theory, or shortens proofs. (2001, 74)

But *pretending*, *pretending to assert*, and making or acting *as if*, are also, presumably, folk psychological notions. If such activities assume the integrity of FP, or require representation in some way, then they are apparently in conflict with some of the eliminativist claims above—in particular, (NoRep), (NoProp), and perhaps (NoQualia).

However, the above point is contingent on whether ‘acting as if’ and ‘pretending to assert’ involve folk psychological concepts. Against this assumption, a mental fictionalist could insist that we are *acting as if* when we engage in folk psychological talk without having anything further to say about what this ‘acting as if’ involves. This would be a way of avoiding charges of blatant self-undermining or incoherence: if such a mental fictionalist refuses to say what we are in fact doing when we use fictional discourse, then an objector cannot insist that doing so presupposes folk psychology. This would be a *mental instrumentalist fictionalist*, to adopt and modify Yablo’s terminology (2001).

An obvious problem with this move would be what Yablo calls the *problem of real content*: we certainly *seem* to be asserting or claiming or stating something when we say things such as “Granny believes there is beer in the fridge.” What exactly are we doing if not asserting or stating something? The instrumentalist doesn’t say. And while this view would be coherent, it is nonetheless incomplete and inadequate; it is silent on the one aspect of the theory of which a mental fictionalist surely needs an account.

Other problems that Yablo lays out for the instrumentalist generally apply to the mental instrumentalist as well. We accept that Granny believes there is beer in the fridge, but reject that she believes there is beer under the couch. If the instrumentalist has no account of what we are doing

when we act as if folk psychological statements are true, then we have no explanation as to why certain utterances are correct (or acceptable) and certain others are not. Yablo calls this *the problem of correctness*.

Because of the problem of correctness and the problem of content, being a mental *instrumentalist* fictionalist is not going to be a viable option for the mental fictionalist.

An alternative is to claim, like Field or van Fraassen, that we are only *quasi-asserting* sentences of folk psychology. But what is quasi-asserting? Is *quasi-asserting* a folk psychological concept?

Van Fraassen (1980) elaborates: “When a scientist advances a new theory, the realist sees him as asserting the (truth of the) postulates. But the anti-realist sees him as displaying this theory, holding it up to view, as it were, *and claiming certain virtues for it*” (57). But this explanation is metaphorical. We *display* a theory, *hold it up to view*, and *claim* certain virtues for it. Is it a work of art? Are we observing it as we would a painting in a gallery? If so, does such an activity presume folk psychology? Here we are, looking at a theory, forming certain beliefs and opinions about it, and having qualitative reactions to it. If *this* is what quasi-assertion amounts to, it certainly *seems* as if folk psychological notions are in play. So appealing to *quasi-assertion* will not be a viable move for (some) mental fictionalists.

To put the point another way: appealing to *quasi-assertions* may be helpful for fictionalists who merely need a way to distinguish literal from non-literal content. But a mental fictionalist has to do more than that—she must distinguish between literal and non-literal content such that *the non-literal content does not presuppose the truth of FP*. It is not obvious that this can be done (although more on this in [section 3](#)).

Moreover, if a quasi-assertion that *S* is just a genuine assertion that *C(S)*—where *C* is the condition needed to make the statement correct¹⁸—then we haven’t eliminated assertions. So if one is a mental fictionalist of the sort that eliminates propositional content across the board, or assertions across the board, then appealing to quasi-assertions in this way will be undermining.¹⁹

We might, following Field (1980), take fictions to be the literally false (or unaccepted) content, the base discourse to be the literally true (or accepted) discourse, together with a set of bridge principles that link the base discourse to the fiction. So we get something like:

X believes *P* iff according to the folk psychological fiction,
X believes *P*.

¹⁸ Yablo (2001, 75) attributes this view to van Fraassen (1980).

¹⁹ If one is a mental fictionalist because one endorses (DualOnt), (Vague), (NoQualia), (BadTheory), for example, then appealing to an account of fictions that assumes propositional attitudes may not be problematic. This is because each of these mental eliminativist positions reject FP on grounds other than FP’s presupposition of the legitimacy of propositional content. I save elaboration on this point for [section 3](#).

The mental fictionalist is then left with the task of analyzing what it is for something to be true *according to the folk psychological fiction*. In particular, whatever analysis she gives cannot be undermined by her brand of mental eliminativism. This may not be difficult to do, however, since we have yet to establish that *according to the folk psychological fiction* is a folk psychological concept as opposed to a purely logical notion.

A Lewisian about fictional truth, for example, maintains that fictional statements have a suppressed hidden logical operator: “In the fiction, $F . . .$ ” This operator is then analyzed as a restricted quantifier over possible worlds:

In the fiction F , P iff P is true in all worlds in a set, S , of possible worlds which is determined by F .

Such an analysis does not (obviously) require mental sensations, qualia, or even propositional attitudes—in short, it need not require mental representation or content.

Yet Lewis ultimately suggests that the relevant set of possible worlds required for interpreting the fictional operator involve “worlds where the fiction is told, but as known fact rather than fiction” (1978, 40). But this brings us back to analyzing fictions in terms of story-telling and *known* facts, both of which seemingly require mental attitudes or folk psychology.

Moreover, one might argue that Lewis’s account requires representation, since possible worlds (for Lewis) are *ways of representing how the world could be*. Lewis claims the following about *de re* modality:

Humphrey may be represented *in absentia* at other worlds, just as he may be in museums in this world. The museum can have a waxwork figure to represent Humphrey, or better yet an animated simulacrum. Another world can do it still: it can have as part a Humphrey of its own. . . . By having such a part, a world represents *de re*, concerning Humphrey . . . that he exists and does so-and-so. By waiving its arm, the simulacrum in the museum represents Humphrey as waiving his arm; by waiving his arm, or winning the presidential election, the other-worldly Humphrey represents the this-worldly Humphrey as waiving or winning. (194)

Is such *de re* representation of the problematic sort that is rejected by (certain) mental eliminativists? According to (NoRep), for example, there is no representation *tout court*. Does this include other-worldly individuals *representing* this-worldly individuals *in absentia*? Does it include wax figures representing real people? Most likely the latter is problematic, since depictions or sculptures are *intended* to represent, which clearly involves mentality. But the modal representing that Lewis discusses—despite the misleading wax simulacrum analogy to the contrary—does not so obviously

involve mentality. An other-worldly tree with six branches may represent a this-worldly tree with five branches (making it the case that this-worldly tree could have had six branches), *even if there had never been any minds* on any world whatsoever. So a mental eliminativist—including one who endorses (NoRep)—may allow that other-worldly objects *represent* this-worldly objects. It's just that the representation here is of an innocuous sort. Admittedly, the plausibility of this line may depend on the strength of one's interpretation of (NoRep).

Alternatively, one could take a Lewisian-like analysis of the fictional operator, only diverging from Lewis by rejecting his claim that the relevant worlds are ones where the fiction is *told*, or told as *known* truth. This addresses one of the worries I raised above. If one endorses (DualOnt), for example, then endorsing a possible world analysis of fiction, where the relevant worlds are ones where an individual tells a story, may not be problematic—unless *telling* presupposes the folk psychological notions that one is an eliminativist about (a point which would need to be argued for). A certain kind of mental eliminativist, then, can accept that individuals in possible worlds move their mouths and utter sounds and go through certain motions (and perhaps even utter propositions, if one does not endorse [NoProp], for example), without having to accept that such individuals *have* propositional attitudes, or *know* so-and-so to be true, or have any mental qualia related to the sounds or utterances they are making. So a mental fictionalist could interpret Lewis's possible world analysis of the fictional operator without thereby committing to the very entities she is an eliminativist about.

Or a mental fictionalist could take a purely formal analysis of the fictional operator. Suppose one accepts: in the fiction *F*, *P* iff *F* logically entails *P*. Such an analysis does not involve intentionality; on the contrary, it is a purely logical relationship. However, such an analysis of fictional content is oversimplified and incomplete. It is never explicitly stated—and arguably not logically entailed—in the Sherlock Holmes fictions that the characters are on planet Earth. Our usual understanding of truth in fiction relies on the intentions and beliefs of the author and audience to settle contextual, background details. Most mental fictionalists will reject this understanding of fictional content. But having an analysis that is broadly logical (and non-intentional) is nonetheless a plausible option for the mental fictionalist, even if she is then under some pressure to provide more detail and a more thorough account of what, exactly, a fiction is.

To put the above points more generally, a mental fictionalist grants that we should keep folk psychological discourse around because it is useful to do so. But she may insist that it is useful because there is an objective, non-intentional relationship between the folk psychological 'story' and the genuine, material facts of the mind—where 'story' here is also understood in an objective, non-intentional sense. So it is plausible that the mental fictionalist has an account of the fictional operator that is logical, not

mentalistic. But if so, then such a mental fictionalist will not suffer cognitive collapse.

3 Defense Against Cognitive Collapse

It was stated previously that if one is motivated toward the kind of mental eliminativism that rejects *any* intentional content, representation, etc., then it would be incoherent to adopt a representational fiction as a way out of making a commitment to mentality—one has just swapped a discourse riddled with intentionality (folk psychology) for another (fictions). Out of the frying pan into the fire, it's hard to see how such a fictionalist view could even get off the ground.

But this seems to eliminate only mental fictionalist views that accept (NoRep); the remainder are still open possibilities. Take, for example, someone who endorses:

(DualOnt): Some of the entities posited by folk psychology presuppose a dualist ontology.

Endorsing (DualOnt) leaves open whether there might be some other notion of content (or representation or meaning) that is non-dualistic. Talk of a 'false' theory is not self-refuting, because it may be that propositional talk does not presuppose non-physical entities in the way that certain folk psychological notions do (e.g., beliefs and desires). Accepting (DualOnt), in other words, does not mean that *all* talk of content, meaning, or aboutness should be abandoned, for it remains to be seen whether there are some representational concepts that do not presuppose a non-physicalist ontology. It is coherent to think that there are no mental states as described by common-sense psychology, but that there are representational items such as sentences and propositions.

If this is right, then one could adopt a notion of a fiction that is non-dualistic as well. One might accept that it is not relevant that we believe such-and-such in order to have fictional content. A fiction may just be a set of sentences that are false. One could divorce the *pretense* and *make-believe* concepts from purely propositional notions, in order to have a theory of fiction that does not presuppose a dualist ontology.

Alternatively, suppose one accepts:

(Vague): Some of the entities posited by folk psychology are too vague, sloppy, indeterminate, or inaccurate to pick out anything real.

(Vague) does not entail that there isn't any kind of representation across the board—that is, that there is no concept of meaning, content, or representation that refers to something real in the world. It is only the claim that folk psychological concepts are somehow inferior or inaccurate relative to neuro-scientific explanations. If you think that FP is too sloppy to have terms that pick out anything real, but neuroscience is not, then you still

think there is a way to capture cognitive activity. It's not as if *nothing* is going on inside! This cognitive activity may include something like 'engaging in pretense,' 'acting as if,' or 'making believe,' only these sorts of concepts would be explained "at a more revealing level." Just because we reject FP on the grounds that it is too simplistic to capture the complex neural activity that actually occurs does not mean that we won't, in the end, have a concept of fiction and fictional content that is likewise more complex than FP currently supposes.

Suppose, for example, that instead of the folk psychological notions of *pretending*_{FP}, we have a scientifically respectable notion, *pretending*_{SR}, which is analogous to *pretending*_{FP} in appropriate ways. The suggestion is that the mental fictionalist appeals to *pretending*_{SR} instead of *pretending*_{FP}.²⁰ Moreover, the mental fictionalist would maintain that engaging in fictions, as required by any fictionalist account, involves *pretending*_{SR}, not *pretending*_{FP}.

One might object, however: in what sense would this be a *fictionalist* account, if we insist that fictions involve *pretending*_{FP} and not *pretending*_{SR}?²¹ To assume that the mental fictionalist must accept a folk-psychological notion of fiction is question-begging against her view, since the very topic at issue is the legitimacy of folk psychological notions. If a mental fictionalist insists that folk psychological notions should be abandoned, then she likewise rejects a folk-psychological notion of fiction. To assume that fictions can *only* be understood in terms of *pretending*_{FP} and not *pretending*_{SR}, in other words, is stacking the deck against the mental fictionalist. At the very least, the burden of proof is on those who make the stronger claim that our notion of fiction *must* assume the legitimacy of FP, in the face of a mental fictionalist who offers an alternative understanding. (I will have more to say about this below.)

The strategy by now should be relatively clear: the mental fictionalist may avail herself of a non-mentalistic understanding of fictions, thus avoiding cognitive collapse. Say, for example, you endorsed any of the following:

(NoQualia): Private sensations and qualitative feels—as posited by folk psychology—do not exist.

(BadTheory): Folk psychology is a stagnant, infertile theory that has no hope for integrating with other advancing theories.

²⁰ Joyce (2013), following Wallace (Unpublished), suggests something similar. However, Joyce maintains the following: "inasmuch as the thesis of fictionalism is essentially characterized with reference to folk psychological entities . . . then psychological fictionalism simply cannot be literally correct." As I argue below, however, the assumption that the "thesis of fictionalism is essentially characterized with reference to folk psychology" is exactly what the mental fictionalist should deny. Moreover, to insist that this assumption is correct is question-begging against the mental fictionalist. So, I do not think that mental fictionalism is as vulnerable to cognitive collapse as Joyce does.

²¹ Thanks to Ted Sider for pressing me on this point.

(NoUS): Private sensations and qualitative feels—as posited by folk psychology—do not exist.

Accepting (NoQualia) does not entail denying content and representation across the board. Or fictions. Lewis's analysis of truth-in-fiction, for example, uses sets of possible worlds. It does not require sensations or qualia. If (NoQualia) is consistent with thinking that human beings understand and communicate with themselves and each other propositionally, then such a view is also consistent with a propositional analysis of fictions (and truth-in-fiction).

Likewise, (BadTheory) and (NoUS) leave open the possibility of, say, a connectionist model version of fiction or pretense, or a version of fiction that does not presuppose a stagnant, infertile theory. Recall the proposal of *pretending*_{SR} versus *pretending*_{FP}. A mental eliminativist might have a connectionist-friendly notion of fiction, involving *pretending*_{SR}, which is somehow analogous (or a successor) to the folk psychological notion of fiction. It may be objected that such a move understates the gulf of difference between folk psychological concepts and complex neural reality. Those who endorse connectionism generally think that there is nothing at all in the brain that will correspond to discrete, localized mental states as folk psychology presumes. This is why they are eliminativists and not reductionists (Ramsey et al. 1990). Moreover, if a mental fictionalist inclined toward connectionism makes use of a connectionist-friendly notion of fiction, then why wouldn't such a person also invoke connection-friendly notions of beliefs and desires, and dispense with eliminativism?

As stated earlier, a mental fictionalist need not appeal to an account of fiction that appeals to pretense or pretending. She may avail herself of a purely logical relation, or one that invokes possible worlds, none of which (obviously) undermine a mental eliminativist who endorses (NoUS). I admit that more needs to be said on behalf of the fictionalist than I have the space to address here. But I see at least two plausible options for the mental fictionalist who adopts (NoUS): (i) she may grant that there is some notion of fiction that uses *something like* pretense, but this involves a non-FP notion of pretense—and she is then left with working out the details of why she is still eliminativist about beliefs and desires even though she thinks there is a salvageable, acceptable notion of pretense; (ii) she could deny that fictions require mental attitudes at all, and that the *in the fiction F* operator is a purely logical relation. Either way, a mental fictionalist who endorses (NoUS) avoids cognitive collapse.

Even if one accepts a mental eliminativism that is relatively strong, such as—

(NoProp): There is no such thing as propositional content—hence, there are no propositional attitudes or states, as FP claims.

—there is still a way to avoid cognitive collapse. To see this, let us look at a particular instance of the argument from cognitive collapse against

the mental fictionalist, and contrast this with the material eliminativist's response to self-refutation worries. Parent (2013) defines mental fictionalism as accepting the following ontological and linguistic theses, respectively:

(Elim): The mental states posited by folk psychology do not exist.

(SPS): Sentences of mentalistic discourse can be true in a fiction relative sense. More precisely: If 'm' is a term for a mental state and ' $\Box(x)$ ' is a formula, ' $\Box m$ ' is true iff, according to the fiction of folk psychology, there is a mental state m such that $\Box m$.

Yet Parent argues that the above combination is incoherent:

(SPS) speaks in earnest about a 'mental fiction,' yet the notion of a 'fiction' itself seems to be a folk-psychological notion. On its face, a 'fiction' is characterized by a certain attitude we take toward it . . . the act of fictionalizing that a is *F* consists in the following rule:

(*): Regard A as being *F*, regardless of whether A is *F*.

But to 'regard' an object as being *F* is to adopt a certain attitude toward it, which again, is a mentalistic phenomenon.

(610)

But the mental fictionalist should rightly insist that the problem is with Parent's interpretation of (SPS). Parent has assumed that a mental fictionalist must interpret fictions as something that we have a mental attitude toward. The above discussion has repeatedly shown that this need not be the case. Moreover, I have suggested ways to interpret operators such as 'in the fiction, *F*' that need not involve mentalistic representation, or (perhaps) representation at all. So a mental fictionalist has very good reason to reject Parent's (*), yet she may nonetheless have a non-mentalistic and acceptable interpretation of (SPS), undermining Parent's argument from cognitive collapse.

This move is similar to one Churchland (1981) makes in response to critics who think that eliminative materialism is self-refuting. Churchland imagines three scenarios of how life might be without FP. In the first scenario, humans adopt a much more complex theory about human cognition, which is compatible with all cognitive activity—humans and non-humans alike. From the perspective of this new theory, Churchland argues, FP will seem simplistic and superficial.²² Such a theory would supplant FP, but could nonetheless provide an account of how we communicate, exchange information, and how we may even have 'knowledge.' Only,

²² "suppose that research into the structure and activity of the brain . . . finally does yield a new kinematics and correlative dynamics for what is now thought of as cognitive activity. . . . From the perspective of the newer theory, however, it is plain that there simply are no law-governed states of the kind FP postulates. The real laws governing out internal activities are defined over different and much more complex kinematical states and configurations, as are the normative criteria for developmental integrity and intellectual virtue" (Churchland 1981, 85).

communication, information exchange, and knowledge are (somehow) all non-propositional.

According to Churchland, a more sophisticated, alternative theory of cognitive activity would involve entities (propositions) that are not “evaluated as true or false, nor are the relations between them remotely analogous to the relations of entailment, etc., that hold between sentences. They display a different organization and manifest different virtues” (87). So while an alternative theory of the mind would not be propositional, would not involve entities that are evaluated as true or false, nor bear relations such as entailment, the theory would presumably involve entities displaying *different virtues*.

While it is admittedly difficult to evaluate what such a theory would look like, given that it is only presented in the abstract, and in contrast to what it is *not*, let us grant Churchland this (admittedly nutty) idea for the moment. If an alternative theory of cognitive activity is possible—one that is more complex and radically different from FP—and if such a theory provides us with a full explanation of how we communicate and relay information, then such a theory might also have an explanation of fictional content. The distinction between fiction and non-fiction, in other words, need not be tied to a *propositional* theory of content or meaning.

What would a non-propositional, non-mental theory of fiction look like? I have no idea. But if Churchland thinks that he can imagine at least three different ways we might supplant FP with a more complex, non-propositional theory of cognitive activity, then I do not see why a theory of fictional content couldn’t possibly undergo a similar upgrade.

One of the scenarios that Churchland discusses involves imagining that human beings eventually learn to communicate in the immediate way that two brain hemispheres communicate by way of pathways in the subcortical regions. Churchland imagines that human beings may eventually learn to communicate with each other in the way that our bodies already automatically relay information internally:

Libraries become filled not with books, but with long recordings of exemplary bouts of neural activity . . . they do not consist of sentences or arguments.

How will such people understand and conceive of other individuals? To this question I can only answer, ‘In roughly the same fashion that your right hemisphere ‘understands’ and ‘conceives of’ your left hemisphere—intimately and efficiently, but not propositionally!’ (88)

But what goes for folk psychological concepts such as ‘understand’ and ‘conceive of’ will go for folk psychological concepts such as ‘fiction,’ ‘truth in fiction,’ ‘pretense,’ etc. We might imagine someone asking: *how will people discuss fictional content without FP?* To which we could answer: *intimately and efficiently, but not propositionally!*

Again, it is difficult to give much detail here since Churchland is generalizing from merely possible theories that we only have a glimpse of in the abstract and by contrast (i.e., by what they are *not*). Nonetheless, the following seems right: if we grant Churchland his three scenarios in defense of his eliminative materialism (and how a world might be without FP), it seems we can similarly grant a mental fictionalist the possibility of the concept of a fiction that does not presuppose FP.

But, one might object, how is appealing to fantasies of future languages of any help to a mental fictionalist, who is pressed to give a fictionalist theory of mental discourse *as it is used here and now*?

I do not propose that the only way to save a mental fictionalist from arguments from cognitive collapse is to invoke fantasies about future people and possible descendants of communication that are non-propositional. This is one way that a mental fictionalist—even one who went to far as to endorse (NoRep)—*could* avoid cognitive collapse. Granting wild possibilities about alternative forms of communication is helpful in emphasizing my broader point (which stands independently of Churchland's examples): arguments from cognitive suicide must assume that fictions and fictional content are mentalistic, but since the legitimacy of mental talk is the very point at issue, such attacks are question-begging against the mental fictionalist.

To modify Churchland's response to charges of cognitive collapse, one cannot simply assume a particular theory of fictions and fictional content without begging the question at issue—namely, the legitimacy of FP. Admittedly, many fictionalist views do include mentalistic concepts such as make-believe, acting as if, and pretense. But the forgoing discussion hopefully demonstrated that this is unnecessary. If arguments from cognitive collapse against the mental fictionalist assume a mentalistic concept of fiction, then it is this notion of a fiction that must be rejected, not mental fictionalism.

In summation, the argument from cognitive collapse concludes that mental fictionalism is incoherent because it assumes the legitimacy of folk psychological concepts (by invoking fictions) when these folk psychological concepts are the very thing the mental fictionalist denies. Yet we have seen that the details of what, exactly, the mental fictionalist denies matters greatly. Some of the ways of being a mental eliminativist are compatible with accepting fictions, so long as she accepts an FP-free notion of fiction. For example, (DualOnt) and (NoQualia) reject FP for reasons that do not obviously conflict with accepting fictions and truth-in-fiction. If one accepts a Lewisian notion of truth-in-fiction, for example—one that uses possible worlds and individuals uttering true statements in those worlds—then no appeal to a dual ontology or ineffable qualitative states are needed, making these kinds mental fictionalism immune to cognitive collapse worries. Such a mental fictionalist would have to commit to being a certain kind of Lewisian about fictional content, of course, which makes for an additional theoretical commitment that one might not have anticipated. But I take it

that anyone endorsing a fictionalist view needs to take a stand on what, exactly, fictions are, or what it is to be true-in-a-fiction. The arguments of this paper—if correct—show that being a mental fictionalist carries with it a demand to adopt a notion of fiction, pretense, and truth-in-fiction that is compatible with one’s specific brand of mental eliminativism.

Certain kinds of mental fictionalism, then, are coherent, even if such theories need to provide a more thorough account of fictions and fictional content. So at best: arguments from cognitive collapse put pressure on the mental fictionalist to provide more detail about what, exactly, a fiction is. But they do not straightway collapse her view.

Meg Wallace

E-mail : meg.wallace@uky.edu

References:

- Baker, Lynne Rudder. 2004. “Cognitive Suicide.” In *Philosophy of Mind: A Guide and Anthology*, edited by J. Heil, 401–413. Oxford: Oxford University Press.
- Boghossian, Paul. 1990. “The Status of Content.” *Philosophical Review* 99 (2): 157–184. <http://dx.doi.org/10.2307/2185488>.
- Brock, Stuart. 2002. “Fictionalism About Fictional Characters.” *Noûs* 36 (1): 1–21.
- Burgess, Anthony and John Burgess. 2011. *Truth*. Princeton, NJ: Princeton University Press.
- Churchland, Patricia. 1986. *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: MIT Press.
- Churchland, Paul. 1981. “Eliminativism and Propositional Attitudes.” *Journal of Philosophy* 78 (2): 67–90. <http://dx.doi.org/10.2307/2025900>.
- Dennett, Daniel. 1978. *Brainstorms*. Cambridge, MA: MIT Press.
- Dennett, Daniel. 1988. “Quining Qualia.” In *Consciousness in Contemporary Science*, edited by A. J. Marcel and E. Bisiach, 42–77. New York: Oxford University Press.
- Dorr, Cian and Gideon Rosen. 2002. “Composition as a Fiction.” In *The Blackwell Guide to Metaphysics*, edited by Richard M. Gale, 151–174. Oxford: Blackwell.
- Eklund, Matti. 2009. “Fictionalism.” In *Stanford Encyclopaedia of Philosophy*, Spring 2009 edn., edited by Edward N. Zalta. Stanford Encyclopaedia of Philosophy. <http://plato.stanford.edu/archives/spr2009/entries/fictionalism/>.
- Feyerabend, Paul. 1963. “Mental Events and the Brain.” *Journal of Philosophy* 40: 295–296. <http://dx.doi.org/10.2307/2023030>.
- Field, Hartry. 1980. *Science Without Numbers*. Princeton, NJ: Princeton University Press.
- Field, Hartry. 1989. *Realism, Mathematics and Modality*. Oxford: Blackwell.
- Hutto, Daniel. 2013. “Fictionalism about Folk Psychology.” *The Monist* 96 (4): 585–607. <http://dx.doi.org/10.5840/monist201396427>.
- Joyce, Richard. 2013. “Psychological Fictionalism, and the Threat of Fictionalist Suicide.” *The Monist* 96 (4): 517–538. <http://dx.doi.org/10.5840/monist201396424>.
- Kalderon, M. 2005. *Moral Fictionalism*. Oxford: Clarendon Press.
- Lewis, David Kellogg. 1978. “Truth in Fiction.” *American Philosophical Quarterly* 15 (1): 37–46.
- Mackie, J. L. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin.
- Marton, Miklos and Tozser Janos. 2013. “Mental Fictionalism as an Undermotivated Theory.” *The Monist* 96 (4): 622–638. <http://dx.doi.org/10.5840/monist201396429>.

Acknowledgements Thanks to Sam Cowling, Daniel Nolan, Ted Sider, Keith Simmons, Tim Sundell, and two anonymous reviewers for helpful feedback, comments, and discussion on earlier drafts. I’d also like to thank the audience and participants at the Mental Fictionalism Conference at the University of Edinburgh in Edinburgh, Scotland, July 2014.

- Nolan, Daniel. 1997. "Three Problems for 'Strong' Modal Fictionalism." *Philosophical Studies* 87 (3): 259–275. <http://dx.doi.org/10.1023/A:1004218817629>.
- Nolan, Daniel. 2011. "Modal Fictionalism." In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Originally published in 2002 with revisions in 2007 and 2011. Stanford Encyclopedia of Philosophy. <http://plato.stanford.edu/archives/win2011/entries/fictionalism-modal/>.
- Nolan, Daniel, George Restall, and Caroline West. 2005. "Moral Fictionalism Versus the Rest." *Australasian Journal of Philosophy* 83 (3): 307–330. <http://dx.doi.org/10.1080/00048400500191917>.
- Parent, Ted. 2013. "In the Mental Fiction, the Mental Fictionalism is Fictitious." *The Monist* 96 (4): 605–621. <http://dx.doi.org/10.5840/monist201396428>.
- Quine, W. V. O. 1960. *Word and Object*. Cambridge, MA: Harvard University Press.
- Ramsey, William, Stephen Stich, and Joseph Garon. 1990. "Connectionism, Eliminativism and the Future of Folk Psychology." *Philosophical Perspectives* 4: 499–533. <http://dx.doi.org/10.2307/2214202>.
- Rorty, Richard. 1970. "In Defense of Eliminative Materialism." *The Review of Metaphysics* 24 (1): 112–121.
- Rosen, Gideon. 1990. "Modal Fictionalism." *Mind* 99 (395): 327–354. <http://dx.doi.org/10.1093/mind/XCIX.395.327>.
- Rosen, Gideon. 1993. "A Problem for Fictionalism about Possible Worlds." *Analysis* 53 (2): 71–81. <http://dx.doi.org/10.1093/analys/53.2.71>.
- Rosen, Gideon. 1995. "Modal Fictionalism Fixed." *Analysis* 55 (2): 67–73. <http://dx.doi.org/10.1093/analys/55.2.67>.
- Sprevak, Mark. 2013. "Fictionalism About Neural Representations." *The Monist* 96 (4): 539–560. <http://dx.doi.org/10.5840/monist201396425>.
- van Fraassen, Bas. 1980. *The Scientific Image*. Oxford: Oxford University Press.
- Wallace, Meg. Unpublished. "Mental Fictionalism."
- Walton, Kendall. 1990. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, MA: Harvard University Press.
- Woodbridge, James. 2005. "Truth as a Pretense." In *Fictionalism in Metaphysics*, edited by Mark Eli Kalderon, 132–177. Oxford: Clarendon Press.
- Yablo, Stephen. 2001. "Go Figure: A Path Through Fictionalism." *Midwest Studies in Philosophy* 25 (1): 72–102. <http://dx.doi.org/10.1111/1475-4975.00040>.