Abstract: Mental fictionalism maintains that: (i) folk psychology (FP) is a radically false theory, but (ii) we should nonetheless keep using it, because it is useful, convenient, or otherwise beneficial to do so. We should (or do) treat FP as a useful fiction—false, but valuable. Some argue that mental fictionalism is incoherent: if a mental fictionalist rejects FP then she cannot appeal to fictions in an effort to keep folk psychological discourse around, because fictions presuppose FP. In this paper, I defend the mental fictionalist against such objections. Whatever disadvantages mental fictionalism may have, it need not suffer from 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. 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) Kahleesi has three dragons.

Your utterance is acceptable—perhaps even true—even though you do not think dragons (or Kahleesi) 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 phenomena 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—e.g., in our conversations about numbers, unobservable entities, possible worlds, etc.

Suppose, for example, that you are a 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) Kicking puppies is morally wrong.

One way to reconcile the acceptability of (2) with a metaphysics that does not quantify over moral properties is to claim that an utterance of (2) is similar to an utterance of (1). Both seemingly

1 Many thanks to Sam Cowling, Ted Sider, Keith Simmons, and Tim Sundell for helpful comments, feedback, and discussion.

2 Hermeneutic vs. revolutionary fictionalism. See Burgess (1983), e.g.

3 Where an X-committing discourse existentially quantifies over Xs.

4 See Nolan, Restall, and West (2005), e.g.
involve ontologically suspicious entities—dragons and moral properties, 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 \( D \), (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.\(^5\)

Philosophers have been fictionalists about fictional characters and content\(^6\), possible worlds\(^7\), mathematical objects\(^8\), scientific unobservable entities\(^9\), moral truths\(^10\), composite objects\(^11\), truth\(^12\), etc.

Many find fictionalism attractive because it allows us to continue talking as if there are Xs, even though we (strictly speaking) 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, aesthetic valuable, it may be more expressive to do so, or psychologically and sociologically convenient, etc.

Motivating moral fictionalism, Nolan, Restall, and West (2005) 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 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: ?)

---

\(^5\) See Yablo (2001), for example, and his distinction between Instrumentalism, Meta-fictionalism, Object-fictionalism, and Figuralism.

\(^6\) Lewis (1978), e.g.

\(^7\) Rosen (1990), (1993), and (1995), and Nolan (1997a) and (2002).

\(^8\) Field (1980) and (1991).

\(^9\) van Fraassen (1980).

\(^10\) Nolan, Restall, and West (2005).


\(^12\) Burgess and Burgess (2011), Woorbridge (2005).
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. 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 has reason to doubt that they do. 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.

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, propositional attitudes, etc. 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 mental states and activities, then any eliminative materialist who denies folk psychology 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.

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.

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

---

13 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—Churchland P.M. (1981), Churchland P. S. (1986), e.g. I take mental eliminativism to be a broader category, of which eliminative materialism is a type. See section 2.1.

14 “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)

15 This is similar to Baker’s ‘cognitive suicide’ argument against eliminative materialism in Baker (2005) See Parent (2013) and Sprevak (2013) for variations of this argument against mental fictionalism.
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.\(^\text{16}\) The ontological thesis is that entities of a certain kind in a given discourse do not exist.\(^\text{17}\) 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 thesis will be relevant to understanding mental fictionalism, I discuss each in turn.

2.1 Mental Eliminativism

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

\[
\text{(NegEx)} \quad \text{Some of the entities posited by folk psychology (FP) do not exist.}\(^\text{18}\)
\]

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:

\[
\text{(DualOnt)} \quad \text{Some of the entities posited by folk psychology presuppose a dualist ontology.}
\]

We might extend Feyerabend’s reasoning as a version of Mackie’s argument form queerness against moral properties. Mackie (1977) maintains that our commonsense moral notions imbue moral facts or properties with 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...

\(^{16}\) Nolan, Restall, and West (2005), Eklund (2011), Parent (2013), et. al., for example.

\(^{17}\) Or that they may not exist, and we thus should—for whatever epistemic reasons—be agnostic about whether they do or not exist.

\(^{18}\) This is a variation of Parent’s (Elim) in Parent \textit{ibid.}
for the truth of our moral statements, in which case both FP and moral discourse should be jettisoned.\(^\text{19}\)

Quine (1960) and parts of Churchland (1981), however, 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.

Take, for example, a theory of the world that claims only the following: *everything is the way it is because something else made it so.* This may be true, but it is a terrible explanation. It is trivial and oversimplified. Analogously, one might argue that folk psychology is a less accurate or detailed description than, say, neuroscience.\(^\text{20}\) If so, you might endorse:

\[
\text{(Vague)} \quad \text{Some of the entities posited by folk psychology are too vague, sloppy, indeterminate, or inaccurate to pick out anything real.}
\]

Another eliminativist position 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:

\[
\text{“‘sensation’ might lose its reporting role as well as its explanatory role, just as ‘demon’ had lost both its roles, and that both of these roles might be taken over by reference to brain-processes.” (1970: 112)}
\]

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

\[
\text{“My claim, then, is not just that the various technical or theoretical concepts of qualia are vague or equivocal, but that the source concept, the ‘pre-theoretical’ notion of which the former are presumed to be refinements, is so thoroughly confused that, even if we}
\]

---

\(^{19}\) 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.

\(^{20}\) “…FP is \textit{at best} a highly superficial theory, a partial and unpenetrating gloss on a deeper and more complex reality.” Churchland (1981: 74)

“A serious advance in our appreciation of cognitive virtue would thus seem to \textit{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.” Churchland (1981:84)
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.” (1988: 382-3)

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.21

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 might admit that folk psychology is rife with explanatory failures, infertility, 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, for example, may hold this. 22 Ramsey, Stich, and Garon 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

21 See Marton and Tozser (2013), for example, who 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.

22 Ramsey, Stich , and Garon (1990), e.g.
distinct, identifiable sub-structures that are semantically interpretable as representations of individual propositions.” (1990: 514)

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.  

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.

Or even stronger still:

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

(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 (à 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, etc. 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, or 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 intensionality,

---

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

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

25 Boghossian (1990), for example, argues that certain mental eliminativists are committed to something like (NoRep).
representation, or propositional content across the board, then it will be inconsistent to also maintain that we should be fictionalists about the mental—where fictions are 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, 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 and 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.” (Lewis 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.” (Yablo (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), perhaps (NoQualia), etc.

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.

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 technically be coherent in that it is not obviously contradictory, it is incoherent in another way: it is silent on the one aspect of the theory that a mental fictionalist surely needs to have an account of.

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

A mental fictionalist could, like Field or van Fraassen, maintain 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.” (van Frassen (1980: 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, having qualitative reactions to it, etc. 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 correct—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.

But, of course, 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.

We might (α la 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.

---

26 Yablo (2001: 75) attributes this view to van Frassen (1980).
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.

But, again, this may not be problematic for certain of the mental eliminativisms detailed above—in particular, (DualOnt), (Vague), (NoQualia), and (BadTheory). Moreover, in order for this point to be a challenge for mental fictionalism in general, it must be shown that a sentential operator such as according to the folk psychological fiction is indeed 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 in 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.

But, you might think, it requires some kind of representation, since possible worlds (for Lewis) are ways of representing how the world could be.27

Lewis claims the following about de re modality, for example:

   “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.”
   (Lewis (1986: 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 museum analogy to the contrary—does not so obviously involve mentality. An other-worldly tree with 6 branches may represent a this-worldly tree with 5 branches (making it the case that this-worldly tree could have had 6 branches), even if there had never been any minds on any world whatsoever. So a mental eliminativist—including one who endorses (E7)—may allow that other-worldly objects represent this-worldly objects. It’s just that the representation here is of an innocuous sort. But, admittedly, the plausibility of this line may depend on the strength of one’s interpretation of (NoRep).

27 Thanks to Ted Sider for discussion in this section.
Alternatively, one could take a Lewisian-like analysis of the \textit{according to the fiction, $F$} operator, diverging from Lewis only where \textit{de re} modality is concerned. Or one could take a purely formal analysis. Suppose, for example, one accepts: In the fiction $F$, $P$ if and only if $F$ logically entails $P$. Such an analysis does not involve intensionality; on the contrary, it is presumably involves only a logical relationship.

True, such an analysis of fictional content is oversimplified. 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-intensional) is nonetheless a plausible option for the mental fictionalist, even if she is then under some pressure to provide a more thorough account.

In other words, 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-intensional sense. So it is plausible that the mental fictionalist has an account of the fictional operator that is logical, not intensional. In which case, the mental fictionalist will not suffer cognitive collapse.

Of course, all of this is assuming that logical relations are themselves non-intensional. But let us save further discussion of these topics for section 3.

3. \textbf{Defense Against Cognitive Collapse}

It was stated previously that if one is motivated towards the kind of mental eliminativism that rejects any intensional content, representation, etc., then it would be incoherent to adopt a representational fiction as a way out of making a commitment to intensionality—one has just swapped a discourse riddled with intentionality (i.e., folk psychology) for another (i.e., 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:

$$(\text{DualOnt}) \quad \text{Some of the entities posited by folk psychology presuppose a dualist ontology.}$$

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—talk of sentence or propositions that have content, are about things, have meaning, and are true or false—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 \textit{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—the folk psychological notions—from the 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—i.e., 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 neuroscientific 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.” That is, 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 that instead of the folk psychological notions of pretending_{FP}, we have a scientifically respectable notion, pretending_{SR}, that is analogous to pretending_{FP} in appropriate ways. The suggestion is that the mental fictionalist appeals to pretending_{SR} instead of pretending_{FP}. But then in what sense is this a fictionalist account, if we insist that fictions involve pretending_{FP} and not pretending_{SR}?²⁸

I will discuss this issue more thoroughly below in relation to (NoUS) and (NoProp), but for now the following may suffice: there are many options for the fictionalist about how she defines fictions. To assume that she 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 thinks that folk psychological notions should be abandoned, then she is clearly not going to accept 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. But, again, I will have more to say about this below.

Or suppose one is a mental fictionalist who accepts:

²⁸ Thanks to Ted Sider for pressing me on this point.
Private sensations and qualitative feels—as posited by folk psychology—do not exist.

Just because you deny sensations and qualia, you may not deny 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).

One may also endorse either:

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

Then surely our common sense, folk psychological notion of fiction is incorrect, if you insist that fictions are stories that we have certain mental attitudes toward. But if one embraces either of the above claims, this leaves 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. For example, as discussed a few paragraphs back, a mental eliminativist may claim that that fictions involve a scientifically respectable notion, pretendingSR, which is analogous to the folk psychological notion pretendingFP in appropriate ways. The mental fictionalist who is at heart committed to connectionism, for example, might have a connectionist-friendly notion of fiction, which is somehow analogous or a successor to the folk psychological notion of fiction.

However, it may be argued that a connectionist would insist that this 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. Moreover, if a mental fictionalist inclined towards 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?

But as stated earlier in section 2.2, 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

---

29 See Ramsey, Stitch, and Garon (1990), e.g.
invokes possible worlds, etc., 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), for example, 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 ‘Ф(х)’ is a formula, ‘Фm’ is true iff, according to the fiction of folk psychology, there is a mental state m such that Фm.

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.” (Parent 2013: 610)

But the mental fictionalist may rightly insist that the problem is with (SPS). As Parent has pointed out, (SPS) is laden with folk-psychological concepts. To accept (SPS), in other words, is question-begging against the mental fictionalist who denies the very folk psychological concepts imbedded in (SPS). So a mental fictionalist has very good reason to reject (SPS).
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. 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, 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.” (1981:87)

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, etc., 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 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 may 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-FP 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

30 “…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. The theory is uniform for all terrestrial brains, not just human brains, and it makes suitable conceptual contact with both evolutionary biology and non-equilibrium thermodynamics…Regarding long-term activity, the theory provides powerful and unified accounts of the learning process, the nature of mental illness, and variations in character and intelligence across the animal kingdom as well as across individual humans.

“Moreover, it provides a straightforward account of ‘knowledge,’ as traditionally conceived…a system of propositional attitudes, such as FP, must inevitably fail to capture what is going on here, though it may reflect just enough superficial structure to sustain alchemylike tradition among folk who lack any better theory.”

“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.” (1981: 85)
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. The idea is supposed to be that human beings may eventually learn to communicate with each other in the way that our bodies already automatically relay information internally. Churchland imagines:

“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.

Churchland summarizes an argument from cognitive collapse against his view:

“…the reductio proceeds by pointing out that the statement of eliminative materialism is just a meaningless string of marks or noises, unless that string is the expression of a certain belief, and a certain intention to communicate, and a knowledge of the grammar of the language, and so forth. But if the statement of eliminative materialism is true, then there are no such states to express. The statement at issue would then be a meaningless string of marks or noises. It would therefore not be true. Therefore it is not true. Q. E. D.”

He responds, however:

“The difficulty with any nonformal reductio is that the conclusion against the initial assumption is always no better than the material assumptions invoked to reach the incoherent conclusion. In this case, the additional assumptions involve a certain theory of meaning, one that presupposes the integrity of FP. But formally speaking, one can as well infer, from the incoherent result, that this theory of meaning is what must be rejected….one cannot simply assume a particular theory of meaning without begging the question at issue, namely, the integrity of FP.”
But likewise with arguments from cognitive collapse against mental fictionalism, which insist that fictions presuppose FP. To modify Chuchland’s response: one cannot simply assume a particular theory of fictions and fictional content without begging the question at issue—namely, the legitimacy of FP!

If arguments from cognitive collapse against the mental fictionalist assume a FP-friendly concept of fiction, then it is this question-begging notion of a fiction that must be rejected, not mental fictionalism.

This does not mean that I think 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. But even without granting wild possibilities about alternative forms of communication, the broader point still stands: 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.

Even so, if the mental fictionalist cannot accept an FP-laden notion of a fiction (or truth-in-fiction), how may she state the linguistic portion of her view? She maintains that something is going on when we are engaged in folk psychological discourse. What is it? Answers will vary, and will depend on the particular mental fictionalist’s preferred mental eliminativism. But I hope the above discussion has shown that the mental fictionalist has options.

[…]

In sum: 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 very folk psychological concepts are the very things she denies exist. Yet we have seen that the details of what, exactly, in FP she denies matters greatly. Some of the ways of being a mental eliminativist are not incompatible with accepting fictions. Indeed, it is not implausible that a mental fictionalist may have an entirely FP-free notion of fiction. So at best: arguments from cognitive collapse put pressure on the mental fictionalist to provide more detail about what, exactly, a fiction is. But this does not threaten doom for her view.

References


Hume, D. (????) ??????


Philosophy 50: 149-59.

Mackie, ???????


