

# Mental Fictionalism<sup>1</sup>

in [\*Mental Fictionalism: Philosophical Explorations\* \(Routledge\)](#)

Tamas Demeter, Ted Parent and Adam Toon (eds.)

PRE-PUBLICATION DRAFT - Please cite original

Meg Wallace - meg.wallace@uky.edu

## 1. Introduction

Suppose you are somewhat persuaded by the arguments for eliminative materialism, but are put off by the view itself. For instance, you might be sympathetic to one or more of the following considerations: (1) that folk psychology is a bad theory and will be soon replaced by cognitive science or neuroscience, (2) that folk psychology will never be vindicated by cognitive science, (3) that folk psychology makes ontological commitments to weird or spooky things that no proper science will admit the existence of, (4) that folk psychology seems to lead to a sort of epiphenomenalism (which is yet *another* thing that's weird and spooky), and (5) that folk psychology seems to lead to the conclusion that mental content is either determined by things outside the head or is completely *indeterminate*, neither of which is appealing.<sup>2</sup> Yet in spite of your sympathy for any one of (1)-(5), you may nonetheless cringe at the consequence of them—that is, you may be unwilling to accept the Eliminative Materialist's radical claim that (i) there are no beliefs, desires, etc., and (ii)—even more radically—we should stop all talk that quantifies to the contrary.

Such a dilemma is typical in metaphysics, especially for those guided by a healthy Quinean methodology for ontological commitment. One popular solution to such a dilemma has been to turn to *fictionalism*. The fictionalist strategy suggests: take the suspicious entity, *E*. Take all sentences, *S*, which seemingly ontologically commit one to *Es*. Reinterpret these sentences as sentences that are *not* ontologically committing.<sup>3</sup> Provide a brief explanation as to why these

---

<sup>1</sup> This paper was originally written in 2007. I initially made it publicly available on my website and distributed it freely whenever requested. Some of it was updated, modified, and used for Wallace (2016), but otherwise, I left this paper aside for many years as I pursued other projects, and as the literature on mental fictionalism grew. Unfortunately, in the wake of such rapid growth of quality work produced on this topic, much of the content in this paper is out of date. Nonetheless, I think there is value in understanding its place in a burgeoning field, so I've decided to leave the paper relatively unchanged from 2007. I have lightly edited and updated the main content, added footnotes and commentary when necessary, and referenced recent work when appropriate. For those interested in seeing the larger landscape of philosophical debate on mental fictionalism, I hope this paper serves as a decent place to start. It was originally inspired by Daniel Nolan's impressive work on moral fictionalism and fictionalism in general, and Bill Lycan's array of arguments against eliminative materialism. See Nolan (2002), (2004) and Lycan (2005). A huge intellectual debt and special thanks are owed to both of them for their work in this area, and for personal comments on several drafts.

<sup>2</sup> These reasons (1)-(5) in support of Eliminative Materialism are borrowed from Lycan (2005).

<sup>3</sup> I am intentionally leaving "reinterpret" underspecified. One may either be a *hermeneutic* or a *revolutionary* fictionalist. If one is a hermeneutic fictionalist, then one thinks that our language in a particular discourse is already

sentences, *S*, are useful—or should still be uttered—yet are nonetheless false (if taken merely at face-value).

Fictionalism of various sorts has been proposed, endorsed, and discussed; it has its fair share of fans and foes.<sup>4</sup> Yet despite fictionalism's broad application, there is a noticeable gap in the literature. There is no discussion of *mental fictionalism*—the idea that we are (or should be) fictionalists about mental entities.<sup>5</sup> This is surprising on its face, since philosophers of mind have battled long and hard about how best to save folk psychology and our ordinary belief-desire talk in light of the overwhelming success and advances of neuroscience. As suggested above, the considerations for eliminative materialism *are* fairly persuasive. Yet if the cost of accepting these arguments means forfeiting our ordinary, folk psychological vocabulary, then we might think twice about the legitimacy of the eliminative materialist's arguments.<sup>6</sup>

My job in this paper is to put mental fictionalism on the map, albeit tentatively. I *do* think that mental fictionalism has some advantages over other fictionalist views in the literature. I also think that it has some advantages over other theories of the mind—most notably, eliminative materialism. However, I also think that there are worries which plague mental fictionalism that do not plague other fictionalist accounts, and which do not plague other theories of the mind. In the latter sections in this paper, I suggest some ways the mental fictionalist might overcome these worries, hopefully paving the way for further development of various kinds of mental fictionalism as legitimate, plausible positions in the philosophy of mind. Please do note: this is merely a conversation *starter*, not a battle-hardened defense of a view. There is much more to say - both for and against mental fictionalism - than what is said here.

---

fictionalist; it is a descriptive thesis about language. If one is a revolutionary fictionalist, then one thinks that our language in a particular discourse *should* be fictionalist; it is a prescriptive thesis about the language. Both of these will be discussed in more detail below. See Burgess and Rosen (1997) and Stanley (2001) for more on the distinction. The details of how one might sever our *talk* from our *commitments* can be done in a couple of ways. One way would be to introduce a hidden fictional operator (Lewis 1978), which is the strategy I follow here. But alternatives are possible; see below section 3 for discussion.

<sup>4</sup> See Lewis (1978), Van Frassen (1980), Field (1980), Rosen (1990), Noonan (1994), Crimmins (1998), Divers (1999), Melia(2000), Stanley (2001), Yablo (2001), Brock (2002) , Dorr and Rosen (2002), Joyce (2005), Nolan, Restall, and West (2005) for some of the proposals, endorsements, and explorations of various fictionalist views. See Rosen (1990), Brock (1993), Rosen (1994), Divers (1995), Hale (1995), Nolan (1997), MacBride (1999), Stanley (2001), Divers and Hagen (2002), Rosen (2005), et. al. for critical discussions.

<sup>5</sup> At least, there wasn't in 2007. Nowadays, there is plenty of great work on the topic, much of it included (or referenced to) in this volume.

<sup>6</sup> I allude here to Lycan's Moorean argument, which will be discussed in more detail below, section 4.5.

## 2. Fictionalism

Take your average metaphysician, Moe, who wrestles with following dilemma: Moe frequently engages in discourse, *D*, where he finds himself saying such things as (1)-(12):

- (1) There are fire trucks.
- (2) There are photons.
- (3) Tables exist, and they are made up of parts.
- (4) It is possible I could be riding a pink scooter right now.
- (5)  $2+2 = 4$ .
- (6) It is morally impermissible to hurt others.
- (7) Pegasus doesn't exist.
- (8) The average metaphysician worries about ontological commitment.
- (9) The sunset is at 7:52.
- (10) The cat is out of the bag.
- (11) Superman wears blue tights.
- (12) Wolverine is cooler than Superman.

We can assume that Moe utters (1)-(12) in all seriousness, and that he only utters what he takes to be true. Yet Moe is also a die-hard Quinean about ontological commitment: he believes that he is ontologically committed to those entities which his best theory says there are (Quine 1960, 1980). In addition, he believes that his everyday engagement in discourse *D* should reflect his ultimate theoretical commitments. He would like as little mismatch between his fundamental metaphysical beliefs and his ordinary talk as possible.

If Moe is a cautious metaphysician, easily spooked by the ontologically mysterious, we should already see his dilemma. Sentences such as (1)-(12) not only seemingly commit Moe to fire trucks and photons (which may not be so controversial), but also to composite objects, possibilia, mathematical objects, moral properties, fictional characters, the property of negative existence, the average *F*, setting suns, metaphors, and modal properties of fictional characters across fictions. It may not be clear where exactly Moe is having a problem—e.g., whether he cringes at the idea of composite objects, of moral properties, of non-existent objects, etc.—but if Moe is representative of a certain kind of metaphysician, Quinean at heart, then we can bet that he has a problem uttering at least one or more of (1)-(12), especially if he is in the throes of a philosophical discussion.

As explained in the introduction, one move to make in light of such a dilemma is to become a *fictionalist*. One kind of fictionalist strategy suggests: take the suspicious entity, *E*. Take all sentences, *S*, which seemingly ontologically commit one to *Es*. Reinterpret these sentences as sentences that are *not* so ontologically committing. Provide a brief explanation as to why these

sentences,  $S$ , are useful—and so should still be uttered—yet are nonetheless false (if taken at face-value).<sup>7</sup>

So, suppose Moe, like Hartry Field, thinks that our best, overall theory of the world can do without mathematical objects or numbers, but that it cannot do without mathematical discourse. Then he might opt for *mathematical fictionalism* (Field 1980). He may (re)interpret all of our mathematical talk as buffered by fictional-relative clauses or fictional operators. A sentence such as (5), for example

$$(5) 2+2 = 4$$

may be interpreted as (5<sub>f</sub>)

$$(5_f) \text{ In the mathematical fiction, } 2+2 = 4.$$

(5<sub>f</sub>) does not entail that there *is* a number, 2, or that there is a number such that when added to itself, yields another number, 4. It does not entail that there are numbers at all. Nor does it entail that there are functions or any kind of mathematical object. This is because (so the story goes) a sentence that is prefixed with an operator such as “In the fiction  $f$ ” is only making claims about what is or is not true *according to some fiction or other*. What is true in a fiction need not be true, *simpliciter*. In this way, being a fictionalist about mathematics allows one to talk the mathematical talk without walking the mathematical walk.

And so on for other fictionalist views, and Moe’s discomfort with any of (1)-(12). Moe might think, along with van Fraassen, that our best, overall theory of the world can do without unobservable entities, yet that talk of such entities is theoretically useful. If so, then he might be a fictionalist about unobservable entities (Van Fraassen 1980). Or he might think, like Gideon Rosen suggests, that our best, overall theory of the world can do without possible worlds, even though our modal knowledge relies indispensably on talking about them (Rosen 1990).<sup>8</sup> And so on. All of the sentences (1)-(12) express a seeming commitment to some entity or other that some philosopher has found ontologically abhorrent at some point or another. A fictionalist move is available in each case, which is a diagnosis by analogy.

---

<sup>7</sup> There are other ways to be a fictionalist, and other ways one might be motivated to accept a fictionalist position, than the Quinean/eliminativist reasons I propose here. See Kalderon (2005), e.g. Nonetheless, my interest is in mapping out an alternative for those who do find arguments for eliminative materialism compelling, yet also cannot stomach eschewing our folk psychological discourse as the eliminative materialist suggests.

<sup>8</sup> Rosen (1990) stops short of endorsing modal fictionalism, but he certainly succeeds in convincing many that it is a view worthy of serious philosophical consideration, even if false.

The fictionalist suggests:

“Look at sentences such as (8) through (12). These are representative of the kinds of sentences that most of us utter every day. All of us find ourselves talking about the average person, the average car, the average American, etc. We talk about sunsets and sunrises and often act as if we are in the midst of a geocentric universe. We use metaphors and idioms and illustrative, poetic language. We talk about comic book characters, literary characters, and people and things in movies and plays. We talk about what these characters do and what they are like and what they wear and who they are bested by on relevant grounds of comparison. When engaged in such talk, we somehow take our utterances to be true, but we don’t take our utterances as ontologically committing. We don’t get confused, for example, when someone says, “Barb has got Bob on a tight leash”; we don’t think that there really is a leash, which Barb has attached to Bob, and that it’s tight. So there is *some* way [insert details of your favorite account of fictional truth here] that we make statements such as (8)-(12) true but not literal (or true but not ontologically committing; or acceptable but not true but pragmatically appropriate and also not ontologically committing, etc). Whatever way *that* is, that’s the way we should treat seemingly ontologically committing sentences such as (4), (5), and (6). In other words, whatever is going on in sentences such as (8)-(10) to make them appropriate to utter is *just like* what’s going on with sentences such as (4)-(6). And *that* is how we can all keep our ordinary, everyday talk, yet still be good, rigorous Quineans.”

Or so claims the fictionalist.

But let us return to Moe, our metaphysician. In addition to sentences such as (1)-(12), Moe also utters sentences such as (13)-(15):

- (13) Joe believes there is beer in the fridge.
- (14) Jen desires some nachos.
- (15) Deb feels on the brink of a meltdown.

As before, we can assume that Moe utters (13)-(15) in all seriousness, and that he only utters what he takes to be true. We can also assume that he still upholds his Quinean standards for ontological commitment, and that he is interested in getting his ordinary, everyday utterances to match his fundamental metaphysical beliefs about what there is. However, let us also assume that Moe has been persuaded by many of the considerations for eliminative materialism. Clearly, Moe is in a dilemma: he would like to keep folk psychology (FP)—whether because FP is extremely useful, or because FP is too well-entrenched in our world view to give up, or because *it just cannot be* that there is such massive, widespread error about what’s going on in our heads (especially given the relative success of FP), etc. But he also thinks the eliminative

materialist has got unflappable grounds for her position. So what is Moe to do?

### 3. Mental Fictionalism

One way to solve Moe's dilemma is to turn to *mental fictionalism*—the idea that we are (or should be) fictionalists about the mental. Mental fictionalism (MF) will mimic the other fictionalist views already populating the philosophical landscape. Statements such as (13), (14), and (15), for example -

(13) Joe believes that there is beer in the fridge.

(14) Jen desires some nachos.

(15) Deb feels on the brink of a meltdown.

- might be recast as (13<sub>f</sub>), (14<sub>f</sub>), and (15<sub>f</sub>):

(13<sub>f</sub>) In the fiction *f*, Joe believes that there is beer in the fridge.

(14<sub>f</sub>) In the fiction *f*, Jen desires some nachos.

(15<sub>f</sub>) In the fiction *f*, Deb feels on the brink of a meltdown.

Like (5<sub>f</sub>), (13<sub>f</sub>) is true just in case according to the fiction *f*, Joe believes there is beer in the fridge. The difference, of course, is that in (13<sub>f</sub>) "in the fiction *f*" will pick out, not the mathematical fiction, but the *mental fiction*. And the mental fiction, in this case, will be *folk psychology*.

How do we do that? One way is to take ordinary folk psychology as a fictional story about how things are in our heads and treat it as we would any other fiction. Accordingly, all of our discourse about mental activity—beliefs, desires, first person qualitative feels, etc.—will *appear* as if we are making claims about the world. Instead, however, we have a nice way of (re)interpreting such talk as involving fictional operators such that no unsavory ontological claims are entailed by our ordinary talk about the mental.

One might initially worry about whether folk psychology can indeed be taken as a fiction *tout court*. After all, someone might argue, folk psychology is a product of the folk, and the folk are notoriously in wild disagreement about what exactly is going on—especially when the folk are holding an opinion about what is going on *inside our heads*.

Moreover, it is not as if any theory the folk produce will be as cohesive and determinate as, say, a *mathematical fiction*. That is, we can probably all easily agree on simple mathematical claims such as "2+2 = 4" and "there are infinitely many prime numbers greater than 11." There's not

too much disagreement here. Furthermore, we can easily rely on the experts to fill in the more difficult details of the mathematical story.

To use another example,<sup>9</sup> genuine fictions like the Sherlock Holmes novels have an uncontroversial resource for deciding whether certain claims are true or false in the fiction—we can open up the book and *look*. That there are certain facts about the fiction that are in print—e.g., that Sherlock Holmes lived at 221b Baker Street, that he had a sidekick named Watson, etc.—indicates that there is a particular, determinate fiction that we are all talking about when we use a prefix “In the Sherlock Holmes fiction.”

So how is folk psychology like the Sherlock Holmes fiction or the mathematical fiction? What could possibly qualify as the *folk psychological canon*—the uncontroversial, hardback equivalent to Conan Doyle’s novels or a collection of uncontroversial *a priori* mathematical truths?

First, it is important to keep in mind that both the *eliminative materialist* and the *intentional realist who accepts folk psychology* at least agree that folk psychology is a *theory*. One crucial difference between them is that the former believe that folk psychology is a bad, dispensable theory, while the latter think it is a good, *indispensable* one. So, given the dialectic at play here, there need not be an actual book in print delineating all of the nitty-gritty, agreed-upon details of folk psychology. Since both parties involved in the debate over the ontological status of the mental *agree* that folk psychology is a unified theory (Churchland 1981, Stich 1978, Horgan and Woodward 1985)—regardless of whether it has any other virtues apart from being unified—it is an easy step from there to take such a theory *as a fiction*. After all, it is plausible to think that a theory is a proposed story about how things are, in fact. So, we take the proposed theory of how things are, in fact, as a proposed theory of how things are *in a particular fiction*. That seems easy enough.<sup>10</sup>

Moreover, we do not want to have such strict standards for what counts as a fiction such that only those things that are bound or in print count. This would result in way too narrow of a

---

<sup>9</sup> I.e., you may not be convinced that mathematics *could* be a cohesive and determinate fiction.

<sup>10</sup> There may be some who deny that folk psychology (FP) is a unified theory. My suggestion here is that many of those who are already inclined towards some kind of mental eliminativism accept that (FP) is a theory, but deny that it is a good (true) one; this is who I have in mind when I suggest that they treat such a theory as a *fiction*. So my claim can be understood conditionally: *if* one already accepts that FP is a unified theory (even if false), *then* it should be easy matter to treat that theory as a fiction. Of course, the antecedent condition may be unsatisfied, and numerous alternative approaches are available. See Ratcliffe (2006) and (2007), for example, who denies that folk psychology is a unified theory, but also endorses a kind of fictionalist approach (different from the one I suggest here). See Hutto and Ravenscroft (2021) for an overview and discussion of options. Thank you to János Tózsér and Márton Miklós for comments on this section.

classification of fictions: it would leave out oral fictions and myths, fictions generated by pantomime, visual fictions, fictions created in the imagination without any external expression, and the like. The mental fictionalist (and all those in the debate over the ontological status of the mental) should have no trouble treating folk psychology as a unified theory. If so, it should be a short step from there to treat folk psychology as a fiction.

That's the easy part. The harder part is in recognizing that, once we lay out a basic strategy for mental fictionalism, there are many options for how the details of such a view could go. For example, following Stanley (2001), we may either endorse a *hermeneutic mental fictionalism* or a *revolutionary mental fictionalism*. A hermeneutic mental fictionalist claims that our discourse about the mental is *already* fictional. Folk psychology, in other words, just *is* a fiction, much like the world of Sherlock Holmes, and so statements about it—like statements about the world of Sherlock Holmes—are *already* ontologically innocuous. This is a descriptive view about our everyday use of folk psychology that will not require anyone to change how they talk, or how to interpret what they say. Contrast this with a revolutionary mental fictionalist, who claims that our folk psychological discourse is not in fact fictional, but *should* be. This is a prescriptive claim about how we *should* use a particular language.<sup>11</sup>

One could also choose between different interpretations of a semantic fictional operator, such as “In the fiction *f*” (or decide against a semantic operator altogether). Lewis (1978) suggests that the semantics of “In the fiction *f*” are ultimately understood in terms of concrete possible worlds. Since Lewis’ concrete modal realism is seen by many as ontologically extravagant, and since many of those who are sympathetic to fictionalism in the first place are motivated by general ontological thriftiness, I suspect most fictionalists will not want to opt for an analysis of the fictional operator in terms of possible worlds.

One way around this would be to cash out the semantics of the fictional operator in terms of syntactic or logical consequence (e.g., Field 1980). Another way would be to lean heavily on the attitudes we take towards certain sentences, making the distinction between *acceptance* and *belief*, where acceptance falls short of belief. Van Frassen suggests this in Van Frassen (1980). Of course, this approach will be (seemingly) problematic for mental fictionalism, since propositional attitudes are precisely the sorts of entities the mental fictionalist wants to fictionalist about. (I will discuss this potential problem below.)

Yet another way would be to endorse a kind of semantics/pragmatics distinction, and claim that

---

<sup>11</sup> Following Yablo (2001), one might also distinguish between different kinds of mental fictionalist views such as *instrumentalism*, *meta-fictionalism*, *object-fictionalism*, and *figuralism*. I won't explore these options in any detail here, but I do acknowledge the broad map of the theoretical space available. See Eklund 2007 for a summary and discussion of these alternatives for fictionalism in general (even if not applied to mental fictionalism in particular).



fictional contexts are ones where certain assumptions are pragmatically presupposed (Hinckfuss 1993). Finally, another way might be to rely on the distinction between talking *loosely* and talking *strictly*. Donald Baxter does this in Baxter (1988), where he is (arguably) endorsing a sort of fictionalism about the identity between parts and wholes. And no doubt there are other ways besides.

I do not wish to settle on any particular mental fictionalist account here; the details of a successful mental fictionalism account will have to wait until after we have discussed whether mental fictionalism *in general* is a viable philosophical position. My main purpose in the foregoing paragraphs is simply to outline some of the options for the mental fictionalist, making it clear that many varieties and variations are possible. Going forward, when I talk of *mental fictionalism* as a unified view, I mean the general strategy of taking our folk psychological talk as a fiction (in some way), not as literal statements that ontologically commit us to mentality. Any other details can, at this point and for my purposes in this paper, be left open.

#### **4. Advantages of Mental Fictionalism**

In this section, I canvass some of the advantages of mental fictionalism, both as a view of the mental and also as a fictionalist account in its own right. Sections 4.1-4.4 briefly canvass several advantages of mental fictionalism over *eliminative materialism*. Section 4.5 is an overview of the benefits of mental fictionalism over other fictionalist accounts. I raise and address some potential problems for mental fictionalism in section 5.

##### *4.1 Having Our Cake and Eating it, Too!*

One obvious advantage of mental fictionalism is that one can be an eliminative materialist at heart, yet still keep belief-desire talk around. Mental fictionalism proposes a paraphrase of our claims about beliefs and desires—a paraphrase that allows these statements to be *true* (or at least: not false). We get to keep all of our folk psychological talk, and there is no need to try and get rid of it, or think that it will, or should, ever go away. Sure, we have amended folk psychological talk *a little* by claiming that there are hidden fictional operators at play. But this is a minor adjustment. And a small price to pay.

This, of course, is the general advantage of any fictionalist account in metaphysics: it allows us to keep some useful discourse around, without taking on any onerous ontological commitments that such a discourse may seem to require.

##### *4.2 No Cognitive Suicide*

Another advantage of mental fictionalism is that it avoids a common argument against

eliminative materialism: the cognitive suicide objection (e.g., Boghossian 1990, 1991 and Baker 1987). The cognitive suicide objection runs roughly as follows.

“In order for eliminative materialists to even be able to argue against their opponent, or to argue for their position, they must *use* belief-desire talk. Communication, assertion, and argumentation *require* propositional attitudes. But the eliminative materialist conclusion is that all belief-desire talk is *false*—that is, that propositional attitudes do not exist! In order for eliminative materialists to argue for their view, they would suffer cognitive collapse; they would be committing cognitive suicide!”

In contrast, because the mental fictionalism *allows* us to keep belief-desire talk around (albeit *amended* belief-desire talk), it avoids such (first-pass) cognitive suicide objections. If some objector insists that communication and assertion require propositional attitudes, the mental fictionalist need not have any problem agreeing. It is just that the mental fictionalist has recast all of our propositional attitude talk as fictional talk, thus saving herself from literally quantifying over mental entities, but still retaining the ability to talk *as if* there were such things around. So, unlike the eliminative materialist, the mental fictionalist does not have to worry about committing cognitive suicide.

It is worth noting, however, that I am not convinced that the cognitive suicide objection is particularly effective against the eliminative materialist, as it seems to be a misinterpretation of the eliminativist position. It is true that the eliminative materialist will make ontological claims such as “there are no beliefs or desires,” and may proceed to give arguments supporting this conclusion. But the *act* of endorsing a statement such as “there are no beliefs” or the *act* of providing reasons in support of this conclusion does not entail that the eliminative materialist is subjecting herself to a cognitive collapse. The eliminative materialist does not think, for example, that absolutely *nothing* is going on when, according to folk psychology, we are saying something, advancing beliefs, proposing arguments etc. Even the eliminative materialist will admit that there is *some kind of cognitive or brain activity* occurring when folk psychology describes what we are doing as believing or desiring something. It’s just that, according to the eliminative materialist, our ordinary, everyday belief-desire talk is wildly disparate from the complicated cognitive activity that is in fact going on—so much so, that “belief” and “desire” fail to pick out any activity or process that’s actually in the world. But that doesn’t mean that there isn’t *any* cognitive activity occurring. And this cognitive activity can account for what’s going on when the eliminative materialist claims that there are no beliefs.

In fact, looking a bit ahead, there is going to be an analogous “cognitive suicide” worry for the mental fictionalist, which I think is similarly misguided (section 5.5). To see how one might think there is a problem, simply reflect on the fact that so many fictionalists rely on mental

activity—such as pretense or make-believe or our *attitudes* towards fictions, etc.—to motivate their particular brand of fictionalism. This will be problematic for mental fictionalism insofar as such activities are *mental* activities - if pretending, make-believing, imagining, etc, are the very sorts of things the mental fictionalist wants to be a fictionalist about. However, the mental fictionalist's response can echo the eliminative materialist here, so I think that this sort of self-refuting worry will ultimately be unsuccessful (see section 5.5 below and Wallace 2016).

Yet given that there are many who think that the “cognitive suicide” objection *is* devastating for eliminative materialism, let us leave this section as follows: *if* you think that eliminative materialism suffers from cognitive collapse, then one advantages of mental fictionalism is that is it not subject to this particular worry.

If you think that mental fictionalism may avoid this *particular* self-refuting worry, but that another kind of self-refuting worry awaits (of the sort I just mentioned), then hold tight. I come back to this issue in section 5.5.

#### 4.3 *Keeping Belief-Desire Talk in Explanation*

A third advantage of mental fictionalism over eliminative materialism is that mental fictionalism can maintain that belief-desire talk is indispensable for explanation and prediction. A claim such as (16), for example,

(16) Joe's belief that there is beer in the fridge explains why he just ran to the fridge.

would be rejected by an eliminative materialist. This is because (16) not only quantifies over 'Joe's belief', but also characterizes such a thing as the cause or explanation as to why Joe ran to the fridge. What it takes for (16) to be true, in other words, is (i) that Joe has a belief and (ii) that this belief caused him to run to the fridge. These are two things that the eliminative materialist will deny.

In contrast, for the mental fictionalist, (16) might get recast as:

(17) In the fiction *f*, Joe's belief that there is beer in the fridge explains why he just ran to the fridge.

And all it takes for (17) to be true is that *according to the folk-psychology fiction*, Joe's belief that there is beer in the fridge *explains* why he just ran to the fridge. In this way, the mental fictionalist can easily show how we get to keep all of the 'indispensable' belief-desire talk

around.<sup>12</sup>

Admittedly, things get a bit tricky here. This is because one might not want to accept that explanations such as (17) are causal, whereas the original (16) does seem to be causal. However, I see no reason why (17) can't be a causal explanation. Consider (18) and (19):

- (18) Superman got his cape stuck in the phone booth, and that explains why Jason was laughing.
- (19) According to the fiction, Superman got his cape stuck in the phone Booth, and that explains why Jason was laughing.

If (19) is the appropriate gloss on (18), then this is one example of how we can have a causal explanation *via* a fiction: it was the fact that something funny happened *in the fiction* that *caused* Jason to laugh.

More needs to be said here, of course. And we might wonder whether the mental fictionalist's gloss on causal and explanatory claims will ultimately be satisfactory. However, the mental fictionalist certainly has an advantage over the eliminative materialist here, and that is the primary point being made here.<sup>13</sup>

#### 4.4 Avoids Lycan's Moorean Objection

Yet another advantage of mental fictionalism over eliminative materialism is that it avoids Lycan's Moorean objection to the latter (Lycan 2005). Lycan proposes that, like Moore, we should modus-tollens any arguments for eliminative materialism. He claims,

"Numerous common-sense mental ascriptions, such as that Granny wants a beer and believes there is one under the sofa, are individually more plausible, and always will be more plausible, than are the purely philosophical premises of any argument designed to convince us to the contrary. As Moore saw, purely philosophical assumptions have very weak epistemic credentials and cannot by themselves outweigh simple common-sense facts... In order to reach the staggering conclusion that there has never been a belief, a desire, or any other propositional attitude, any argument for Eliminativism will have to rest on one or more a priori principles connecting scientific truths to negative ontology. And it is terminally unlikely that any such principle could be more credible for me than that Granny

---

<sup>12</sup> For a similar response to (two) different kinds of fictionalism, see Balaguer (1998). Balaguer (1998: p.811) argues that fictionalists can admit the indispensability of certain theories yet "simply account for this in fictionalist terms." He is discussing mathematical and semantic fictionalism in these papers, not mental fictionalism, but the same strategy can be applied here.

<sup>13</sup> Many thanks to Bill Lycan for discussion here.

wants beer. ” (Lycan 2005, p. 201)

Irrespective of whether this sort of objection is ultimately successful against the eliminative materialist, notice that such an attack is immediately avoided by the mental fictionalist. This is because the eliminative materialist’s conclusions such as “No one has ever believed anything” or “Granny has never wanted a beer” are *false* when properly amended by the mental fictionalist: according to the folk psychological fiction, lots of people *have believed* lots of things, and Granny *has wanted* a beer!

It might be objected that according to the mental fictionalist, it is true that no one has ever believed anything *strictly speaking*. That is, *absent the fictional operator*, the mental fictionalist must admit that no one has believed anything. And it is *this* conclusion that Lycan would probably attack in his Moorean way.

However, one of the advantages of mental fictionalism as I’m proposing it would be that, because our slip into fictional talk is often subtle, our ability to detect when we are talking fictionally and when we are not is not so obvious. So, for example, Lycan’s claim that he is more certain that the eliminative materialist’s conclusion is false than he is certain that any of the premises are true might be called into question.

The mental fictionalist might argue as follows: “your certainty that the eliminative materialist’s conclusion is false is fueled by the slip between the conclusion *without* the hidden fictional operator and the conclusion *with* the hidden fictional operator. Since the conclusion *with* the hidden operator—viz., “in the fiction *f*, no one has ever believed anything”—is clearly false, this might easily account for why you are so certain that the eliminative materialist’s conclusion must be false.” The burden of proof would then be on Lycan to show that he is *not* making a slip between the conclusion without the operator and the conclusion with it.

#### 4.5 General Fictionalist Worries: Arbitrariness, Consistency, and Incompleteness

As discussed at the start of this paper, there are many fictionalist accounts about many different kinds of entities. In response to many of these accounts, however, are several general problems such as arbitrariness, contingency, and incompleteness worries. I (briefly) argue below that mental fictionalism rivals other fictionalist accounts when it comes to addressing these worries; in some cases, mental fictionalism avoids them all together.

To begin, let’s return to another fictionalist view that we mentioned at the start of this essay: *modal fictionalism*. One of the problems facing modal fictionalism involves the fact that there

are several adequate candidates to serve as our possible worlds fiction. Should we take Rosen's advice and use Lewis's *On the Plurality of Worlds*? Or should we use an abstract, actualist model? Or something else entirely? There are several equally adequate, yet incompatible, candidates to stand in as our possible worlds fiction, which makes our choice between them somewhat arbitrary, and renders the resulting fictionalist view less plausible. A similar worry arises for *moral fictionalism*. There are many moral theories out there; which one are we (or should we be) appealing to when we make moral claims? Should we adopt the utilitarian story? Or the Kantian one? Or something else entirely? The more equally adequate, yet incompatible, candidates we have for the relevant moral fiction, the less plausible fictionalism in these areas appears.<sup>14</sup> Call this the *Arbitrariness Worry*.

Fortunately, mental fictionalism can avoid the arbitrariness worry. As mentioned in section 3, both the eliminative materialist and the mental realist agree that there is a unified, folk psychological theory. It is just that one of them thinks this is a good (true) theory, while the other thinks that it is a bad (false) one. Yet so long as they are in agreement that there is a *unified* theory—whether good or bad, true or false—the mental fictionalist can simply (non-arbitrarily!) settle on this *one* theory as her fiction.

Related to the above arbitrariness worries are concerns involving the seeming contingent nature of fictions. Fictions are man-made creations that, intuitively, need not have been created. Yet consider *modal* fictionalism. Modal fictionalism claims that something of the following sort holds: Possibly *p* iff according to the possible worlds fiction, there is a possible world where *p*. If the possible worlds fiction that we are using to analyze this bi-conditional is a *contingent* fiction (as so many, if not all, fictions are), then it seems that the following is true: if no one had ever been around to tell the possible worlds story, then nothing would have been possible (Nolan 2002). This result seems to cut sharply against our modal intuitions.

A similar problem arises for mathematical fictionalism. Mathematical fictionalism holds something like the following bi-conditional: *x* is 6ft tall iff according to the mathematical fiction, *x* has a height of 6 ft. Yet if the mathematical fiction employed is a contingent, man-made fiction, then the following counterfactual seems to hold: (i) If there had been no one around to tell the mathematical story, then nothing would be 6ft tall. This is certainly counterintuitive: even if one is an eliminativist about mathematical entities, one need not be (as the fictionalist in this case shows) an eliminativist about *predications* of height or measurements. And you need not think that, just because stories or fictions are contingent, that predications of height

---

<sup>14</sup> It should be noted that these "arbitrariness" worries are not necessarily *devastating* for either modal or moral fictionalism, but they are worrisome enough. That mental fictionalism can avoid them is certainly a bonus for the view.

and measurements are.<sup>15</sup>

If fictions are man-made, contingent creations, then any fictionalist view that relies on them will have to make sure that the consequences of such a reliance aren't problematic. However, in the case of modal and mathematical fictionalism, where the relevant fictions result in modal truth or mathematical predicates, respectively, we might legitimately worry that such reliance on contingent fictions *are* problematic.<sup>16</sup>

Yet mental fictionalism avoids this sort of worry. For the mental fictionalist, it won't matter that folk psychology is contingent—i.e., it won't matter that it is a theory that was artificially created by contingent beings who might not have existed, and who might not have created folk psychology even if they *had* existed. This is because nothing (intuitively) necessary hangs on the fiction. *Modal* fictionalism, in contrast, claims that modal truths hang on what is said in the possible world fiction.<sup>17</sup> But modal truths are (intuitively) necessary. If modal fictionalism is true, then our claims rely on something that isn't necessary, and - as a result - such claims themselves are not necessary. *Mental* fictionalism, however, doesn't make a parallel claim in this respect; as such, contingency worries are not a problem for the mental fictionalist.

Finally, a typical problem for other fictionalist views involves the fact that fictions are often incomplete. For example, Conan Doyle does not tell us what kind of soap Sherlock Holmes uses to wash himself, nor how many ounces of opium he smokes per week. That fictions are typically silent on certain matters or incomplete in this way generates a problem for some fictionalist views—in particular, modal and mathematical fictionalism. If the possible worlds fiction is incomplete, then there will be some modal truths that are unaccounted for, which would be problematic for modal fictionalism.

However, there is no parallel problem for the mental fictionalist, since it won't matter to her

---

<sup>15</sup> This contingency worry might seem less problematic for the mathematical fictionalist than it does for the modal fictionalist. For instance, you might think that the counterfactual *if there had been no one around to tell the mathematical story, then nothing would be 6ft tall* is less counterintuitive than *if no one had ever been around to tell the possible worlds story, then nothing would have been possible*. Fair enough. It doesn't really matter for my point here, so long as one of them seems problematic.

<sup>16</sup> This objection could be avoided if one is a Platonist about fictions. One could insist that the Sherlock Holmes stories exist somewhere in Platonic heaven, independent of Conan Doyle's existence, or his initiative to write such adventures down on paper. Of course, since most people are driven to fictionalist accounts because of their heightened susceptibility to the ontological heebie-jeebies, I doubt that many would be fond of this sort of maneuver. See Nolan (1997a) and (2002) for more discussion of this problem, which he calls "the problem from artificiality."

<sup>17</sup> This isn't true of all modal fictionalist accounts, but it is true of some. See Rosen (1990) and Nolan (1997) for more on the distinction between "strong" and "timid" modal fictionalism.

that folk psychology is incomplete. In fact, it is probably a bonus for the view that folk psychology is silent on some matters. You presumably don't have a belief—even according to folk psychology—about how many hairs I have on my head or whether I like to dip pickles in spicy mustard. (Well, you probably have a belief about these things *now*, but you didn't prior to me making such topics relevant.) The point is that there are numerous things that we have no opinion about, and numerous things we have neither a desire for nor a repulsion against. That folk psychology may be silent about such matters, then, seems to make mental fictionalism more plausible than other fictionalist in this respect.

## 5. Potential Problems for Mental Fictionalism

If mental fictionalism might be welcome in an area of inquiry such as philosophy of mind, where many have anguished over the ontological status of the mental, then why haven't there been any serious endorsers of mental fictionalism to date? One reason might be that there are several legitimate and somewhat obvious worries that have been anticipated for mental fictionalism, truncating at the start any chance for the view's success. I will discuss some of these below, with suggestions about how a mental fictionalist may respond.

### 5.1 *It's Phenomenologically Unmotivated!*

Here's one worry: the introduction of hidden fictional operators is unmotivated or ad hoc, and lacks any phenomenological support. In the case of genuine fictions, for example, there seems to be *prima facie* motivation for introducing fictional operators—namely, we have obvious conflicting intuitions about the truth-value of sentences like (11).

(11) Superman wears blue tights.

Away from philosophical discussions, a sentence such (11) may initially strike us as true. But then suppose someone says: "Hey, you say (11) is true, right? But guess what? I've gathered up all the blue-tighted beings and Superman is not among them!" At this point, we realize something is different about (11) - something that distinguishes it from other intuitively true assertions such as "this is a philosophy paper." There is a noticeable shift. And the suggestion that (11) is *true in a fiction* as opposed to *true simpliciter* becomes a plausible story of how to account for the phenomenological difference.

Not so with sentences involving folk psychology, one might argue. Suppose that Joe is running towards the fridge, uttering "Man, I could really use a beer right now!", reaching for the door handle, exclaiming "I'm so glad there's beer in the fridge - I'm super thirsty!", and so on. Given



Joe's behavior, and away from philosophical discussions, (13) seems true:

(13) Joe believes there is beer in the fridge.

But then suppose someone says: "Hey, you say (13) is true, right? But guess what? I've gathered up all the cognitive activity in Joe's brain and beliefs are not among them!" At this point, what do we do? Where's the noticeable *shift*? Even assuming the brain-facts are as the eliminative materialist suggests, where's the phenomenological difference between (13) and a sentence such as "this is a philosophy paper"? Should we find it at all plausible that (13) is *true in a fiction* as opposed to *true simpliciter*?

The mental fictionalist has a couple of responses available. First, she might appeal to the fact that very few other fictionalist accounts—e.g., modal, moral, mathematical fictionalism, etc.—are *prima facie* phenomenologically motivated either. All of these face the above worry, which has been called the Phenomenological Objection (Stanley 2001, Brock 2014, Eklund 2019). So even if this *were* a problem, the mental fictionalist could deny that it is a problem for her view alone. To the extent that someone found any of the other fictionalist accounts convincing, then so much the better for mental fictionalism, and so much the worse for our *prima facie* intuitions.

Also, recall that while sentences such as (11) or (12)

(11) Superman wears blue tights.

(12) Wolverine is cooler than Superman.

may have the phenomenological "stink" of fictional talk, other sentences such as (8)-(9) are much more subtle.

(8) The average metaphysician worries about ontological commitment.

(9) The sunset is at 7:52.

Since there are already plenty of *these* kinds of uncontroversial statements that do not admit of an obvious fictional "feel," and that nearly everyone (after some reflection) would likely be fictionalist (or fictionalist-like) about, the lack of a phenomenological shift in sentences such as (13) should not be worrisome.<sup>18</sup>

Secondly, however, the mental fictionalist could simply deny that our intuitions about sentences

---

<sup>18</sup> Compare Crimmins (1998)'s distinction between "shallow speaking as-if" and "imaginative play."

such as (11) and (12) really are all that obvious.

(11) Superman wears blue tights.

(12) Wolverine is cooler than Superman.

After all, it probably took some time (and perhaps even a philosophy course) to realize that a sentence such as (11) is any different from a sentence such as “Mikal Barishnikov wears blue tights.” For many, it isn’t until after we recognize that Superman is just a fictional character and not a real guy, that we start having second thoughts and phenomenological shifts about sentences such as (11).

Likewise with belief and desire talk. If we are *truly convinced* by some of the considerations for eliminative materialism (as was assumed at the start), then we will just be coming to “realize” that beliefs and desires (as described by folk psychology) simply do not map on to reality. And if we *have* come to this conclusion, then perhaps we will have the same sorts of conflicting intuitions concerning sentences such as (13)

(13) Joe believes there is beer in the fridge.

as we do with (11) and (12). It may be subtle, and we may admit that not much hangs on it at the end of the day. But if we do assume that the brain-facts are as the eliminative materialist says, then perhaps we will also admit of a subtle phenomenological shift when we compare sentences such as (13) against more literally-true-sounding claims such as “this is a philosophy paper.”

### 5.2 The Phlogiston Worry

Here’s another worry: if we can save belief-desire psychology by appealing to a hidden fictional operator, what’s to stop us from doing the same in the case of phlogiston? Why not say that all phlogiston-talk also has hidden operators such that we get sentences like: in the *phlogiston* fiction, *p*. If so, then it seems we have rampant fictionalism about all sorts of archaic theoretical entities—which seems not only unnecessary and excessive, but implausible, semantically speaking. It is pretty unlikely that there are really *that* many cases of hidden fictional talk!

While a variety of responses are available here, I suggest that the mental fictionalist simply bite the bullet here and admit that, yes, phlogiston-talk does—or, rather, *did*—include hidden fictional talk. The important difference, however, between phlogiston-talk and belief-desire talk is that *we no longer have any use for phlogiston-talk*. So there’s no need to worry overly much about it.

Fictionalists are drawn to their position by at least two things: (i) eliminativist leanings about some class of entities, *E*, and (ii) acknowledgement that certain discourses that quantify over *Es* are *useful*. To the extent that such talk stops being useful, that's the extent to which a fictionalist stops being a fictionalist and starts being an eliminativist. It is relevant, then, that phlogiston-talk has ceased to be useful, for it means that we no longer need to cushion such talk with fictional operators. But if it *were* useful, I see no problem with a fictionalist treating such talk like any other fiction: it should be no different than our talk about Superman, sunsets, the average metaphysician, etc.

### 5.3 This is Just Dennett's Instrumentalism!

A more substantial objection than the previous two—and one that is particular to mental fictionalism—is the charge that mental fictionalism is nothing more than Dennett's instrumentalism (Dennett 1978 and 1987). To the extent that one can accurately pin down Dennett's view,<sup>19</sup> it seems to go something like this: for any entity whatsoever, if its behavior can accurately and reliably be predicted by taking the intentional stance towards it, and if the predictive power gained by taking such a stance is an advantage over alternative stances (e.g., a physical or design stance)<sup>20</sup>, then it *is* an intentional system. That is, it is a system that we can *correctly* attribute beliefs and desires *to*. To put it another way, if attributing beliefs and desires to a system helps in predicting its behavior, then it is appropriate to claim that such a system *has* beliefs and desires.

First, it should be noted that there is precedence in the literature for distinguishing instrumentalism (in general) from fictionalism (in general) (Nolan 2002). This distinction is grounded in part on the fact that instrumentalists typically claim that the statements that they are instrumentalist about *lack* a truth-value. This is not the case with fictionalists. Indeed, as we've already discussed, one of the primary motivations for fictionalism (in general) is to *keep* our ordinary talk rather than eliminate it. And, of course, we want to keep our talk because we think it is (for the most part) *true*. Or else: we want to refrain from attributing massive, widespread error. According to Dennett's instrumentalism, in contrast, it is not so clear whether our ordinary belief-desire ascriptions admit of a definitive, unambiguous truth-value. After all, whether a system can legitimately be ascribed beliefs and desires depends on someone *else* taking an intentional stance *towards it*. In this respect, mental fictionalism differs from Dennett's Instrumentalism.

---

<sup>19</sup> See Stich's (1981) criticisms of the slipperiness of Dennett's position. I'm not endorsing this criticism, just acknowledging it.

<sup>20</sup> This caveat is added to avoid "lectern" objections. (Dennett 1987, p. 23).

In addition, the mental fictionalism I have outlined here is making a semantic, as well as an ontological, claim. Mental fictionalism (as suggested here) is the proposal that (i) strictly speaking, there are no beliefs and desires, and (ii) talk that (seemingly) commits us to the contrary is really just fictional talk. This second point is a positive, semantic claim that distinguishes it importantly from Dennett's Instrumentalism.<sup>21</sup>

Another difference to consider is that for Dennett, so long as taking an intentional stance towards a system results in predictive power with respect to that system, it is legitimate to claim that that system *has* beliefs and desires. However, this leads to somewhat counterintuitive results, such as admitting that thermostats, plants and even flashes of lightning have mental states (Dennett 1987, pp. 21-31). In contrast, mental fictionalism relies on what the folk say - i.e., folk psychology - as its fiction. The folk do not usually attribute beliefs and desires to things like thermostats, plants, or flashes of lightning, so the mental fictionalist will not be able to say—even within the scope of the fictional operator—that such things *have* beliefs and desires. The most that a mental fictionalist may be willing to admit is that we sometimes anthropomorphize certain objects—such as thermostats, plants, and lightning—just for fun, or as a playful joke, or for whatever reasons motivate us to speak metaphorically. But she might reasonably insist that this kind of playful talk is importantly different from normal belief-desire attribution, since in the former (she might argue) is intuitively false, whereas the latter is not.

Finally, one of Dennett's requirements for taking an intentional stance towards a certain system is that we must assume that the system is ideally rational. That is, in order to figure out whether a certain system has beliefs and desires, we should see what it *ought* to believe or desire, under the assumption that it is ideally rational, and only then do we predict its behavior. Admittedly, it is not uncontroversial whether this "ideally rational" requirement ultimately proves problematic for Dennett's account (Stich 1981, Dennett 1991). The point to be made here is that mental fictionalism is not subject to such worries. Mental fictionalism relies on what folk psychology says about mental states, and folk psychology does not assume that intentional agents are ideally rational.

#### *5.4 Mental Fictionalism is Not Novel*

Another objection particular to mental fictionalism is the charge that this view is no different than being a fictionalist about propositional attitudes or attitude reports, which has already

---

<sup>21</sup> Notice, too, that even if one is a different sort of mental fictionalist - for example, one who didn't introduce fictional operators, but relied on conversational presuppositions to introduce the fictional context - it will *still* differ from Dennett's Instrumentalism, since it is appealing to fictional contexts to justify (seemingly) false assertions, and Dennett does not depend on such a move.

been proposed and discussed (Crimmins 1998, Balaguer 1998).

Crimmins (1998) suggests that we adopt a pretense account for propositional attitudes that seemingly use “modes of presentation.” His account is motivated by a puzzle about identity statements—in particular, Frege puzzles, involving sentences such as “Hesperus is Phosphorus.” According to Crimmins, whenever we seemingly distinguish between several different modes of presentation of a single thing, we talk *as if* the one object is many. His view might be dubbed “Mode of Presentation Fictionalism.”

Balaguer (1998), in contrast, proposes what he dubs “semantic fictionalism.”<sup>22</sup> This is a fictionalism about propositions in particular, which is committed to the following three claim: “(a) platonists are right that ‘that’-clauses purport to refer to propositions, but (b) there are no such things as propositions, and (c) ‘that’-clause-containing sentences...are not true—they are useful fictions.” (Balaguer 1998, p. 805)

To motivate this view, Balaguer draws an analogy between mathematical fictionalism and semantic fictionalism, and then canvasses some arguments against the existence of abstract entities in general. One line of reasoning Balaguer gives runs roughly as follows: since propositions are purportedly abstract, they are causally inert. Yet propositional attitudes are assumed to be a relation between an individual and a *proposition*. If propositions are causally inert, then they cannot have a causal effect on anything physical; so, they cannot have a causal effect on an individual or an individual’s belief state. Further, abstract entities such as propositions are unnecessary, making a nominalistic world-view more likely. Finally, we can account for the illusory reference to abstracta such as propositions by a fictional account. Thus, semantic fictionalism is coherent and plausible.

While it may be obvious to some already, let me nonetheless lay it out explicitly: there are at least three ways in which the mental fictionalist account presented here differs from Crimmins’ (1998) proposal, and at least two ways in which it differs from Balaguer’s (1998) account. Let me run through these differences briefly.

First, Crimmins’ account is specifically a fictionalism about propositional attitudes that seemingly involve “modes of presentations.” Crimmins states that “our talk about what people say and think *often* involves semantic pretense.” (Crimmins 1998, p. 9) So, in Crimmins’ view, not *all* of what people say and think is going to be talk *as-if*, or fictional talk. In contrast, mental fictionalism (as I’ve been presenting it) claims that what people say and think *always* involves

---

<sup>22</sup> Balaguer does not endorse semantic fictionalism; he merely thinks that it is an alternative to Platonism in light of Frege-Bealer-Schiffer arguments for Platonism. (Balaguer 1998, p. 810).

fictional talk. That is, *all* propositional attitudes are products of the folk psychological fiction, and so *all* of them (including those not presented in multiple modes of presentation )are talk *as-if*.

More succinctly: Crimmins' account is specifically a fictionalism about mode of presentation reports; mental fictionalism is specifically a fictionalism about mental entities and folk psychology. Second, Crimmins is motivated by Fregean puzzles of identity statements. Mental fictionalism is motivated by a more general worry—the overwhelming success of neuroscience and the highly credible arguments for eliminative materialism. Finally, Crimmins makes heavy use of pretense and games of make-believe. I have downplayed this aspect of mental fictionalism. The exact reasons for my doing so will be laid out below, in section 5.5, but it is enough for our purposes here to recognize that very little emphasis on *pretending* or *make-believe* need be made for a mental fictionalist view, in contrast to Crimmins.

Mental fictionalism differs from Balaguer's semantic fictionalism in at least two ways. First, Balaguer's account is specifically a fictionalism about propositions and abstracta; mental fictionalism is a fictionalism about mental entities and whatever problematic intentional entities folk psychology seemingly quantifies over. Traditionally, the mental entities of folk psychology are not abstract—they may be spooky and weird because they are ghost-like, dualist-y sorts of things—but they are *not* abstract.

More importantly, they are not non-causal, which is the one particularly unsavory feature of abstracta that Balaguer targets. This leads us to the second distinction between Balaguer's semantic fictionalism and mental fictionalism—Balaguer's view is motivated by an abhorrence of abstracta; mental fictionalism is (again) motivated by the overwhelming success of neuroscience and the highly credible arguments for eliminative materialism.

So, looking carefully at the details of the views, and the particular worries motivating them, we can easily distinguish mental fictionalism from related (but different!) fictionalist accounts about modes of presentation or semantics.

### 5.5 Cognitive Suicide - Again

I turn now to one last problem. On almost all contemporary fictionalist accounts, most noticeably beginning with Walton (1985) and (1990), much is made out of the *act* of pretending or games of make-believe to generate support for fictionalism as a philosophical position. Walton (2000), for example, talks about a game where children pretend that the bikes in a garage are horses in a corral. A fiction's content is often generated by real world facts; the

number of bikes in the garage, for example, may dictate how many horses we are imagining are in the corral. After drawing our attention to imaginative games, it is then proposed that we often engage in a similar kind of pretense all of the time—e.g., we use metaphors and idioms, we talk about the average man and sunsets, etc. Then, depending on which fictionalism that is being endorsed, arguments are made to show how we might be engaged in games of make-believe or pretense in a particular kind of discourse—one involving mathematics, possible worlds, morality, etc. While the details of just how this is done vary from account to account, the common thread through fictionalist accounts is the endorsement of a distinction between fictional talk and non-fictional talk that *relies on mental attitudes*. We either engage in *imaginative* pretense or play games of make-believe or *imagine* that things are thus-and-so or *believe as if* things are a certain way, etc. Fictional accounts seem to rely crucially on our mental attitudes, and in particular *beliefs*. Yet this is the very thing that the mental fictionalist is proposing we are fictionalists about.

Analogous to the cognitive suicide worry launched against the eliminative materialist (discussed above in section 4.2), is the following objection against mental fictionalism:

“In order for there to be a distinction between fictions and non-fictions, or fictional-talk and non-fictional talk, we must be able to make sense of the act of pretending or make-believe. But these acts *presuppose* that folk psychology is *not* a fiction—i.e., they presuppose that there actually *are* beliefs and desires. Yet strictly speaking, the mental fictionalist thinks that all belief-desire talk is *false*—i.e., that propositional attitudes do not exist! According to the mental fictionalist, we only *pretend* that there are beliefs, desires, and other mental states when we are *engaged in the folk psychological fiction*. Yet it is unclear what *being engaged in folk psychological fiction* could possibly mean without appeal to mental activity. Engagement in fictions - pretending, imagining, behaving *as-if*, etc. - are activities that presuppose realism about mental activity or propositional attitudes, the very things that the mental fictionalist (when speaking strictly) denies. *This is cognitive collapse.*”

As I mentioned in 4.2, I think that this objection—like the cognitive suicide argument launched against the eliminative materialist—is a misguided one.

To see this, first consider how the eliminative materialist might answer the following (admittedly bad) modus tollens argument against her view:

“If eliminative materialism is true, then (i) there are no beliefs and desires, (ii) folk psychology is false, and (iii) we are all subject to massive, widespread error. As a consequence, there can be no genuine distinction between when we legitimately take something to be true and when we are simply *pretending as-if* it is true. If there are no beliefs or desires, then there is nothing to distinguish the mental activities of imaginative

make-believe from believing something in earnest; there is no difference between telling a tall tale and telling it straight; there is no difference between fiction and non-fiction. All of these distinctions presuppose that there are indeed mental states, which the eliminative materialist denies. This is unacceptable. So, eliminative materialism is false.”

The eliminative materialist’s response to this sort of objection will repeat her response to the cognitive suicide objection (section 4.2). She will explain that eliminative materialism does not think that *absolutely nothing is going on* when, according to folk psychology, we are pretending, make-believing, sincerely avowing, etc. There *is* cognitive activity when we are (seemingly) doing these things, she admits. It’s just that our ordinary, everyday pretense (make-believe, sincere-avowal) talk is wildly disparate from the cognitive activity that *is* going on—so much so, that terms such as “pretense”, “make-believe”, “sincere avowal”, and the like, fail to uniquely pick out any brain activity that is in fact occurring. But that doesn’t entail that there isn’t *any* cognitive activity occurring. And, the eliminativist may insist, it is *this* cognitive activity that can account for what’s going on when folk psychology distinguishes between, say, make-believing and sincere-believing.

In a similar way, the mental fictionalist can make use of the cognitive activity that is (strictly speaking) going on whenever folk psychology describes us as being engaged in mental activity. Like the eliminative materialist, the mental fictionalist doesn’t think that absolutely *nothing* is happening in the world when someone claims to be make-believing or sincere-believing. There *is* cognitive activity when these things are seemingly going on. It is just that our ordinary, everyday pretense (make-believe, sincere-avowal) talk is wildly disparate from the cognitive activity that *is* going on—so much so, that terms such as “pretense”, “make-believe”, “sincere avowal” , and the like, fail to uniquely pick out any brain activity or process that is in fact occurring. However, the advantage of mental fictionalism is that it allows one to still *talk as-if* these terms do pick out something, even though this “talk as-if” will ultimately (and strictly speaking) get cashed out in terms of some sort of complicated cognitive activity that is unsuitable as a legitimate element of folk psychology. In this way, then, mental fictionalism, like eliminative materialism, can evade cognitive collapse worries.<sup>23</sup>

## 6. Concluding Thoughts

My aim here is primarily to put the outlines of a view on the map. I hope I have shown that mental fictionalism is a viable alternative for those initially sympathetic to the arguments for eliminative materialism, but unwilling to accept the radical claims of such a view. Being a fictionalist about mentality has some advantages over eliminative materialism, as well as some

---

<sup>23</sup> See Joyce (2013), Parent (2013), Wallace (2016) for further discussion on this issue.



advantages as a fictionalist view in its own right. I have attempted to show how mental fictionalism can circumvent certain worries, or - at least! - how it can start to circumvent those worries, with the promissory note that much more needs to be said. It remains to be seen—and remains a further project—whether mental fictionalism will rival alternative views about the mind or mentality, and also whether such a view will ultimately withstand further and deeper scrutiny on its own. Until then, I hope this has been a solid first step.

## References

- Baker, L. (1987). *Saving belief*. Princeton: Princeton University Press.
- Balaguer, M. (1998) "Attitudes without propositions" *Philosophy and Phenomenological Research*, 98: 805-26.
- Boghossian, P. (1990) "The status of content", *Philosophical Review*, 99: 157–84.
- Boghossian, P. (1991) "The status of content revisited," *Pacific Philosophical Quarterly*, 71: 264–78.
- Brock, S. (2002) "Fictionalism about fictional characters", *Noûs*, 36: 1-21.
- Burgess, J. and G. Rosen (1997) *A subject with no object*. Oxford: Clarendon Press.
- Churchland, P. M. (1981) "Eliminative materialism and the propositional attitudes, *Journal of Philosophy* 78: 67-90.
- Churchland, P. M. (1988). *Matter and consciousness* (revised edition). Cambridge, MA: MIT Press.
- Crimmins, M. (1998) "Hesperus and Phosphorus: sense, pretense, and reference", *Philosophical Review*, 107: 1-48.
- Dennett, D. (1987) *The Intentional Stance*. Cambridge, MA: MIT Press.
- Dennett, D. (1991) "Real Patterns" *The Journal of Philosophy* 88(1): 27-51.
- Divers, J. and J. Hagen (2006) "The Modal Fictionalist Predicament", in F. MacBride (ed.),

*Identity and Modality*, Oxford: Oxford University Press.

Dorr, C. and G. Rosen (2002) "Composition as a fiction", in R. Gale (ed.), *The Blackwell Guide to Metaphysics*, Oxford: Blackwell.

Eklund, M. (2005) "Fiction, indifference, and ontology", *Philosophy and Phenomenological Research*, 71: 557-79.

Everett, A. and T. Hofweber (eds.) (2000) *Empty names, fiction and the puzzles of non-existence*. Stanford: CSLI Publications.

Field, H. (1980) *Science without numbers*. Princeton: Princeton University Press.

French, P. and H. Wettstein, (eds.) (2001) *Midwest studies in philosophy volume XXV: figurative language*. Oxford: Blackwell.

Hinckfuss, I. (1993) "Suppositions, presuppositions, and ontology", *Canadian Journal of Philosophy*, 23: 595-618.

Horgan, T. and Woodward, J. (1985). "Folk psychology is here to stay", *Philosophical Review* 94: 197-226.

Hutto, D. and I. Ravenscroft, "Folk psychology as a theory", *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/fall2021/entries/folkpsych-theory/>

Joyce, R. (2005) "Moral fictionalism", in Kalderon (ed.) *Fictionalism in metaphysics* (2005) pp. 287-313.

Kalderon, M., (ed.) (2005) *Fictionalism in metaphysics*, Oxford: Clarendon Press.

Lewis, D. (1978) "Truth in fiction", *American Philosophical Quarterly*, 15: 37-46.

Lycan, W. (2005) "A particularly compelling refutation of eliminative materialism" in D.M. Johnson and C.E. Erneling (eds.), *Mind as a Scientific Object: Between Brain and Culture* (Oxford: Oxford University Press, 2005), pp. 197-205.

MacBride, F. (1999) "Listening to fictions: a study of fieldian nominalism", *British Journal for the*

*Philosophy of Science*, 50: 431-55.

Márton, M. and J. Tőzér (2013) "Mental fictionalism as an undermotivated theory", *The Monist* 96 (4): 622–638.

McCormick, P., (ed.) (1985) *The reasons of art/L'art a ses raisons*. Ottawa: University of Ottawa Press.

Melia, J. (2000) "Weaseling away the indispensability argument", *Mind*, 109: 453-79.

Nolan, D. and J. O'Leary-Hawthorne (1996) "Reflexive fictionalisms", *Analysis*, 56: 26-32

Nolan, D., G. Restall and C. West (2005) "Moral fictionalism versus the rest", *Australasian Journal of Philosophy*, 83: 307-30.

Parent, T. (2013) "In the mental fiction, mental fictionalism is fictitious." *The Monist* 96 (4): 605–621.

Quine, W. v. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Ratcliffe, M. (2006) "'Folk psychology' is not folk psychology", *Phenomenology and the Cognitive Sciences*, 5(1): 31–52. doi:10.1007/s11097-005-9010-y

Ratcliffe, M. (2007) *Rethinking commonsense psychology: a critique of folk psychology, theory of mind and simulation*, London: Palgrave Macmillan UK. doi:10.1057/9780230287006

Rosen, G. (1990) "Modal fictionalism", *Mind*, 99: 327-54.

Rosen, G. (1994) "What is constructive empiricism?", *Philosophical Studies*, 74: 143-78.

Rosen, G. (2005) "Problems in the history of fictionalism", in Kalderon (ed.) *Fictionalism in metaphysics* (2005), pp. 14-64.

Stanley, J. (2001) "Hermeneutic fictionalism", in French and Wettstein (eds.) *Midwest studies in philosophy volume XXV: figurative language* (2001), pp. 36-71.

Stich, S. (1981) "Dennett on intentional systems" in *Philosophical Topics* 12(1): 39-62.

Van Fraassen, B. (1980) *The scientific image*, Oxford: Oxford University Press.

Wallace, M. (2016) "Saving mental fictionalism from cognitive collapse", *Res Philosophica* 93(2): 405-424.

Walton, K. (1985) "Fictional entities", in McCormick (ed.) *The reasons of art/L'art a ses raisons*. (1985), pp. 403-13.

Walton, K. (1990) *Mimesis and make-believe*, Cambridge, Mass.: Harvard University Press.

Walton, K. (2000) "Existence as metaphor?", in Everett and Hofweber (eds.) *Empty names, fiction and the puzzles of non-existence*. (2000), pp. 69-94.

Woodbridge, J. (2005) "Truth as a pretense", in Kalderon (ed.) *Fictionalism in metaphysics*(2005), pp. 134-77.

Yablo, S. (2001) "Go Figure: A Path Through Fictionalism", in French and Wettstein (eds.) *Midwest studies in philosophy volume XXV: figurative language* (2001), pp. 72-102.