Mental fictionalism: a foothold amid deflationary collapse[1]

Meg Wallace (forthcoming)
in *Mental Fictionalism: Philosophical Explorations* (Routledge)
Tamas Demeter, Ted Parent and Adam Toon (eds.)

I.   Serious Ontology, Deflationism and Fictionalism

Some ontologists take their work very seriously. They aim to discover the fundamental truths of reality, to figure out how the world is and what is in it, by way of *a priori* reflection and argumentation. Such philosophers are *serious ontologists*.[2] Their seriousness is reflective of the attitude they take towards the metaphysical conclusions they endorse, the dialectical landscape they accept, including the ground rules for philosophical debate and available moves in logical space. They may not agree with their opponents' views, but they nonetheless think that it is well worth the effort to sit down with them and debate about why not.

Let's identify three distinguishing features of the serious ontologist. First, serious ontologists are proposing ontological theories. They present and defend theories about the world with the aim of getting it right. There is an assumption of match, correctness, truth, or accuracy. In tandem with this assumption of accuracy is the second feature: the assumption that the world is in fact a certain way. It's an assumption that there's a *fact of the matter* about the world.[3] Third, there

---

[1] Many thanks to Tamas Demeter, Ted Parent, and Adam Toon for organizing two conferences on Mental Fictionalism: one at the University of Edinburgh in Edinburgh, Scotland (2014) and the other at the Hungarian Academy of Sciences in Budapest, Hungary (2019). Helpful feedback from audience members and participants at both events helped hone the ideas and arguments in this paper. Much gratitude is also owed to all of those involved in this present volume, contributors and editors alike. Thanks also to Tim Sundell and Ted Parent for personal discussion and comments.

[2] Terminology borrowed from Thomasson (2009) and Jenkins (2014). Jenkins defines "serious ontological debates" as those that are "neither trivially easy to solve nor merely verbal." (p. 454) This way of capturing the relevant class of ontologists isn't quite right for my purposes here, as it would include *deflationists* such as Thomasson (2017), who diagnoses certain ontological debates as *verbal but valuable*. Thomasson thinks that certain ontological debates are neither trivially true nor *merely* verbal, yet she also takes a deflationary stance towards such debates in that she thinks that they do not involve metaphysical or philosophical issues. In order to capture this kind of deflationary view (among others), my understanding of 'serious ontologists' will differ.

[3] Some might say this is just ontological realism at the meta-ontological level. That roughly captures the views I have in mind, but not entirely. After all, some may claim that - as ontologists - we (i) propose theories about the world with the aim of getting it right and (ii) the world may be a certain way which explains, in part, how the theory can get it right, but they may also claim that (iii) accomplishing (i) and (ii) is *utterly trivial*. In short, the delfationists I have in mind may very well include meta-ontological realists themselves who think that their brand of realism isn't really all that metaphysically deep or interesting. Chalmers (2009) calls such views 'lightweight realism.' Hirsch (2002), (2011) uses the term 'shallow ontologist.' These are not serious ontologists as I am understanding them here.

is a shared attitude that these ontologists have towards what they are doing, the methodological practices in place, and the aims of doing metaphysics more broadly. They take their projects to be philosophically important. We might even go so far as to say that it is a kind of value judgement about the metaphysical oomph of their claims about the world.[4] Serious ontologists take the conclusions that they endorse and the activity they are engaged as being *metaphysically important*.

In contrast, certain deflationists insist that many ontological debates are merely verbal, trivial, nonsensical, off-topic, or pragmatic.[5] These deflationists maintain that there is something deeply wrong with serious ontological debates - either with the questions being asked or in the answers being given. For my purposes here, the crucial unifying feature of these deflationary critiques is the *rejection* of the claim that debates about these topics are in any way *philosophically substantial*. Carnap (1950), Hirsch (2011), and Thomasson (2017) are three examples of proponents of this kind of deflationism. They all agree - for one reason or another - that the arguments and conclusions endorsed by serious ontologists are decidedly *not* metaphysically important.

Such deflationists typically direct their complaints toward a certain subset of metaphysical inquiry, not *metaphysics-at-large* or *ontology as an entire discipline*. Some direct their complaints toward discussions over ordinary, humdrum objects such as tables, chairs, statues, cats, or persons. Others complain about debates involving ontological oddballs such as mereological sums, numbers, properties, classes, propositions, and so on. Yet as useful as it may be to recognize which ontological debates are in fact targeted by such deflationists, it is also illuminating to see which debates are *not*. Notably, what's missing are ontological debates involving *consciousness* or *mental entities*. This absence is curious. According to their own deflationary criteria, ontological debate about consciousness or mental entities should qualify as a candidate subject to their critiques. Yet they aren't.[6] Why not?

---

[4] These views roughly align with what Chalmers (2009) has called 'heavyweight realism', and what Manley (2009) has called 'mainstream metaphysics,' although some details differ.

[5] Bennett (2009) uses 'dismissivism' instead of 'deflationist' to characterize the generalized view that there is "*something* deeply wrong" with certain ontological debates. Her interest concerns the motivations behind the deflationist position, however, and her primary aims in that paper do not entirely align with mine here. So I will stick with using the term 'deflationism' since it better suits my purposes.

[6] There are exceptions. Sidelle (2007) explicitly discusses disputes in philosophy of mind and suggests that certain debates, such as those between identity theorists and functionalists, may be open to deflationary critiques, and uses this particular debate as a "philosophical case study." Sidelle categorizes disputes in philosophy of mind as one subcategory (of many) under a broader class of disagreements, so his aim is much broader in scope than some of the views I consider here. Sidelle also doesn't insist that these disputes are in fact verbal, but rather that it "can be legitimately asked whether they are…" (84) His project is fairly exploratory and non-committal relative to the other deflationary views I will be examining in this paper; his discussion will not be part of my focus.

Is it that debates in the metaphysics of mind do not involve participants taking ontological issues *seriously*? Clearly not. *Is there a mind? Is it material or immaterial? What is the relation between the mind and the body? Are mental events or properties reducible to physical events or properties? Are our subjective first-person experiences reducible to empiricist-friendly objective facts?* These are basic, familiar questions in philosophy of mind that involve deep questions about existence, about what there is or isn't, and about the metaphysical make up of what certain entities are like. Such questions have been (and still are) taken to be metaphysically meaningful by many who engage in debate about these topics.

So is it plausible that those engaged in debate about the mind or mentality are deeply mistaken in their assumption that the ontological issues in philosophy of mind are really worth arguing about? After all - and as we shall see - *according to the deflationists' own criteria*, many ontological debates in philosophy of mind share certain salient features with other debates that, deflationists argue, are susceptible to their critiques. Surely that is reason enough to merit further investigation into the depth (or shallowness) of disputes about mentality.

In what follows, I examine three meta-ontological deflationary approaches and apply them to ontological debates in philosophy of mind. An intriguing consequence of this, I argue, is that it leads to systematic problems for *deflationism* - specifically, problems of *cognitive collapse*. This should be surprising: cognitive collapse problems are usually reserved for serious ontological views such as *eliminative materialism* and *mental fictionalism*, not deflationism.

Mental fictionalism is the thesis that we are (or should be) fictionalists about mentality. Statements made about the mind, consciousness, or whatever it is that is going on inside our heads are not *literally* true, the mental fictionalism claims; rather, they are *fictionally* true. Such talk, despite being strictly speaking false, is nonetheless useful and should be kept in use.[7]

Mental fictionalism may seem an odd fit for a paper about *serious ontology* and *deflationism*. After all, fictionalists in general may seem to be neither quite deflationist (they take an ontological stance, if speaking literally) nor serious (they avoid an ontological stance, if speaking

---

[7] For my purposes here, I am interested in mental fictionalist accounts that have both an ontological thesis and a linguistic thesis. The ontological thesis is either eliminativist or agnostic with respect to mentalistic entities (or whatever entities are presupposed by folk psychology). The linguistic thesis is a claim about the mentalistic language - the folk psychological discourse - and avows certain non-truth-aiming virtues for it. Perhaps, like Lewis (1978), the linguistic thesis involves appeal to hidden fictional operators. Perhaps, like Walton (1990), utterances in the folk psychological fiction involve some kind of subtle pretense and are only pseudo-statements. Options abound. What matters is that the mental fictionalist view explicitly separates our talk of mentalistic entities from our ontological commitments to them. See Wallace (this volume) and (2016), e.g.

fictionally). It may also be thought that fictionalism is just a kind of deflationary position, or that the distinction between deflationism and fictionalism is negligible. Indeed, even Carnap (1950), while arguing for a meta-ontological deflationary view, claims that one of his primary goals is to show how using a language that refers to certain metaphysically suspicious entities - *abstracta* such as numbers, properties, etc. - can nonetheless be "perfectly compatible with empiricism and strictly scientific thinking." (Carnap 1950, p.21) This certainly seems to be the fictionalist's aim, too: finding a way to separate our ontological discourse from our ontological commitments.

Nonetheless, my focus here involves serious ontologists who propose theories about the world, who assume that the world is in fact a certain way such that a theory about the world could get it right, and assume that engaging in discussions about both of these things is *metaphysically important*. Generally, fictionalists tend to meet these three conditions. A fictionalist about *X* is often motivated by the assumption that discourse about *X*s is ontologically committing. Yet she also admits that ontological commitment to *X*s is somehow problematic, even though discourse about *X* is indispensable. This last point is crucial. The fictionalist does not think that debate about *X* is shallow, silly, or pointless; on the contrary, she *endorses X*-talk. She thinks that talk about *X*s is so important that, despite the fact that many claims about *X*s are strictly speaking false, X-talk should nonetheless continue.

One might be tempted to draw a distinction between *first-order talk about Xs* and *second-order meta-ontological debate about the existence of Xs*. Perhaps, one might argue, a fictionalist endorses the former but is happy to reject the latter. This seems implausible. Fictionalism (as I am understanding it) consists of a linguistic thesis and an ontological thesis. Where does the ontological thesis come from? Presumably it comes from *discussion about the very entities that the fictionalist is fictionalist about*. So, she must think that discussion about the existence of Xs is worth the trouble of doing, since she eventually comes to some conclusion about it (even if it's an epistemically cautious one - i.e., agnosticism about *X*s). Put another way: that the fictionalist about *X*s *has an ontological stance about Xs* indicates that she thinks that discussion about the existence of *X*s is worthwhile - at least, worthwhile enough to come to a particular conclusion about the ontological status of Xs.

It is very often *because* fictionalists assume that certain metaphysical debates are substantial and worth doing that they are motivated to embrace the fictionalist view that they favor. They very often start from an assumption of Quinean meta-ontological methodology, which maintains that any ontologically committing language is worth investigating both in and of itself (to see if what our ontological commitments are) and in consideration of its integration with other theories (as part of its potential fit into our best overall theory of the world). Fictionalism,

as oddly as it may seem, is one way to take ontology seriously. Consequently, deflationary critiques against serious ontological debates are *ipso facto* critiques against fictionalism.[8]

Here is the plan. Section II examines meta-ontological deflationism as applied to mentality. I examine three deflationary views, one by one: Carnap's (1950) *framework* proposal, Hirsch's (2011) *verbal dispute* thesis, and Thomasson's (2017) *metalinguistic negotiation* view.[9] I apply each of them to ontological debates in philosophy of mind, and show how this leads to systematic cognitive collapse. Throughout this discussion, I discuss the advantages of mental fictionalism over deflationism with respect to these worries. In section III, I explain why deflationism about mentality leads to systematic collapse, and expand on why mental fictionalism is in principle better equipped to address collapsing problems than deflationism is.

II. <u>Mental Deflationism</u>

i. The Folk Psychological Framework

Carnap (1950) is interested in explaining why existence questions as asked by the ontologist are misguided or confused. In particular, he is interested in ontological debates concerning *abstracta* - numbers, properties, propositions, and so on. He aims to show how we can accept a language that seemingly quantifies over abstract entities and yet still adhere to our "nominalistic scruples." We can do this, Carnap proposes, by properly understanding the distinction between *internal* and *external* questions with respect to *linguistic frameworks*. If we are asking internal questions relative to a framework, then the answers are straightforward, easy, or trivial, making the questions relatively uninteresting - and not worth asking. If we are asking external questions relative to this framework, then the answers are meaningless, nonsensical, or merely practical, making the questions non-theoretical at best - and (metaphysically) not worth asking.

Carnap proceeds by a strategy of examples. He provides a healthy sample of frameworks - the world of things, the system of numbers, the framework of propositions, the framework of thing properties, and so on - and then outlines what each framework looks like, the seeming

---

[8] To be clear, my primary focus here will be on mental fictionalism. I aim to show how the mental fictionalist has the resources to avoid cognitive collapse, while the mental deflationist does not. Yet - as we will see - some of what I say can be carried over to eliminative materialism, too.

[9] Thomasson's (2017) deflationism is distinct from the 'easy ontology' deflationism she defends in Thomasson (2009), (2015), etc.

commitments involved in accepting them, and various internal and external questions relative to each. This suggests a helpful diagnostic going forward. For any area of ontological dispute about *X*s, we can suppose that there is a relevant framework with new vocabulary and rules - *the X-framework*. We can then ask ourselves: Are existence questions about *X* internal or external to the X-framework? If they are internal, then the answers can be determined either empirically or logically, depending on whether the *X*-framework is factual or theoretical, respectively (Carnap 1950, p.22). If the questions are external - if the questions are about whether *X*s exist in some sense that is distinct from "Is *X* an element of the *X*-framework?" - then either the question is absurd and meaningless, or else it is a mere practical question about whether we should adopt the *X*-framework. Whatever our answers, we either have a straightforward diagnosis about the existence of *X*s, or we have an explanation about why the question is either malformed or merely practical. Importantly, in no case are the questions (or answers) theoretical or deeply philosophical.

So let's consider the *folk psychological framework*. This framework is similar to Carnap's *world of things framework* in that it is concerned with everyday language. It includes the ordinary, everyday talk of whatever is going on inside our heads - our thoughts, beliefs, desires, choices, intentions, etc. Given how commonplace this framework is, it may be considered the mental correlate of the *world of things framework*. Ordinary folk use this language and talk this way all of the time to describe the contents of our own heads. We have all accepted this framework early on in our lives. Many of us (unless we are philosophers) do not even bother asking ourselves whether we should accept this framework or not. We merely find ourselves having already accepted it.

Once we have the folk psychological framework in place, can we use it as a way to capture the ontological questions in philosophy of mind? It initially seems so. The mind/body debate, as far as the disputes between material monists and dualists go, is nicely analogous to disputes between nominalists and Platonists with respect to abstract entities. Canrap's stated purpose is to show that using a particular language that seemingly commits to abstract entities does not in fact commit us to such things in our ontology. So we should be able to use his deflationary approach to show how we can, in principle, use a language that *seemingly* commits to dualistic or non-empiricist *mentalistic* entities without thereby *in fact* committing ourselves to such entities.

According to framework deflationism, here is how the process works: first, we suppose that an individual *wishes to speak* about a new kind of entity. In order to do so, they *introduce* a new framework, "a system of new ways of speaking, subject to new rules." (Carnap 1950, p. 21) Others can then *accept* the framework or not, depending on what she *wants* to do, which may

come down to a *practical decision* about the utility of doing so. There may be some *deliberation*, but in the end we are *free to choose* whether to accept a framework or not. With this framework in place, we can then *raise questions* relative to it, which we can *intend* to be either internally or externally *understood*.

Perhaps now we start to see a problem. The action items involved in the framework procedure, indicated by italics, are taken nearly directly from Carnap himself. This is the process that's described. It's how frameworks are constructed, how we respond to them, how we ask questions in light of them, etc. Yet, importantly, this description involves *folk psychological notions*. In order to have a clear description of what frameworks are and how they work, we have to assume certain folk psychological notions such as what the speaker desires, wishes, intends, chooses, etc. Yet if these folk psychological notions are a *prerequisite* for framework acceptance, how is acceptance of the folk psychological framework coherent?

Let's tread carefully. Carnap explicitly denies that in accepting a framework we are then committed to an "acceptance of a *belief* in the reality" of whatever entities that framework is about (Carnap 1950, p.23, emphasis his). So, for example, if the system of numbers says "there are infinitely many prime numbers", then in accepting this framework we are not thereby committed to *believing in the reality of infinitely many prime number*s. We can accept the framework without thereby acquiring a belief that the entities circumscribed by the framework *exist* in some robust, external sense. But note that this is a claim about the connection (or rather: disconnection) between accepting a framework and having or acquiring a belief about the existence of the entities described by the framework. In contrast, the problem I am interested in is whether, in accepting a framework, we are presupposing mentalistic activity *at all*. Put another way: Is the process of *accepting a framework* itself a mentalistic activity? If it isn't, then what exactly is it? If it is, then the folk psychological framework seems problematic.

In response to the ontologist who maintains that certain existence questions must be answered prior to the adoption of a framework, Carnap explains that activities such as accepting a framework are merely a way of adopting a new language, or a new way of describing things, not the adoption of new beliefs about how things really are. Moreover, any question prior to accepting the framework about whether we *should* accept it is a practical question, not theoretical.[10]

---

[10] "...the introduction of the new ways of speaking does not need any theoretical justification because it does not imply any assertion of reality. We may still speak...of 'the acceptance of the framework' or 'the acceptance of the new entities' since this form of speech is customary; but one must keep in mind that these phrases do not mean for us anything more than acceptance of the new linguistic forms. Above all, they must not be interpreted as referring to an assumption, belief, or assertion of 'the reality of the entities'. There is no such assertion. An alleged statement of the reality of the framework of entities is a pseudo-statement without cognitive content. To be sure,

Ok, fine. But the acceptance of a framework is still an activity. So what *is* this activity? And what is it in relation to the folk psychological framework? Carnap says that acceptance of a framework is a judgement, but without cognitive content. Perhaps it is something like *entertain* or *presuppose* or whatever it is that is not quite belief and is perhaps also non-propositional. Perhaps it is a kind of *knowledge-how* or skill acquisition state or some emotive, non-cognitive mental attitude. We don't need to specify what it is exactly, so long as we are careful about what it is *not*. And, importantly, we can easily grant that any alleged 'statement' about the reality of the framework entities is a pseudo-statement without cognitive content while also acknowledging that *the act in accepting this pseudo-statement is itself a mentalistic activity*.

If accepting a framework is *not* a mentalistic activity, then what else could it be? What exactly are we *doing* when we accept a framework? Is it exercise? Is it bodily movement in space? Are we looking at a work of art? Evaluating it aesthetically but non-propositionally? Is it something that animals can do? Is it something that my smartphone can do? Presumably, no; it's none of these things. It is an activity that we (human beings?) engage in as a response to a linguistic framework before us.[11] One may be tempted to explain: these frameworks are themselves an (abstract?) new language with new rules that someone constructs and introduces for the purpose of having a new way of speaking, to describe a new kind of thing. Yet all of these actions seem to presuppose folk psychological notions - *understanding* and *constructing a language*, *intending* it to *mean* or describe things in a certain way, etc. So, whatever it is that accepting a framework involves, it is at the very least some kind of mentalistic activity, as understood by folk psychology. But then the activity of accepting a framework presupposes folk psychological notions.

This may be fine in most cases, and for most frameworks. There is no incoherence, for example, in presupposing folk psychological notions when we explain how it is that we accept the *world of things framework*, or the *system of numbers framework*, and so on. But it is thoroughly problematic to try to explain the activity of accepting the folk psychological framework when the activity of acceptance itself presupposes those very folk psychological notions.

---

we have to face at this point an important question; but it is a practical, not a theoretical question; it is the question of whether or not to accept the new linguistic forms. The acceptance cannot be judged as being either true or false because it is not an assertion. It can only be judged as being more or less expedient, fruitful, conducive to the aim for which the language is intended. Judgments of this kind supply the motivation for the decision of accepting or rejecting the framework." (Carnap 1950, pp.31-32)

[11] And even this talk of 'before us' is metaphorical; frameworks are not paintings in a gallery. So what are they? And what are we doing when we contemplate, accept, or reject them?

To be clear, the conflict comes from what we are *doing*, not what we believe. So we can agree with Carnap that no one is committed to *believing* in the existence of folk psychological notions when they accept the folk psychological framework. But we are nonetheless committed to *engaging in folk psychological activities* when accepting the folk psychological framework. Furthermore, just because some activity is pragmatic or non-theoretical does not mean that it is non-mentalistic, or free from folk psychological notions. So it won't help to appeal to the pragmatic nature of the process of framework acceptance.

Again: unlike the ontologist who Carnap responds to, I am not insisting that certain existence questions must be answered prior to the adoption of a framework. Rather, I am insisting that certain *activities* must be accounted for prior to the adoption of a framework. In particular, our mentalistic activities, circumscribed by the folk psychological framework, must be accounted for prior to adoption of this very framework. But this, in essence, is to adopt the framework before we've adopted the framework, which is nonsensical.

Perhaps the reason Carnap's deflationary views have not, in fact, been applied to ontological questions in philosophy of mind is that it is incoherent to do so. Not incoherent in the way that, according to Carnap, asking theoretical (non-practical) external questions are, but incoherent as in: *conceptually undermining*. This kind of incoherence may sound familiar - it is similar to "cognitive collapse" or "cognitive suicide" worries launched against eliminative materialists and mental fictionalists.[12] Let's take a moment to review arguments from cognitive collapse in order to clarify how such arguments apply to mental deflationism, too.

ii.   Cognitive Collapse

Mental fictionalists (when speaking strictly) are often eliminative materialists at heart: they deny the existence of - or are agnostic about - the entities described by folk psychology. Many mental fictionalists are sympathetic with the eliminativist's claim that folk psychology is a sloppy or outdated theory that is very likely false. They also very often agree with the eliminativist that the ultimate *true* theory - the accurate, literal description of the reality of the world - is much more scientifically rigorous than folk psychology, and is very likely still in progress. This is their ontological thesis. However, unlike the eliminative materialist, the mental fictionalist does not think that we should thereby abandon folk psychology. She does not think that we should stop talking about beliefs, desires, etc., and whatever it is that ordinary folk say is going on inside our heads. Folk psychology, however flawed and false, is also incredibly useful - *indispensably* so. It

---

[12] See Baker (1987), (1988), Boghossian (1990), Reppert (1992), etc. for discussion of collapsing arguments against eliminative materialism. See Parent (2013) and Wallace (2016) for discussion of collapsing arguments against mental fictionalism.

is a helpful fiction that boasts many other virtues besides truth. As such, it should be kept in use and treated *as a mental fiction*. This is her linguistic thesis.

Against mental fictionalism, the *argument from cognitive collapse* is roughly this: in order to propose her theory, argue for her conclusions, to object to competing theories, to generally engage in philosophical debate at all, the mental fictionalist must presuppose the very thing she is, when speaking strictly, eliminativist about. She has to *propose* her theory, *believe* certain propositions, *disbelieve* others, *argue* for her thesis, etc. - all of which are folk psychological activities - the very things the mental fictionalist insists do not exist. Moreover, the mental fictionalist proposes that we treat folk psychological discourse *as a fiction* - something false but valuable. Yet the notion of a fiction *also* presupposes folk psychological notions. To engage with a fiction, we *entertain* it, we *imagine*, we *presuppose*, we *pretend*. These are all folk psychological concepts. So the mental fictionalist cannot propose either her ontological or her linguistic thesis, for doing so presupposes the very things she claims do not exist. The very thing she describes us *doing* - engaging in a folk psychological fiction - when it seems as if we are genuinely participating in mentalistic activity, presupposes the legitimacy of mentalistic activity. This is the problem of cognitive collapse.

Yet notice that the collapsing problem that I've raised for the framework deflationist is much worse than the collapse worries facing the mental fictionalist. One of the defining features of the deflationary position is that it does not take an ontological stance with respect to the ontological issues at hand. Carnap is very clear that his framework deflationism does not need to take an ontological stance - or commit to any existence claims - prior to the adoption of a framework. This is (in part) *why* it is deflationary. But nevermind the ontological, existence claims. What are we *doing* when we accept the folk psychological framework? What are we doing prior to accepting it? The fact that there is no answer here - or, that the answer only makes sense once we've presupposed a framework - suggests that cognitive collapse worries are particularly thorny for the deflationist.

In contrast, for the mental fictionalist, there *is* an answer as to what is going on when we are engaging in the mental fiction. Indeed, there are *two* answers - one literal, one fictional. The literal, strict answer involves a description about whatever non-psychological activities are occurring when we make noises, flap our mouths, wave our arms, gesticulate at powerpoints, and otherwise exchange audible murmurings with fellow philosophers. This literal answer will not involve any reference to folk psychological entities or activities, and it will likely only appeal to empiricist-friendly observable behavior or scientifically respectable neural activity. But we will nonetheless have an explanation for the event taking place when it seems for all the world that we are engaged in the very folk psychological activities that the mental fictionalist

(speaking strictly) denies exists. Moreover, the mental fictionalist also has a nonliteral, fictional answer to describe this activity. This fictional answer *will* involve folk psychological notions. It will include a description of the participants in a debate making statements, giving arguments, defending claims, proposing beliefs and so on, in just the way you would expect a mental realist to give. Only, according to the mental fictionalist, none of these descriptions or statements are literally true or ontologically committing; it is merely a fiction.

One might think that this is problematic, since such an explanation presumes the coherence of a fiction, which also presupposes mentalistic activity. But the mental fictionalist, when speaking strictly, can have a non-psychological description about fictions and engagement with fiction that involves the empiricist-friendly and scientifically respectable entities, events, or activities she accepts. The mental fictionalist can give an account of fictions and our interactions with fictions that is consistent with her mental eliminativist ontology. In this way, because the mental fictionalist *has* an ontological claim - depending on the details of the view - cognitive collapse is avoidable (Wallace 2016).

Yet this kind of move is in principle unavailable to the framework deflationist. Because the framework deflationist doesn't *take* an ontological stance - or rather, because she doesn't think there is one to be had - she doesn't *have* an explanation as to what, exactly, is going on when we accept the folk psychological framework. What's more, those of us who are launching the objection against the framework deflationist do not seem to be saddling the deflationsist with commitments she herself does not already accept. Indeed, we've merely taken the framework deflationist's description of the framework process at face value, which already comes laden with folk psychological terminology and activities.

As may be evident, I don't think that the cognitive collapse worry is particularly troubling for mental fictionalism.[13] But this is because, so long as the mental fictionalist thinks that, strictly speaking, *something* is going on when it seems as if we are engaging in folk psychological behavior, including engaging with a fiction, she will have plenty of resources at her disposal to avoid cognitive collapse. To put it another way, the mental fictionalist is only an eliminativist about a small subsection of ontology: mentalistic entities. She is still an ontological realist otherwise. She is, in line with the categories delineated at the outset of this paper, a *serious ontologist*. So she can gain a solid foothold of reality *somewhere* to explain what is going on when she is *fictionally-but-not-literally* engaging in folk psychological activities - i.e., making statements, arguing for conclusions, proposing, holding, or defending beliefs, engaging with fictions, etc. In this way, the argument from cognitive collapse fails to gain traction against the mental fictionalist. To be sure, the argument is still important, even if it is not effective in

---

[13] For similar reasons, I don't think it's particularly devastating for the eliminative materialist, either.

undermining the view. For it puts pressure on the mental fictionalist to say something (more) about what is literally going on when we are using folk psychological discourse. It forces her to reaffirm her ontological position and give us more of a compelling story about how that fits up with her linguistic thesis, and the folk psychological discourse she wants to keep in use *as a fiction*.

Yet analogous moves are not available to the framework deflationist. The meta-ontological deflationist *by design* does not take an ontological stance. Carnap's framework deflationism in particular is supposed to be a diagnosis of how language works, not how the world works. His view, I take it, is supposed to show us that if we concentrate on words and language carefully, we will see that ontological worries about the world will resolve themselves fairly straightforwardly, or else we will have a satisfactory explanation for our confusion. But investigating framework deflationism as it applies to ontological issues about mentality reveals that this language-world divide diminishes. When it does, there is no obvious foothold to grab onto here. The deflationist, unlike the mental fictionalist, does not have a non-folk-psychological explanation for what is going on outside of the folk psychological framework. He doesn't think that *something else is really going on* when we are using the language of folk psychology, because he doesn't accept the legitimacy of the idea of "*something else is really going on.*" That's what being a deflationist is all about.

I had said at the outset that very often meta-ontological deflationists do not give deflating arguments across the board, against *metaphysics-at-large* or *ontology as an entire discipline*. Rather, they seem to aim their complaints at certain subsets of ontology, depending on their particular interests. Carnap (1950) is primarily concerned with ontological debates about *abstracta*, for example. Hirsch, as we shall see in a moment, is primarily concerned with ontological debates about ordinary objects, and so on. Perhaps this fact can be of use here. Perhaps the framework deflationist can appeal to whatever it is that she is not a deflationist about to aid her response to objections from cognitive collapse. Perhaps she can somehow non-circularly re-describe the activities and behavior involved in accepting the folk psychological framework so that it is not subject to collapse. Or, perhaps she can find something malformed about the cognitive collapse objection itself. Perhaps it begs the question against deflationism *at-large*, or presupposes ontological realism to get it going, making it dialectically ineffective against certain kinds of deflationism. But all of this remains to be seen. As is, the framework deflationist is under pressure to say more, or somehow re-describe the activities and behavior involved in accepting the folk psychological framework, on pain of cognitive collapse. Until then, collapse remains a particularly formidable worry for the mental deflationist.

I had mentioned in the introduction of this paper that cognitive collapse worries are a *systematic* worry for mental deflationism. To show this, we first need to look at a couple of more deflationary proposals. So let's move on to our second one: Hirsch's verbal disputes deflationism. Then we'll look at Thomasson. My ultimate aim is to show how, in the wake of cognitive collapse worries for mental deflationism, mental fictionalism emerges as an admirable alternative theory of mind.

iii.   Verbal Disputes and Disagreements about Mentality

Hirsch (2003), (2011) promotes quantifier variance and common sense. He criticizes contemporary ontologists for endorsing theories about ordinary objects that no ordinary person would likely understand or judge to be true. In particular, Hirsch aims to undermine *revisionary ontologists*, who are defined by the endorsement of a worldview where "…many common sense judgments about the existence or identity of highly visible physical objects are *a priori* necessarily false." (Hirsch 2011, p. 101) This includes the mereological nihilist, who says there are no tables; the compositional universalist, who says there is an object composed of my cat and my coffee mug; the temporal worm theorist, who says that only one small part of my coffee mug is in front of me right now, and so on.[14]

Hirsch's arguments explicitly appeal to a *charitable presumption*: the assumption that ordinary speakers are not asserting necessary *a priori* falsehoods about *ordinary objects*. Importantly, this is not a commitment to the claim that ordinary speakers always speak truly about anything whatsoever. Speakers long ago uttered false statements such as "the earth is flat" or "whales are fish." Hirsch maintains that the most plausible interpretation of such speakers is not that what they said is true, but rather that they had good reasons to believe their utterances (even though they were false). Ordinary folk, Hirsch maintains, can have good reasons for, yet still be wrong about, contingent, empirical claims about *highly visible physical objects*. This does not include conceptual necessary *a priori* mistakes the common folk might make about, say, mathematics or logic; the folk can certainly be wrong about these sorts of things. But they

---

[14] Quick terminological note: as explained at the outset, my aim in this paper is the serious ontologist, not the revisionary one. It is true that, very often, serious ontologists, after spending a lot of time arguing and debating amongst themselves, end up endorsing worldviews that contradict common sense. Yet what makes someone a serious ontologist is not the outrageousness of their claims. Indeed, some serious ontologists explicitly argue *against* extraordinary ontologies, using our unremarkable, ordinary intuitions to support perfectly mild metaphysical views. (See Korman (2016), e.g.) What I am interested in is not the extraordinariness of some conclusion but rather the metaphysical importance the endorser of this conclusion takes her project to have. Nonetheless, for the sake of simplifying the following exegetical portions, I will mostly stick with Hirsch's terminology of "revisionary ontologists" for the moment. There is enough overlap between the theorists we are interested in for it not to make much difference in the section that follows.

cannot be in massive widespread error about necessary *a priori* statements about ordinary objects.

Hirsch's principle of charity is supposed to be the language analogue of a Moorean move in epistemology. A Moorean appeal to common sense knowledge is not decisive. It merely puts epistemic pressure on us to weigh our convictions, pitting ordinary intuitions against all of the premises in an argument that concludes anything to the contrary. Philosophical epistemic principles that lead to absurd conclusions, such as radical skepticism, need to outweigh our ordinary claims about knowledge, such as "I know I have hands." Otherwise, it is the philosophical principles that will be rejected, not common sense. Similarly, Hirsch's appeal to a principle of charity is not intended to be decisive; it merely puts linguistic pressure on us to weigh our understanding of ordinary concepts against revisionary ones. Hirsch maintains that any analysis of 'object', 'thing', or 'table' etc. that flies in the face of our ordinary notions, fails to be an analysis of the very concepts under consideration. So, very often in ordinary object ontology, the thing to go should be our highly theorized principles and analyses, not the original concepts or terms we are trying to analyze or formulate principles about. If ordinary speakers assert "The entire table is in the room", then we are under some linguistic pressure to assume that what they are saying is not necessary, *a priori* false, barring overriding reasons to the contrary.

Given that Hirsch is interested in debates about highly visible ordinary objects such as tables and trees, but not debates concerning necessary *a priori conceptual* matters involving *abstracta* such as mathematics or logic, where does that leave you and me? You and I are at least moderate-sized dry goods, even if, in order to account for consciousness or our inner mental life, we are something else besides. Using Hirsch's own criteria, do ontological debates involving mentalistic entities count as merely verbal disputes? Can we coherently apply his verbal dispute charge to ontological debates in philosophy of mind? If we can, what is the verdict?

One might think that proposals about the mentalistic entities circumscribed by folk psychology will not qualify as revisionary in Hirsch's sense because mental states, beliefs, desires, etc., are not highly visible physical objects like tables and chairs. But why should visibility matter? Is it because highly visible physical objects are more *epistemically uncontestable* than non-visible or non-physical objects? That may be so when it comes to certain *abstracta* such as numbers or properties or sets or sums. But internal mental states are, quite arguably, more certainly known than any mind-independent physical object. Intuitively, we know with more certainty propositions such as "I am in pain" or "it seems as if a table is in front of me" than we do propositions such as "there is a table in front of me" or "my car is in the parking garage." There are plenty of good reasons why the content of our own mind - the Cartesian ego, the thing that

thinks, reasons, wills, has opinions and so forth - is considered the foundation upon which we build knowledge of the non-mental world. So, if the restriction to highly visible physical objects is based on epistemic considerations, it seems Hirsch should be open to applying his brand of deflationism to ontological disputes about the mind.

Furthermore, if Hirsch is primarily interested in deflationism by way of *linguistic pressure*, then ordinary folk psychological discourse should have a similar status as ordinary object discourse. Ordinary uses of folk psychological terms such as 'belief', 'desire', 'pain,' etc. seem just as widespread and accepted as ordinary uses of object terms such as 'table', 'chair,' 'thing,' etc. A charitable presumption of what we say about the content of our own heads will require an interpretation that does not conclude that we are saying wildly false things all of the time. Thus, one might argue, any analysis of folk psychological terms that flies in the face of ordinary usage must fail to be an analysis of the very concepts under consideration. If ordinary speakers assert "I believe the sky is blue", then we are under some linguistic pressure to assume that what they are saying is not necessary, *a priori* false, barring overriding reasons to the contrary.

Most importantly, however, disputes in the philosophy of mind seem to satisfy one of Hirsch's explicit diagnostic criteria:

> "The simplest paradigm of a verbal dispute...is where, for each disputed sentence D, there are two undisputed sentences, U_1 and U_2, one true and one false, such that one side holds that D is (a priori necessarily) equivalent to U_1 and the other side holds that D is equivalent to U_2.

To see how this paradigm is supposed to map onto (what Hirsch takes to be) a merely verbal dispute, take the purported disagreement between the mereological nihilist and the mereological universalist.[15] Let us also assume that the universalist assumes that mereological sums are ontologically innocent - i.e., that sums are nothing over and above the parts.[16] The nihilist claims that there are no composite tables, but grants that there are mereological simples-arranged-tablewise. The universalist grants that there are composite tables (and composite sums of tables and ducks, and so on), but she also grants that there are mereological simples-arranged-tablewise. Given this, it seems we can characterize the debate as follows, where the subscript "c" is to indicate a dispute about composition:

---

[15] The disputants need not be this extreme in their views. The example that follows will work just as well if the disputants were the mereological nihilist and any kind of composite *realist*; the latter need not be a full-fledged mereological *universalist*.

[16] The most straightforward way to do this is to accept composition as identity, of which there are many different types. See Baxter (1988), Lewis, (1991), Wallace (2011), etc.

D_c = There is a table in region R.
U_{c1} = There are simples-arranged-tablewise, in R.
U_{c2} = There is a sum composed of parts in addition to the simples, in R.

The disputed sentence, D_c, between the nihilist and universalist is "there is a table in front of us." The universalist says it is true; the nihilist says it is false. Meanwhile, U_{c1} is undisputedly true according to both sides: both agree that there are simples arranged in a table-like way. And, presumably, both parties agree that U_{c2} is indisputably false - although, admittedly, for different reasons. The nihilist thinks there never are any sums, *ipso facto*, there aren't any sums in region R. The universalist, on the other hand, thinks that there aren't any sums *in addition* to the parts. That is, there are sums, but they aren't something over and above the parts that are already there.

So we've satisfied the template. According to Hirsch, this should indicate that we are in the midst of a merely verbal dispute.[17]

My concerns here are not whether Hirsch's criterion for verbal disputes are *adequate* but rather, assuming that they are, would they apply to ontological disputes involving mentality? Putting it another way, I want to know whether Hirsch, *given his own criterion*, would put disputes about mentality in the same category as disputes about composition and tables. If he thinks that the latter are merely verbal, then given his stated reasons for thinking so, should he think that disputes about the mental are merely verbal as well?

Imagine that we have a dispute between an identity theorist and an eliminative materialist.[18] The identity theorist is, of course, not a substance dualist, but she nonetheless thinks that folk psychological utterances such as "Joe believes the sky is blue" are true. The identity theorist believes that some (most) of our ordinary folk psychological claims are true, whereas the eliminative materialist thinks that they are false. Using Hirsch's paradigm, we might see if we could characterize the debate between these disputants as follows, where the subscript "m" is to indicate a dispute about mentality:

---

[17] Bennett (2009) raises some objections to Hirsch's criteria here, which I am sympathetic to. Her objections may also be behind any forced awkwardness of the example - i.e., why U_{c2} may not *quite* adequately capture what is uncontroversially accepted as false between the nihilist and universalist. We may also worry whether there can even be a diagnostic test, even in principle, to determine whether a certain debate is verbal. See Jenkins (2014a, 2014b), e.g. These issues, while important, are not relevant for my purposes here.

[18] For comparison, consider Sidelle's (2007) discussion of the identity theorist and the functionalist. Sidelle is considering a slightly different verbal dispute than Hirsch (2011); his criteria rests on the fact that the identity theorist and functionalist agree on all the empirical facts and which worlds (on some description) are possible. (Sidelle 2007, p.95)

D_m = Joe believes that the sky is blue.
U_{m1} = Joe's brain is (participating) in a certain physical state, event, or activity.
U_{m2} = There is a mental event in addition to the physical state, event, or activity occurring in Joe's brain.

The disputed sentence, D_m, between the eliminativist and the identity theorist is "Joe believes the sky is blue." The identity theorist says it is true; the eliminativist says it is false. Meanwhile, U_{m1} is undisputedly *true* according to both sides. Both agree that Joe's brain is in a certain physical state or is participating in some physical event or activity. Certain material facts, we can assume, are not in dispute. Moreover, both parties agree that U_{m2} is *false* - although, admittedly, for different reasons. The eliminativist thinks there are never any mentalistic activities or events, *ipso facto*, there aren't any mental events in addition to whatever is going on inside Joe. The identity theorist thinks that there aren't any mental events *in addition* to the physical events occurring in Joe's brain. There are beliefs, but these kinds of mental states aren't *something over and above* the physical stuff that is already there.

So we've satisfied the template. According to Hirsch, this should indicate that we are in the midst of a merely verbal dispute. Thus, the identity theorist and eliminative materialist must mean different things by "belief." They must be speaking different languages, talking past each other, not really disagreeing, etc. They are mistaken in thinking that they are engaged in a genuine dispute.

However, notice that the issue is particularly tricky now that the topic is mentality. According to Hirsch's deflationist story, one disputant *holds* that D_m is equivalent to U_{m1}, the other *believes* that D is equivalent to U_{m2}, both *accept* that D_m is true, but they think that it *means* different things, and so on. This process supposedly shows that the disagreement is over word usage and definitions, not ontology or metaphysics. It's about words, not the world. But all of this activity - the entire verbal deflationist explanation - presupposes mentalistic activity. According to Hirsch's own account, the description of what's going on between the identity theorist and the eliminative materialist *presupposes folk psychological notions*.

Here's another way to put the point. Suppose the deflationist thinks that his deflationist description of the debate is accurate - i.e., that the verbal deflationist theory of what's going on *matches up* with what is really going on, that it truly describes the events occuring between the identity theorist and the eliminativist when they are debating about ontological proposals in philosophy of mind.[19] Then his view presupposes the existence of the very thing at issue:

---

[19] I doubt he will do this, but more on that in a minute.

mentalistic entities. In order to take the process of a verbal dispute at face value, as it is being described by the verbal deflationist, certain folk psychological assumptions have to be made - the very things under dispute. We can imagine that an eliminative materialist, for example, will flat-out reject this picture as question-begging against her view!

When a dispute is assumed to be ontologically substantive, a common deflationist move is to explain how the dispute is, on the contrary, verbal, not metaphysical. We thought that we were talking about the metaphysical nature of reality - tables, chairs, numbers, propositions, free will, etc. But now, if the deflationist has been successful, we've been shown that the dispute is *really* about language, how we use our terms, how we define our concepts, etc. It is about words, not the world. Yet, as we are beginning to see, ontological debates about the mind are importantly different. The very mentalistic activities under scrutiny are presupposed in our (usual) understanding of what a verbal dispute *is*. So to point out that two disputants discussing ontological issues about mentality is *merely verbal* is not to give a diagnosis of the activity free from the issues under scrutiny. Diagnosing a dispute as merely verbal presupposes the very notions under dispute. *Are there beliefs? Are there meanings? Do we have mental attitudes towards propositions?* If a verbal deflationist says 'yes' to all of these in order to show that the particular debate is merely verbal, then this is to assume that eliminative materialism is *false*. This is not deflationary; this is to take an ontological stance. Moreover, it's to take an ontological stance on an issue that, according to the verbal deflationist, is verbal and not ontological. So it is non-deflationary *and* circular. Thus, the only available move for the deflationist is *not* to take an ontological stance - i.e., not to assume the legitimacy of folk psychological discourse. But then her diagnosis of the debate as merely verbal cannot gain traction. This is cognitive collapse.[20]

Hirsch himself does not, in fact, include ontological debates about the mind in his deflationary criticisms. Why not? He doesn't say. In light of the looming threat of cognitive collapse, we now see the importance of having a principled division between debates that are subject to verbal deflationary critiques and those that are not. Yet it's hard to see where any verbal deflationist will be able to draw the line. The verbal deflationist, like the framework deflationist, is under pressure to say more, or at least re-describe the activities and behavior involved in a verbal dispute such that it is not subject to cognitive collapse. But it is difficult to see how this will be possible for a meta-ontological deflationary view since *by design* they do not take an ontological stance. As such, these worries seem particularly challenging for the deflationist in ways that it isn't for serious ontological views, such as mental fictionalism.

I still have not explicitly said why cognitive collapse worries are *systematic* for meta-ontological deflationary views applied to philosophy of mind. Let us discuss one more deflationist position,

---

[20] Thanks to Ted Parent for comments here.

Thomasson's (2017) metalinguistic negotiation thesis, and then we'll back up to look at the bigger picture.

iv.   Metalinguistic Negotiations and Disputes about Mentality

While not mentioned explicitly above, Hirsch launches two complaints against revisionary ontological disputes: (i) that the two ontologists are not really disagreeing, and hence are talking past each other, and (ii) that neither disputant is speaking the language that ordinary speakers are using (English) and hence, that the ontologist and the ordinary speaker are not really disagreeing, and hence are talking past each other. In each case, Hirsch is inferring from (a) to (b):

> (a) Two speakers, X and Y, mean different things by the words they are using.
> (b) Thus, X and Y are not expressing genuine disagreement; it is merely a verbal (trivial) debate.

However, the legitimacy of inferences from (a) to (b) has recently come under criticism. Plunkett and Sundell (2013) challenge Hare's (1991) argument from disagreement to semantic univocality with respect to normative and evaluative terms, which runs roughly as follows:

> "If X says that helping others is good, and Y says that helping others is not good, then X and Y genuinely disagree. But if X and Y genuinely disagree, then they must mean the same thing by 'good'."

But such an argument is an example of the following inference, which is the contrapositive of (a) to (b):

> (c) An exchange between two speakers, X and Y, expresses genuine disagreement.
> (d) Thus, X and Y mean the same things by the words they are using.

Against these inferences, Plunkett and Sundell claim that very often substantive disagreement is metalinguistic, involving what they call *metalinguistic negotiation*. If they are right, two speakers can be engaged in substantive debate, even though they mean different things by their words, because (say) they are implicitly arguing about which is the better language. If implicit disagreement about which language is better qualifies as substantive disagreement, then one cannot move from (c) to (d), nor from (a) to (b). Even if we admit that two serious ontologists mean different things by the words that they are using, such disputants can still be in a genuine disagreement.[21]

---

[21] This point not only cuts against verbal deflationists such as Hirsch, but any others who argue from (a) to (b) or (c) to (d). Siddelle (2007), for example, says "One reason I think verbal disputes are so common is that when I look, I seem to keep finding them! It so often seems that there is a charitable interpretation according to which the claim

Following Plunkett and Sundell, let us characterize a *canonical dispute* as one where speakers disagree over the literal content of what is said, and thus disagree about the truth value of the propositions *literally* expressed. Plunkett and Sundell do not deny that there are canonical disputes; they deny that a dispute must be canonical in order for there to be genuine disagreement. They maintain that there are many instances of non-canonical disputes, metalinguistic negotiation being one among them.

Consider the following dialogue:

(1)
        (a) There is one proton in the nucleus of a helium atom.
        (b) No, there are two protons in the nucleus of a helium atom.

Speakers (1a) and (1b) cannot be in a canonical dispute, since having two protons in a nucleus is consistent with having one. The *explicit* content of (1a)'s statement, in other words, involves a claim that there is at least one proton, which is consistent with there being two protons. But given how we commonly use count statements, (1a) pragmatically conveys that there is *exactly* one proton in the nucleus of a helium atom, whereas (1b) pragmatically implies that there are *exactly* two. So the explicitly conveyed content is compatible, but the pragmatically implied content is not. Granting that the communicative upshot of most exchanges is a combination of *explicit semantic content* and *pragmatically implied content*, it is plausible that speakers non-canonically disagree. That is, the inconsistent content between the speakers is whatever is expressed implicitly and pragmatically, not explicitly and semantically. But if so, then this suggests that quite often (or often enough) disputes are *non-canonical*.

For comparison, consider the following exchange:

(2)
        (a): Waterboarding is torture.
        (b): Waterboarding is not torture.

We can imagine that (2a) and (2b) agree on certain relevant facts: what waterboarding is, what happens to an individual subjected to it, etc. Suppose, however, that (2a) defines 'torture' as "any act inflicting sever suffering, physical or mental, in order to obtain information or punish," whereas (2b) defines 'torture' as "any such act inflicting pain rising to the level of death, organ

---

of each party doesn't conflict with that of the other." (2017: 86). Sidelle is assuming here that because the semantic content of two disputants is *compatible*, they must be having a merely verbal, non-genuine dispute. If Plunkett and Sundell are right, however, this kind of common inference is undermined: absence of conflict over the literal content of what is said does not exclude genuine disagreement between speakers.

failure, or the permanent impairment of a significant body function." Then both (2a) and (2b) speak truly. But if (2a) and (2b) accept different definitions of 'torture,' there is a sense in which their debate is verbal: they disagree how a certain word should be used or defined. But that does not mean that their debate is *merely* verbal, in the pejorative sense in which the charge is often intended. And it certainly does not mean that their debate is not worth having. The speakers in (2) do not explicitly contradict one another, yet they pragmatically imply inconsistent content. In the context of policy-making, for example, (2b) implicitly endorses waterboarding as a tactic for acquiring information or to punish, whereas (2a) opposes it. These are inconsistent positions that will likely have very tangible outcomes in law-making, moral accountability, etc. Thus, we have an example of a debate where (i) the speakers do not mean the same thing by their words, yet (ii) they are having a substantive disagreement (Plunkett and Sundell 2013).

This suggests a meta-ontological deflationary alternative to framework and verbal dispute deflationism. Building off of this suggestion, Thomasson (2017) proposes that certain debates do, in fact, hinge on important linguistic matters (see also Belleri 2017). Understanding serious ontological debates in this way reveals them to be genuine *pragmatic* disagreements, not substantial metaphysical disputes. This deflationist diagnosis of what the serious ontologists are doing does not render their activity pointless. But it is, nonetheless, a deflationist option. It's a way of understanding how it is that the serious ontologist is doing something other than they take themselves to be doing, even if this other thing is important and worth doing. This is *metalinguistic negotiation deflationism*.

To see how this will apply to serious ontological debates, consider the following exchange:

(3)
        (a): The entire table is in the room.
        (b): No, the entire table is not in the room.

Let us assume that the two speakers in (3) do not mean the same thing by 'table': (3a) defines 'table' as whatever it is that ordinary folk countenance as tables - things that look like tables, function as tables, etc.; (3b) defines 'table' as 'a particular trans-world mereological sum of modal and spatio-temporal parts.' Given this difference in meaning, (3a) and (3b) explicitly convey *compatible* propositions. Nonetheless, they may reasonably implicitly convey endorsement of their incompatible definitions. Their disagreement may be over which definition or language is better, where the measure of 'better' depends on context. If the speakers are in the ontology room, for example, and are trying to figure out how best to carve up nature at its joints, avoid constitution puzzles, and so on, then (3b)'s definition may be better. If we are figuring out whether we have successfully moved someone's furniture from

one room to another, (3a)'s definition may be better.

So again: we have an instance where (i) the speakers do not mean the same thing by their words, yet (ii) there is nonetheless a substantive disagreement. Depending on the context, the exchange in (3) may very well be important and worth having, even if the debate is also about words and meanings - i.e., a *verbal dispute*. It's just that it's not *merely* verbal.

Could this kind of deflationism - metalinguistic negotiation - apply to otological debates in philosophy of mind? Consider the following:

(4)
      (a): Joe believes the sky is blue.
      (b): No, Joe does not believe the sky is blue.

As before, we can imagine that speakers (4a) and (4b) agree on certain relevant facts: that Joe is behaving in a certain way, that his brain is in a certain state, that he utters sentences like "I believe the sky is blue", etc. But let us also assume that the speakers have different definitions of 'belief.' (4a) is an identity theorist who takes sentences of the form "x believes p" to be true when x is in a certain brain state (or when the physical events and states of x's brain are thus-and-so). (4b) is an eliminativist who takes "x believes p" to be true only when there is some non-scientific or dualistic entity involved. Given this difference, (4a) and (4b) explicitly convey compatible propositions. But if they implicitly convey endorsement of their incompatible definitions, then their disagreement is over which definition or language is *better*, for the relevant purposes. If the speakers in (4) are in the ontology room, for example, and they are trying to figure out how best to solve substance-dualist problems of interaction, then (4b)'s definition may be better. If we are figuring out whether Joe perceives the world as many others do, then (4a)'s definition may be better. Viewing the debate in this way shows how the speakers in (4) may mean different things by their words, yet nonetheless disagree about which definition is *better* for the relevant purposes.

Notice that, unlike verbal deflationism, there is no dismissal of ontological debates as trivial or pointless according to this deflationist picture. In fact, Thomasson insists that one of the virtues for the metalinguistic thesis is that it captures the fact ontologists take themselves to be doing something important. In contrast, for all of Hirsch's appeal to a charitable presumption of speakers, the one group of speakers he fails to charitably interpret are revisionary ontologists (Thomasson 2017; see also Hawthorne 2006). According to Hirsch, ontologists everywhere are massively mistaken about what they are doing and how important it is: ontologists are *not*

disagreeing, and that whatever it is they are doing is totally trivial.[22] Metalinguistic negotiation deflationism is a vast improvement on this front, for it does not interpret ontologists as doing something utterly pointless. Nonetheless, it is uncharitable in another way: it claims that ontologists are engaged in verbal disputes, even if it is granted that these verbal disputes are important. This is still to grossly mischaracterize what the serious ontologists take *themselves* to be doing.

Notice, too, that unlike framework deflationism where one interprets ontologists as asking *external theoretical* questions, there is no charge here that certain ontological debates are absurd or nonsensical. On the contrary, this proposal provides an explanation of how these debates are quite coherent *at the implied, pragmatic level*. This could still be interpreted as an expansion of Carnap's framework diagnosis of *external pragmatic* questions, however; just not external theoretical ones. Whether to adopt a certain framework, we might say, is akin to endorsing a particular language or vocabulary. Insisting that *this* is the way to define 'table' or 'beliefs', for example, is much like giving practical reasons for endorsing one particular linguistic framework over another. If this is right, then we now have a much more detailed account of what might be involved in accepting a framework.

Even so, cognitive collapse awaits. Take a look again at the dialogue in (4) and the explanation of what's going on, according to the metalinguistic negotiation thesis. The description of two disputants at the metalinguistic level crucially involves folk psychological notions: the speakers *mean* different things by the relevant terms, they *endorse* one definition over the other, they *believe* that one definition is better, where what counts as 'better' depends on contextual features such as what they *desire*, what's *intended*, what they *aim* to accomplish, etc. If we assume that this description is accurate, then it assumes the legitimacy of the very entities at issue. Again, we can imagine that an eliminative materialist would take the metalinguistic explanation of what's going on in (4) as directly question-begging against her view. So in order for a metalinguistic deflationary explanation to get off the ground, it would have to assume that there are, in fact, the very entities as circumscribed by folk psychology - i.e., that eliminative materialism is false. But this is to take an ontological stance and beg the question against the eliminativist, making it non-deflationary and circular. Thus, the only available move for the deflationist is *not* to take an ontological stance - i.e., not to assume the legitimacy of folk psychological discourse. But then her diagnosis of the debate as metalinguistic negotiation cannot gain traction. This, yet again, is cognitive collapse.

---

[22] "The issues being debated by revisionists are not deep; they are completely trivial…Revisionists suffer from the illusion that certain questions are philosophically deep, inviting complicated theoretical debates, when in fact these questions are comically trivial….The argument from charity is a last-ditch effort to bring the revisionists back to their senses...so that they can recognize utter triviality when it stares them in the face." (Hirsch 2011: 102)

We might have thought that moving to the pragmatic, meta-linguistic level would be an improvement on other deflationary approaches. But moving to the metalinguistic level still carries with it the diagnosis that the ontologists' disagreement is over words, not the world. And this very diagnosis still presupposes the legitimacy of folk psychology. On pain of cognitive collapse, the metalinguistic deflationist, just like the verbal and framework deflationist, is under pressure to say more or at least re-describe the activities and behavior involved in giving a metalinguistic analysis. But it is difficult to see how this will be possible for a deflationary view since by design they do not take an ontological stance. This cognitive collapse makes for a particularly formidable worry for the deflationist.

III.   Deflationism, Collapse, and Mental Fictionalism

Let's take stock. We've run through three meta-ontological views and applied them to ontological debates in philosophy of mind. In each case, the meta-ontological deflationist suggests a way to re-diagnose serious ontological debates about mentality. Yet, as we have seen, this rediagnosis assumes the legitimacy of the very entities and activities under consideration. This is why the cognitive collapse worries for meta-ontological deflationists are *systematic*. When the deflationary views are applied to debates about mentality, the deflationary view includes descriptions of disputants who hold beliefs, have desires, make assertions, etc. But this is to assume the legitimacy of folk psychology - the very thing at issue. We cannot assume the legitimacy of folk psychology in order to deflate debates about the legitimacy of folk psychology. If we did, it presumes an ontological stance and is circular. The deflationist *by design* does not take an ontological stance; they are, presumably, providing an explanation as to why there is no ontological stance to be had. But if a stance isn't taken, the deflationary critique suffers cognitive collapse.

Cognitive collapse worries are usually reserved for serious ontological views such as mental fictionalism and eliminative materialism. The fact such worries arise for deflationism about mentality is, itself, quite surprising and worth discussing. Yet even more surprising is the fact that such cognitive collapse worries seem particularly problematic for deflationists. Because the deflationist doesn't take an ontological stance (she doesn't think there is one to be had), she doesn't have the resources that, say, the mental fictionalist does. The mental fictionalist thinks that, strictly speaking, *something* is going on when it seems as if we are engaging in folk psychological behavior, including engaging with a fiction, so she will have plenty of resources at her disposal to avoid cognitive collapse. The deflationsist does not.

Now, true, as I've pointed out several times in this paper, not all deflationary views are deflationary *across the board*. Not all deflationists are *global* deflationsists. Hirsch and Thomasson, for example, do not *in fact* apply their deflationary criticisms to ontological debates in philosophy of mind. They also both explicitly claim they do not intend for their deflationary views to apply to *all of metaphysics*; they are *local* deflationists. So perhaps the cognitive collapse worries I've raised are only a problem for global deflationists, not local ones. Or perhaps collapse is only a problem for *near*-global deflationists - i.e., those who are deflationaist *enough* such that any place where they 'touch down' and make substantive metaphysical claims isn't substantive enough to get them out of cognitive collapse.[23] In this way, perhaps local deflationists will have some resources at her disposal to avoid collapse, just as the mental fictionalist does.

Yet, as we've seen with Hirsch, allowing for some substantive metaphysical debates (and not others) won't help. Hirsch's deflationary criteria on offer - despite being non-global - lands him in cognitive collapse *anyway*. So endorsing a local deflationary position is insufficient for avoiding collapse. Moreover, in Thomasson's case, collapse doesn't occur because of any particular criteria provided, but because the entire deflationary proposal - the overall explanation for why certain debates are deflationary - involves appeal to metalinguistic negotiation. And metalinguistic negotiation (taken at face value) assumes the legitimacy of folk psychological discourse, generating cognitive collapse. So it won't matter if she allows that some metaphysical debates are substantial and others are not; it's the fact that the deflationary proposal on offer is *in principle* inapplicable to mentality that is so problematic.

Moreover, it would surely be an unwelcome consequence that the coherence of an ontological deflationary proposal depends on details of the *non-deflationary* portions of the view. If a deflationary position is only as good (coherent) as it's non-deflationary details, this should be pretty alarming for deflationism generally. Whatever the deflationary position, however global or local and whatever the details, the problem of collapse creates pressure for the deflationist to say what he is *not* a deflationist about, on pain of incoherence. And whatever answer or explanation is given, this should be worrying: realism is a pretty awkward place for a deflationist to hang his hat.

So what? Can't we just take the entire foregoing discussion as a *reductio* on the assumption that meta-ontological deflationism can coherently be applied to ontological debates about mentality? After all, and as I've admitted along the way, none of the deflationary views we

---

[23] I don't know what that would be, but I can imagine cases: if a deflationist thought that the only substantive metaphysical question is whether God exists or not, then this likely won't be enough to get out of cognitive collapse worries.

examined are in fact applied to issues of mentality. If doing so yields cognitive collapse worries, then the solution seems simple: don't do that. Don't apply meta-ontological deflationary views to ontological debates in philosophy of mind. Got it. Let's move on.

Not so fast. We cannot simply refrain from holding a deflationary view of an entire subcategory of metaphysics because of the problems that would arise for the deflationists if we did. This would be theoretically irresponsible (and dialectically suspect). The deflationist needs to provide a *principled* division between those ontological debates that are subject to certain deflationary critiques and those that are not. By *principled*, I mean a way of providing criteria that avoids collapse but isn't merely *ad hocly* designed to avoid collapse. Yet it's hard to see where any deflationist will be able to draw the line. At the very least, neither Hirsch nor Thomasson presently provide us with such a criteria. And neither can merely rely on the fact that their deflationary proposal was not intended to apply to debates about mentality. Given the pressure of cognitive collapse, we are now in need of a full, non-arbitrary explanation of not just why such views *are in fact not* applied to debates about mentality, but why they in principle *can not* be. Why are ontological debates about mentality so resistant to such deflationary critiques? How can we be sure that these collapse problems won't infect other ontological debates? At the bare minimum, our discussion here has put quite a bit of pressure on the deflationists to say more - and, in particular, to say more about the non-deflationary details of their view.

Moreover, whatever the deflationary responses are from here, establishing that certain deflationary arguments cannot apply to ontological debates in philosophy of mind is a huge win for non-deflationary views about the mind - including (and especially!) mental fictionalism. As I said at the start, I am interested in serious ontologists who assume that what they are doing is *metaphysically important*. It is very often *because* fictionalists assume that certain metaphysical debates are substantial and worth doing that they are motivated to embrace the fictionalist view that they favor. Fictionalism, however oddly, is one way to take ontology seriously. Yet meta-ontological deflationary views offer substantial challenges to serious ontology by aiming to undermine the assumption that the issue at hand is a serious ontological one. A deflationary analysis of a particular debate - say, ontological debates in philosophy of mind - is an attempt to undermine *all* of the relevant serious ontological positions, including mental fictionalism. So, if any of the three deflationist views we've examined *can* be successfully applied to ontological debates in philosophy of mind, then this would be a direct challenge to all relevant serious ontological positions, including mental fictionalism. This is how a deflationary challenge for debates about mentality is thereby a challenge for mental fictionalism. So if we've shown - which I hope we have - that certain deflationary views *can not*, in fact, be coherently applied to ontological debates about the mental, then this removes a whole class of (what would have

otherwise been) substantial challenges for mental fictionalism. We will have shown how mental fictionalism, along with other serious ontological views about mentality, has dodged a considerable bullet.

Meanwhile, we have also shown how mental fictionalism has the resources to avoid cognitive collapse worries, whereas certain mental deflationist views do not. While this may not show that mental fictionalism beats out other serious ontological views about mentality, it certainly shows how mental fictionalism has a solid foothold amid deflationary collapse.

REFERENCES

Austin, John Langshaw and Warnock, Jefferey James (1962) *Sense and sensibilia*. Clarendon Press.

Balcerak Jackson, Brendan (2014) "Verbal disputes and substantiveness" Erkenntnis 79(1): 31-54.

Baxter, Donald (1988) "Many-one identity" in *Philosophical Papers* 17(3): 193-216.

Belleri, Delia (2017) "Verbalism and metalinguistic negotiation in ontological disputes" *Philosophical Studies* 174: 2211-2226.

Bennet, Karen (2004) "Spatio-temporal coincidence and the grounding problem" ...

Bennett, Karen (2009) "Composition, colocation and metaontology." *Metametaphysics*. Ed. David J. Chalmers, David Manley, Ryan Wasserman. Oxford: Oxford UP.

Carnap, Rudolf (1950/1956) "Empiricism, semantics, and ontology" *Revue internationale de philosophie* 4(1950): 20-40. Reprinted in the supplement to *Meaning and necessity: a study in semantics and modal logic*. University of Chicago Press.

Chalmers, David (2009) "Ontological anti-realism" in *Metametaphysics* eds. Chalmers, Manely, and Wasserman (2009). Oxford University Press.

Chalmers, David (2011) "Verbal disputes" *Philosophical Review* 120(4): 515-566.

Demeter, Tamas (2013) "Mental fictionalism: the very idea" *The Monist* 96(4): 483-504.

Hawthorne, John (2006) "Plenitude, convention, and ontology" in *Metaphysical Essays*. Oxford: Oxford University Press.

Hirsch (2011) *Quantifier variance and realism: essays in metaontology*. Oxford University Press.

Jenkins, C. S. I. (2014) "Serious verbal disputes: ontology, metaontology, and analyticity" *The Journal of Philosophy* 11:(9/10): 454-469

Lewis, David (1978) "Truth in fiction" *American Philosophical Quarterly* 15(1): 37-46

Marton, Miklos and Tozer, Janos (2013) "Mental fictionalism as an undermotivated theory" *The Monist* 96(4): 622-638.

Parent, Ted. (2013) In the mental fiction, mental fictionalism is fictitious. The Monist 96 (4):605-621.

Plunkett, David, and Sundell, Timothy (2013) "Disagreement and the semantics of normative and evaluative terms," *Philosopher's Imprint* 13: 1–37.

Sidell, Alan (2007) "The method of verbal dispute" *Philosophical Topics* 35(1-2): 83-113.

Thomasson, Aimie L. (2013) "Fictionalism versus deflationism" *Mind* 122(488): 1023-1051).

Thomasson, Aime L. (2015) *Ontology made easy*. New York: Oxford UP.

Thomasson, Amie L. (2017) "Metaphysical disputes and metalinguistic negotiation" *Analytic Philosophy* 58(1): 1-28

Toon, Adam (2016) "Fictionalism and the folk" *The Monist* 99: 280-295

Wallace (2011) "Composition as identity: part 1" *Philosophy Compass* 6(11): 804-816. https://doi.org/10.1111/j.1747-9991.2011.00431.x

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

Walton, Kendall (1990) *Mimesis as make-believe* Harvard University Press.

Yablo, Stephen (1998) "Does ontology rest on a mistake?" *Proceedings of the Aristotelian Society*, Suppl. 72: 229-261