

*Draft of December 2020. Final version to appear in T. Demeter, T. Parent, and A. Toon (eds.), Mental Fictionalism: Philosophical Explorations. New York: Routledge.*

**I Think; Therefore, I am a Fiction**  
T. Parent (Nazarbayev University)

**1. Introduction**

We are all familiar with Descartes' *Cogito* argument for the existence of the thinking self.

It runs something like this:

(Cog) I think.

(\*) So, I exist as a thinking thing.

Even Descartes was aware that (Cog) seemed question-begging re: the self—regardless, the argument strikes many as compelling. (Descartes thought there was some sort of direct intuition of the premise.) However, if the argument is circular, then it seems just as legitimate to run the *Cogito* in reverse, as follows:

(\*) I exist as a thinking thing.

(Cog) So, I think.

Is the reverse *Cogito* at all persuasive? Does the Cartesian “natural light” make the premise attractive? A *mental fictionalist* ought to be concerned. For if (\*) is primitively compelling, then realism about thought follows—and mental fictionalism is false. That is so, regardless of whether ‘mental fictionalism’ is defined as a type of “atheistic” eliminativism, or just as a more modest, “agnostic” view about the mental.<sup>1</sup>

---

<sup>1</sup> For more on such varieties of mental fictionalism, see Parent (2017, ch. 9) and section 2 of the introduction to this volume.

In what follows, I shall explain why I do not find (\*) compelling, and why you should not either. I will try to persuade you that the self is plausibly regarded as a fiction as well, although fictionalism about the self is here merely the “agnostic” and not the “atheistic” type. (Even so, I sometimes give the impression I am an atheist about the self. But while my doubts are strong, I do not believe that they are conclusive.)

I have more than one reason for opposing realism about the thinking self. Besides the threat it poses to mental fictionalism, it is also a threat to the kind of global fictionalism in ontology that I favor. Such global fictionalism, by the way, offers obvious support for mental fictionalism—so a problem here would indirectly bear on mental fictionalism too. At any rate, allow me first a few words about global fictionalism, since this will at least clarify the present motivations, and make clearer how ontological matters are understood. (But if preferred, the reader may skip the next section without losing the main thread of the paper.)

## **2. Background on Ontological Commitment**

When it comes to ontology, I am commitment-phobic—I wish to avoid ontological commitments, strictly so called. That is because it is often right and natural for one’s existential beliefs to be questioned during an inquiry, without advanced warning. So I would not want to thwart such questioning by being categorically wedded or committed to claims about what exists.

But like anyone else, I have ontological *assumptions* in a context of inquiry, even if those assumptions are highly defeasible, and vary from context to context. At least, this is what I want to say in the present context. Yet then, there is a question about how to interpret my existential utterances. Suppose I say ‘there are electrons’ where this is meant to express realism about electrons. But suppose I say this without committing to the existence of electrons across all

contexts. Then, is my existential quantifier commissive or not? Does my statement *entail* that there are electrons?

If it does not, then it seems like my quantifier is not commissive. However, the quantifier must be commissive if I manage to express something ontological at all. If it is non-commissive, my statement would seem like the quantifier in ‘there is an even prime’ when uttered by a mathematical anti-realist. However, when I say ‘there are electrons’ I *am* asserting realism about electrons. And yet, I want to say that I am not *committed* to electrons—at least in the sense that, in some contexts, I would be agnostic about electrons. For instance, I see mereological nihilism as a live hypothesis, where electrons would be mere heaps of more basic entities, superstrings or monads or what have you. I would not take such nihilism as seriously as I sometimes do, if I were simply *committed* to the reality of electrons.<sup>2</sup>

So when I say ‘there are electrons’, does my position entail that there really are electrons? One could simply say that “my position” changes from context to context, but such vacillation in worldviews does not seem very respectable. I think it is at least *possible* for me to have a single (consistent) position, while freely engaging in different inquiries with different ontological assumptions. Sure, the different inquiries will have assumptions that are inconsistent with each other. But *I* am not being inconsistent, if I willingly adopt the different assumptions, relative to the different contexts of inquiry. I am not being inconsistent, just open minded!

But the basic datum remains: If I say ‘there are electrons’ in one context, and then refuse to assert this in another context, it seems I am being inconsistent in realism about electrons. The way to resolve this, I think, is to see my ontological statements as tacitly prefixed with a

---

<sup>2</sup> Quine, in his classic (1948), avoids such issues by speaking of the ontological commitments of a *theory* rather than the ontological commitments of a person. But if preferred, the present issues can be re-framed in Quinean terms by observing that I often adopt different theories (with different ontological commitments) in various contexts.

fictionalizing operator. Thus, when I say ‘there are electrons’, I am really saying “According to the fiction adopted in the present context, there are electrons.” Now in one sense, such a statement is genuinely ontological. Ontologically significant statements *can* occur embedded in fiction, as when a story begins ‘once upon a time, *there was* a dragon...’ The story includes a sentence with a bona fide existential content about a dragon. But at the same time, it is just a story. In precise terms, the statement has an implicit fictionalizing prefix, whereby it expresses something like “According to the present fiction, at some time there was a dragon...” The prefixed statement does not entail realism about a dragon, yet the prefix hooks onto a *bona fide* existential statement—a statement which, apart from the prefix, has a heavyweight quantifier.<sup>3</sup>

This fictionalist proposal is how I hope to construe (nearly) all my ontological assertions. Naturally, there still remains a difference between telling a fairy tale, and giving a theory about electrons in, e.g., the explanation of magnetism. In the fairy tale, it is obvious that we are just spinning a yarn. But with electrons, I am weaving a narrative as well. Yet unlike with dragons, the existence of electrons is something I am *open* to. I am just not committed to it. Put differently, when I say ‘there are electrons,’ it is understood to have a tacit fictionalizing operator—and as such, the statement is not *anti*-commissive; it is only *non*-commissive.

More broadly, I am proposing a kind of *global fictionalism* regarding existential statements. This would be the view that for any sentence of the form “x exists,” its truth condition is a condition of some model or fictional reality, which may or may not “correspond” to reality outside of fiction.<sup>4</sup> Thus, my quantifier is commissive in a *relative* sense—relative to the model that is salient in the context—but non-commissive if my statements are interpreted in an absolute sense. The view is not a hermeneutic fictionalism, but rather a revolutionary one in

---

<sup>3</sup> My understanding of fictionalist semantics is highly influenced by Rosen (1990) and Nolan (2002).

<sup>4</sup> On the metaphysics of nonactual, imaginary realities, see Parent (ms.).

Burgess & Rosen's (2002) sense. It is not the claim that all existential statements in fact have a model-relative commissive quantifier—but rather that inquiry makes the most sense when we interpret them that way.<sup>5</sup>

### 3. Lycan's Realism

There are lots of reasonable objections you might raise at this point. Unfortunately, I cannot address them all here (though I hope to do so in my next book).<sup>6</sup> But there is an objection to such a view that was first raised to me years ago by Bill Lycan. Lycan pressed that every person is saddled with at least one unwavering realist commitment, namely, a commitment to the person herself. The argument he gave was something along the following lines:

- (L1) It is undeniable that I am a thing in the world.
- (L2) If (L1), then I am committed to realism about at least one thing.
- (L3) So, I am committed to realism about at least one thing.

The conclusion obviously follows, and I would accept (L2) as true. Yet it is undeniable that I am a thing in the world? It should be clarified that we are not talking about the well-worn *diachronic* issue about personal identity, about the continued existence of numerically one person over time. Rather, we are talking about the *synchronic* question: Even if there is no such thing as a

---

<sup>5</sup> There are interesting comparisons between the present view and Carnap's (1950) framework-relative metaontology. Carnap could be read as a global fictionalist as, where any existential statement is context-relative due to a fictionalizing prefix ("According to the framework we are presently using, there exists...") However, Carnap would seem to be more of a hermeneutic rather than a revolutionary fictionalist, given that existential statements are *described* as meaningless when used outside their proper framework. For my part, English is flexible enough that existential statements can have a meaning without any fictionalizing prefix (although I deny that it is usually the best way to understand such statements). I hope to elaborate further on these matters in future work.

<sup>6</sup> One footnotable objection: If my ontological assumptions are context-relative, am I committed to the existence of *contexts* across all contexts? What I should say is that, relative to a "context context," I assume the existence of contexts of inquiries. In which case, I will say that in the different contexts, I choose different ontological assumptions. But what if we are not in a "context context?" *Then, I won't have anything general to say about my ontological assumptions in the various contexts.* Regardless, the present context is a "context context."

continuous self (which persists over time), is it *presently* true that a “self” or a person exists? And if so, what conditions undergird its present existence? (Such a self might also persist through time, but we are leaving that aside.)

Having clarified that, why believe (L1)? Knowing Lycan, he might claim a Moorean commonsense intuition in its support.<sup>7</sup> However, it is the pleasure of philosophers to question commonsense, or so they say. And others will naturally look to the *cogtio* to support (L1). But to avoid question-begging, the argument might be reformulated along the following lines:

(Cog') Thinking exists.

(L4) If (Cog') is true, then something produces the thoughts (= the thinking thing).

(\*) So, I exist as a thinking thing.

We can then get to (L1) if we add:

(L5) The previous reasoning has undeniably true premises and is undeniably valid.

But what is the status of the premises? (L4) does not enjoy the indestructible certainty that Descartes sought. Yet (L4) may remain “undeniable” by a reasonable standard. After all, in the normal course of events, things are causally produced by some antecedent condition or state-of-affairs.

Of course, in a volume on *mental fictionalism*, (Cog') hardly seems “undeniable.” That is so, assuming we recognize that mental fictionalism is not self-refuting.<sup>8</sup> But mental fictionalism would have us eschew the *thinking* self straightaway, end of discussion. So I will not be assuming mental fictionalism in what follows. My aim, nevertheless, will be persuade you to resist a commitment to the thinking self, which is ultimately part of a defense of mental fictionalism, and part of my commitment-phobia regarding ontology generally.

---

<sup>7</sup> For examples of Lycan’s Moorean approach, see Lycan (2001; 2007.)

<sup>8</sup> See Parent (2017, ch. 9), Joyce (2013), Wallace (this volume), etc.

#### 4. The Paronymy of NPs (Lycan Rides Again)

In discussing the self, it is necessary to present one further preliminary. It is crucial to appreciate that noun phrases (NPs) in natural language seem to be equivocal in widespread and systematic ways, including NPs that refer to oneself. This something Lycan himself has made clear in recent work (Lycan 2017). Lycan introduces the basic phenomenon as the *paronymy of proper names*, where “paronymy” is ambiguity that is not *sheer* ambiguity (as when different people both share the name ‘Dan’). As an illustration, Lycan (p. 407) asks us to consider the name ‘Germany’ in the following uses:

(G1) Germany is east of France.

(G2) Germany voted Social Democrat.

(G3) Germany invaded Austria.

(G4) Germany won the World Cup.

(G5) Germany loves potatoes.

Thus, consider that the relevant landmass east of France is incapable of voting, invading, winning matches, or loving potatoes.

Lycan observes, moreover, the paronymy of proper names creates confusion about identity in relation to persons (p. 408). Consider that Terry Schiavo died on March 31<sup>st</sup>, 2005. But loved ones might have rightly said that she left us before then, due to entering a persistent, vegetative state. Such a post-psychological condition is sufficient for her absence. In which case, the following has a true reading even if her living body is present in the room:

(TS1) Terry Schiavo is no longer with us.

Even so, it can be simultaneously true that:

(TS2) Terry Schiavo is at Pinellas Park hospice.

The explanation, again, is that the name ‘Terry Schaivo’ is paronymous; in particular, it refers to the living body in one case but not in the other.

For that matter, consider that (TS2) remains true on one reading, even just after the body dies. The name refers to the body regardless. But simultaneously, of course, there would be an even more palpable sense in which (TS1) is true.

Lycan further notices that such paronymy is not limited to proper names, but can arise with a variety of NPs, including the first-person indexical. He has us consider the following sort of contrast (p. 413):

(Me1) After I die, I will no longer exist.<sup>9</sup>

(Me2) After I die, I will be buried in the ground.

Here too, the indexical ‘I’ seems to shift in reference from the mind to the body.

In addition, Lycan observes that such paronymy explains some of the confusion with the “animalism” debate, i.e., the debate on whether persons are living animals. (Carter 1989, Snowdon 1990, Olsen 1997, Blatti 2012.) Animalism has very strong appeal, but it would seem to dictate that (TS1) and (Me1) are false, as long as the living body is present. Lycan’s response (p. 413) is that it simply depends on how interpret the relevant NP, and I think he is right about that.

So again, the paronymy of NPs can cloud the nature and existence conditions of the self, even as concerns synchronic existence. In fact, the paronymy goes beyond the distinction between the body vs. the mind. For there are cases where the self seems to be isolated to a proper *part* of the mind, e.g., when another part of the mind is introspected as an object. In such a case,

---

<sup>9</sup> Questions of post-mortem existence are not meant to be summarily dismissed here. So we might instead see (Me1) as short for: After I die, I will no longer exist in the same way, viz., in an embodied form.

the object appears distinct from the psychological subject, whence “the self” is not identical with the (entire) mind.

More concretely, consider that the following is often true in my own case:

(Me3) When I am anxious, I calmly and mindfully observe the feeling and experience relief.

In such cases, the feeling of anxiety does not completely disappear. So in one sense “I” am continuing to feel anxiety. And yet in another sense, “I” do not feel anxiety, as long as I calmly and mindfully watch the feeling. In such cases, the anxiety seems to become an object of thought, and thus becomes separate from me, as the thinking subject. Once the feeling is thus compartmentalized, there is a clear sense in which “I” am no longer anxious.<sup>10</sup>

If experiments in mindfulness do not appeal, similar examples arise with cases of critical self-reflection. Consider:

(Me4) When I start believing in a perfect God, I catch myself and curtail it.

These are cases where “I” intentionally try to alter what “I” believe. And the “who” that acts against the belief in naïve theism is not a naïve theist. Yet it is “I” who is sometimes a naïve theist, despite myself. The way to make sense of this, again, is to think of ‘I’ as paronymous vis-à-vis different parts of my mind.

The suggestion, then, would be that for each of (Me3) and (Me4), the second indexical differs slightly in reference from the first. After all, it is not as if every single neuron in your brain is needed to think a thought. In the case of naïve theism, then, the belief is likely realized by a subgroup of neurons. Other neurons can then manifest the introspective check on the belief;

---

<sup>10</sup> Lycan (in conversation) reports not hearing the paronymy in (Me3). Perhaps we are habituated to attributing contrary emotions to a single subject (anxiety and calm), where strictly and literally speaking, *I* am both calm and not calm. Nevertheless, I would submit that on at least one interpretation of (Me3), there seems to be a subject of anxiety, and something distinct experiencing calm.

these other neurons are then the “me” who curbs “myself.” In which case, paronymy in ‘I’-talk arises even in reference to the mind.

Such intra-psychic paronymy may also explain, why many researchers now talk of multiple selves. In a widely cited article, Ulric Neisser (1988) distinguishes between the ecological self, the interpersonal self, the extended self, the private self, and the conceptual self. Here are some examples based on Neisser’s observations:

(UN1) “When you touch my shoulder, you are touching *me* even if a shirt and jacket interpose between your fingers and my skin.” (p. 39) [Ecological self]

(UN2) If we write Neisser a letter, we are addressing *him*, even if (unbeknownst to us), he died several years ago. [Interpersonal self]<sup>11</sup>

(UN3) Neisser is (now!) identical to a person who gave a talk in Aberdeen in 1998 (p. 47), even though Neisser died several years ago. [Extended self]

I expect such cases to be controversial, but *if* we accept the different “selves,” a puzzle arises. An individual human organism is supposed to have each kind of “self.” However, if Neisser has five different selves, which one is *him*? (There can only be one.) Neisser seemed to appreciate this conceptual difficulty only dimly; ditto for many of those who cite him.<sup>12</sup>

However, the difficulty can be resolved, once we recognize the paronymy of NPs. Again, it just depends on how we interpret the relevant terms, and as Lycan implies, it can be reasonable to interpret them in different ways, depending on whatever goals and interests we have at the time.

---

<sup>11</sup> Your *idea* of Neisser is perhaps key in explaining your letter-writing behavior. Regardless the letter is being addressed *to him* and not just to one of your own ideas.

<sup>12</sup> Neisser raises something like this issue in his concluding paragraph; however, his answer suggests that he treats the question as “how do I come to *represent* myself as one person?” since he talks here of the cognitive integration of information from disparate sources. But the question I am raising is “How can I *be* a single person if I have five selves?” The answer, I think, is that in one respect “I” am *not* a single person but a disunity (like “Germany”) and that in another respect, which self is “me” depends on the goals and purposes adopted in the context.

## 5. Fictionalism about the Self.

So talk of a “person” or a “self” is rather like talk of “Germany.” Different parts or aspects of the German totality figure into the different present-tense truths about “Germany,” viz., the German landmass, the German electorate, the German military, the German football team, etc. Similarly, different parts or aspects of the “you” totality figure into different present-tense truths about you. Your mind, your body, your body plus your clothes, etc.

In light of such observations, there is a simple argument for fictionalism about “you.” For consider that we should be fictionalists about Germany, if “Germany” is understood to be a single entity which satisfies all the different predicates in (G1)–(G5). (Let us forbid wildly disjunctive or gruesome entities.) We speak as if there is one such entity, but that appears to be just a convenient *façon de parler*. Similarly, there is no single entity which satisfies the different predicates in (Me1)–(Me4), or (UN1)–(UN3), etc. For instance, there is no single spot in the brain which feels anxiety, is mindful of anxiety, occasionally believes naïve theism, occasionally curtails such belief, etc. At best, we have a Dennettian hodgepodge of subsystems or modules, and none is the unequivocal reference of ‘I’.

This simple argument looks convincing as far as it goes. But it may not go that far. After all, even if there is no such thing as “Germany” as such, there still exists a German landmass. And there still exists a German military. Etc. Similarly, the simple argument may just show that our talk of “I” is rather loose and can refer to a variety of things. There may be no one thing that counts as “I,” yet various “I-related” phenomena still exist.

But, waiting in the wings is a more interesting type of fictionalism about the self. Let us focus on the satisfier of (Me1). What is it that ceases to exist at the time of death? Certainly not the body. Yet assume that physicalism is true. Then, if I am not just a fiction, I am just the body,

or at least a part of the body. But no part of the body ceases to exist at death. However, since I cease to exist, this must mean I am not any part of the body. So the thing that ceases to exist at death must have only existed in fiction.

We can reconstruct the reasoning here as follows:

- (1) If I exist outside of fiction, then I am identical to (some part of/) this biomass [= my body].
- (2) If I die at  $t$ , I cease to exist at  $t$ .
- (3) If I die at  $t$ , no part of this biomass ceases to exist at  $t$ .
- (4) Therefore, no part of this biomass is identical to me. [From (2), (3)]
- (5) Therefore, I do not exist outside of fiction. [From (4), (1)]

Some objections and replies are now in order.

The first premise is based on the physicalist idea that if I exist outside of fiction, then I must be something physical, presumably, something bodily. But “physical” existence might not be limited to *bodily* existence. After all, the physical includes not only matter but also forces, including electromagnetism, the force behind the attraction and repulsion of chemical ions within the firing of synapses. Upon death, thoughts stop because the synapses stop firing, meaning ultimately that certain electromagnetic forces disperse or fail to be coordinated in the necessary way. Thus, after death, “I” will not exist, in the sense that there will fail to be the right sort of pattern or arrangement of electromagnetic forces in the brain, of the sort that realizes a mind.

On second thought, however, the pattern of electromagnetism seems constitutive of a *mental state*, and not constitutive of *mind*. If the mental state in question is a thought, for instance, then the electromagnetic pattern in the brain is the token thought state, and not the thing that *produces* the thought. In attending to electromagnetic patterns, we have conflated the product with the producer.

The producer of the electromagnetic pattern, moreover, is the brain. So really, on this physicalist picture, what ceases to exist at death is only the thoughts and not the *thinker*. The brain still exists and is buried in the ground after death; accordingly, if 'I' refers to the brain in (2), then (2) is false. But (2) is not false. Such a physicalist view therefore appears to misfire.

However, is it fair to say that the brain is what *produces* thought? It may be more that the brain is where electromagnetic forces "assemble" into the right sort of pattern, when a thought token comes into existence. True enough, various contours of the brain might be instrumental in bringing about the right electromagnetic pattern. Yet the nature and magnitude of electromagnetism in the brain also partly explains how the patterns come about.

One might compare it to explaining a whirlpool in the bend of a river. The contours of the riverbed are instrumental causing the whirlpool, but it's not the whole story. The physical features of water are also important to explaining the whirlpool. And when those features change in wintertime (when the water turns to ice), the whirlpool ceases to exist.

So what *produces* the whirlpool? Well, really, it's better to say that the whirlpool is the product of various features of the riverbed and features of (liquid) water itself. It's a bit misleading to say there is a separable *thing* that creates the whirlpool. Similarly, when it comes to the "thing that thinks," it seems better to say that a thought is not really created by some distinct entity. A thought is instead a pattern of electromagnetism, explainable by the arrangement of molecules in the brain and their physical features, and by the dynamical features of intermolecular forces. These variegated, multi-level conditions are not *produced* by the brain, except in a loose way of speaking. The brain is more the "favorable environment" in which such conditions all co-assemble.

But if this is what we really want to say, then it supports rather than undermines that the thinker is unreal. For we are now saying that there is no single entity which produces thinking, except in a loose way of speaking.

This might prompt us to backpedal a bit, reverting to the view that the brain, strictly and literally, is the thing that thinks. Even so, perhaps the thinking thing is distinct from the psychological subject, the thing that *experiences* thinking. This would fit nicely with the intuition that only the psychological subject has mortal importance. (What do I care if my *brain* continues to exist after death?) Yet for a physicalist, what in the world is this psychological subject, if not something between the ears?

This can prompt a different line where the self is identified with a *self-concept* or *self-conception*. We certainly have a concept such that (i) the “self” produces thoughts, and (ii) the “self” ceases to exist upon death. Both of these conditions are not satisfied by the brain, but the objector is now suggesting that the *self-concept* indeed satisfies both conditions. However, this conflates use and mention as regards mental representation: The self is not the same thing as the self-concept, but is rather the thing represented by the self-concept. In addition, a self-concept is produced by the thinker, yet assuming there is no circular causation here, a self-concept would not produce itself. Thirdly, the brain can host more than one self-representation at a time, yet a brain is not supposed to host more than one self at a time (paronymy aside).

Yet another objection would be that the self is really the *mereological fusion* of this biomass (or a part thereof) with the neuroelectricity that dissipates upon death. Upon the dissipation, the fusion would cease to exist as well, even though no part of the biomass ceases to exist. The problem, however, would be that the neuroelectricity which comprises a thought would be *produced by an entity which has that very thought as a metaphysical part*. Or, to use

the fashionable lingo, a thought would be partly “grounded” on itself, if the thought is partly constitutive of the thinking self. Such circular grounding is not usually looked upon favorably.<sup>13</sup>

## 6. Self-Location

Fictionalism about the self is counterintuitive to a degree, but why is that? Why, for example, might Lycan have the Moorean intuition that the self is undeniably real? I suspect it owes to the vivid sense that the thinker has a spatiotemporal location. But this needs to be qualified a bit. As Hume famously says, we never catch a glimpse of the self in inner sense. So whence the impression that we can pinpoint its location?

It can help with this puzzle to consider the street-view feature of Google maps. The street-view feature represents a location as if from a particular perspective. One can rotate the frame, thus creating an impression much like moving your head to take a look around. One thus gains the sense that there is a particular *spot* around which the visual representation is rotating, a hub around which the panoramic representation is being rotated, although the hub can never be part of the scene. (If it were, it would no longer be the hub.) But in the visual scene, there are subtle perspective lines that point back toward panoramic center, and this is what seems to provide a sense of self-location in Google maps. Something analogous seems to occur in our own non-virtual experience as well.

When it comes to non-virtual experience, is the panoramic center the location of the self? If so, then the self would have a location, and hence, exist in physical reality. I do not doubt that

---

<sup>13</sup> A better version of the proposal is this. The thinker at  $t_1$  = the thoughts at  $t_1$  fused with the brain at  $t_1$ . That fusion indeed produces the thoughts at  $t_2$ , which in turn, are a metaphysical part of the thinker at  $t_2$  (along with the brain at  $t_2$ ). But there is no circular causation here. For the thinker at  $t_1$  produces only the thinking at  $t_2$ , and not the thinking at  $t_1$ . Nonetheless, in order to avoid a regress, one must posit thinking at  $t_0$  which was not produced by a thinker at a previous time, yet partly constitutes the thinker at  $t_0$ . So the view has the funny consequence that there could be “thinkerless thinking” in the human organism—i.e. a thought without someone thinking the thought.

one's own phenomenological hub corresponds to a location in physical reality. And that location is a point in the brain, which is entirely real (as far as "points" go). But is that the self? I doubt it. At least, this point in the brain does not cease to exist at death, even though something important does.

Yet the phenomenological hub certainly *feels* like the abode of the psychological subject, the thing that is mortally important to "me." But we can erode this feeling by reconsidering the extended thought experiment from Daniel Dennett's (1981) "Where Am I?" The thought experiment is essentially an illustration of how the phenomenological center can patently *fail* to correspond to the self.

One might recall that in Dennett's tale, his brain is surgically removed from his body, to be preserved in a vat of nutrients, and yet:

each input and output pathway as it was severed, would be restored by a pair of microminiaturized radio transceivers, one attached precisely to the brain, the other to the nerve stumps in the empty cranium. No information would be lost, all connectivity preserved. (p. 311)

The philosophical drama starts to unfold when Dennett's body walks over to the vat to take a look at his brain.

I thought to myself: "Well, here I am sitting on a folding chair, staring through a piece of plate glass at my own brain . . . But wait," I said to myself, "shouldn't I have thought, 'Here I am, suspended in a bubbling fluid, being stared at by my own eyes'?" I tried to think this latter thought. I tried to project it into the tank, offering it hopefully to my brain, but I failed to carry off the exercise with any conviction. I tried again. "Here am I, Daniel Dennett, suspended in a bubbling fluid, being stared at by my own eyes." No, it just didn't work. Most puzzling and confusing. Being a philosopher of firm physicalist conviction, I believed unswervingly that the tokening of my thoughts was occurring somewhere in my brain: yet, when I thought "Here I am," where the thought occurred to me was *here*, outside the vat, where I, Dennett, was standing staring at my brain. I tried and tried to think myself into the vat, but to no avail. (p. 312)

The story is quite compelling, yet why should the protagonist be so inclined to see himself as located outside the vat? I suspect it is because the protagonist's phenomenology still points

toward a center that is located inside the skull. And he is habituated into thinking that this phenomenological center is *his* location. But in the story, no point inside the skull corresponds to any region of the brain. And so, the protagonist is simultaneously attempting to understand himself as located elsewhere, viz., inside the vat. But the feat is difficult.

This shows the strength of the belief that the location of the self = the phenomenological center. And it also illustrates that this belief can be resisted: If these locations were literally numerically the same, then it seems unlikely that they could be separated in Dennett's narrative in such a compelling way. (The point is not conclusive, but it provides a basis for doubt).

If so, the phenomenology that allegedly discloses a self-location is non-demonstrative. Per the paronymy of NPs, it *can* correct to say the "self" has a location, e.g., when the relevant NPs refer to the body or a part thereof. But the phenomenological center may not be where the self resides. To use the earlier analogy, the center of the whirlpool is not the location of what produces the whirlpool. Indeed, there is really no single locus where the whirlpool is birthed—and in the same way, I have tried to make plausible that the psychological subject is not located anywhere in particular, except perhaps according to a fiction.

Thus, the sense of a self-location is no proof against fictionalism about the self. Some fictions are more cognitively ingrained than others, and the sense of self-location seems to be a case in point. The upshot, again, would be that the self does not exist outside of fiction. But since talk of "you" and "I" is largely unavoidable, we should therefore be fictionalists about the self.

## References

- Burgess, J. & G. Rosen (1997). *A Subject with No Object*. Oxford: Clarendon Press.
- Carnap, R. (1950). Empiricism, semantics, and ontology. *Revue Internationale de Philosophie* 4: 20–40.
- Carter, W. (1989). How to change your mind. *Canadian Journal of Philosophy* 18: 1–14.
- Dennett, D. (1981). Where am I? In his *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge, MA: MIT Press.
- Joyce, R. (2013). Psychological fictionalism. *The Monist* 96: 517–538.
- Lycan, W. (2001). Moore against the new skeptics. *Philosophical Studies* 103: 35–53.
- \_\_\_\_\_. (2007). Moore's antiskeptical strategies. In S. Nuccetelli & G. Seay (eds.), *Themes from G.E. Moore: New Essays in Epistemology and Ethics* (pp. 84–99). Oxford UP.
- \_\_\_\_\_. (2018). Metaphysics and the paronymy of proper names. *American Philosophical Quarterly* 55: 405–419.
- Nessier, U. (1988). Five kinds of self-knowledge. *Philosophical Psychology* 1: 35–59.
- Nolan, D. (2002). *Topics in the Philosophy of Possible Worlds*. New York: Routledge.
- Olson, E. (1997). *The Human Animal: Personal Identity Without Psychology*. Oxford UP.
- Parent, T. (2017). *Self-Reflection for the Opaque Mind: An Essay in Neo-Sellarsian Philosophy*. New York: Routledge.
- \_\_\_\_\_. (ms.) Conservative Meinongianism: An Actualist+ Meta/Ontology. Available at <http://tparent.net/ConservativeMeinong.pdf>.
- Quine, W. (1948). On what there is. *Review of Metaphysics* 2: 21–38.
- Rosen, G. (1990). Modal fictionalism. *Mind* 99: 327–354.

Snowdon, P. (1990). Persons, animals, and ourselves. In C. Gill (ed.), *The Person and the Human Mind: Issues in Ancient and Modern Philosophy*. Oxford: Clarendon Press.

Wallace, M. (forthcoming). Mental fictionalism. In T. Demeter, T. Parent, & A. Toon (eds.), *Mental Fictionalism: Philosophical Explorations* (pp. xx-xx). New York: Routledge.