



The Very Idea of Mental Anti-Representationalism

Huang R*

Department of Philosophy, University of Florida, USA

*Corresponding author: Rusong Huang, Department of Philosophy, Adjunct Instructor in Philosophy, University of Florida, USA, Tel: 352-871-6558; Email: rusongh@gmail.com

Review Article

Volume 7 Issue 4

Received Date: May 16, 2024

Published Date: December 31, 2024

DOI: 10.23880/phij-16000342

Abstract

In this article, I will introduce the idea of mental anti-representationalism (MAR) that I defended. According to MAR, psychological sentences are not representational. The article has four sections. I will first clarify MAR (“Three Clarifications about the Thesis of MAR”) and explain it with the help of the view of noncognitivism or expressivism in metaethics (“Metaethical Noncognitivism, Expressivism and MAR”). Like noncognitivism, MAR is a negative thesis. However, the positive thesis of MAR is not that psychological sentences express some non-cognitive or desire-like attitudes, but that they are a type of rationality sentence. I will then compare MAR with other views of mind on the market, such as mental eliminativism and mental fictionalism (“Mental Eliminativism, Fictionalism and MAR”). MAR rejects eliminativism and improves fictionalism. Finally, I will outline my main argument for MAR and address some challenges (“My Master Argument Outline and Some Challenges”). My argument relies on the uncodifiability thesis of rationality and my view can avoid what I will call the Question-Begging Problem.

Keywords: Mental Anti-Representationalism; Interpretationism; Eliminativism; Fictionalism

Introduction

It is a widely received view that psychological sentences represent mental facts. I argue against it. I call my view Mental Anti-Representationalism (MAR) [1]. According to MAR, a psychological sentence (e.g., “I believe that the earth is round.”) does not represent, refer to, or correspond to, the fact of my believing that the earth is round. Accordingly, the meaning of the sentence should not be understood in terms of the representational relations between the sentence and what it represents, a belief fact in this case.

Three Clarifications about the Thesis of MAR

First, MAR is not the view that psychological sentences do not represent anything or are not about anything. Or, to put it another way, MAR is not the view that a psychological sentence is not representational simpliciter. For example, the sentence “S believes that snow is white” clearly represents S

or is about S. Rather, what they do not represent are mental facts. The notion of fact here is used in a robust sense in which, for example, it is not a fact that Sherlock Holmes (the fictional character) lives in 221b Baker Street, London¹. By saying “psychological sentences are not representational,” I mean that the meaning of a psychological sentence should not be understood in terms of the representational relation between the sentence and what it represents. And by “what it represents,” it is not supposed to be “what it represents, whatever it is.” It is theoretically loaded. What it represents is

1 The example is from Lewis [2]. We may say that, in the fiction, it is a fact that Sherlock Holmes lives in 221b Baker Street, London. I do not deny that. However, I am only concerned with facts in our world, if a restriction has to be added. In our world, for example, it is a fact that Arthur Conan Doyle wrote some stories about Sherlock Holmes. However, in our world, it is not a fact that Sherlock Holmes lives in 221b Baker Street, London. Or, more precisely, the notion of fact is not applicable to the sentence (see more in “Mental Eliminativism, Fictionalism and MAR”).



supposed to be a fact (in a robust sense) that is ontologically on a par with a physical fact.

Second, the thesis is a linguistic inquiry. It is about psychological sentences. If MAR is correct, then the meaning of psychological sentences cannot be affected, let alone determined by those mental facts in any way, making them categorically different from the meaning of representational sentences, such as descriptive sentences. Since what is in question is about the reference or the meaning of psychological sentences, I cannot presuppose the existence or the non-existence of mental facts. As a methodology, I will start with so-called mental interpretationism, according to which whatever mental states something has are equivalent to those mental states they can be attributed or interpreted as having by us. The view understood in that way is an ontological thesis. Or, it is the view that there is no real gap between the epistemological study of mentality and the ontological study of mentality. At a certain point, they converge together.

The difference between MAR and mental interpretationism is that MAR does not necessarily accept the ontological thesis stated above. For example, MAR can leave room for the possibility that there are some mental states that cannot be accessed by attributions or interpretations in any way. The point of MAR is that even if there are mental states, their existences have no effect on the meaning of any psychological sentence. MAR, as a linguistic inquiry, keeps neutral about all kinds of ontological questions about mind. Or, MAR avoids ontological commitments. This is also where MAR is different from views such as mental eliminativism, which claims that there are no mental states (because they can be eventually eliminated from a mature scientific theory of mind).

What MAR tells us is merely that to understand the meaning of psychological sentences, no mental facts should be consulted. MAR is indeed consistent with the view that there are no mental facts. If MAR is correct, it gives us one reason to support the idea that there are no mental facts. But it does not necessarily compel us to accept that there are no mental facts as a consequence, but only that even if there are mental facts, they will not affect how we understand those psychological sentences.

Third, MAR is clearly a negative thesis. It by itself leaves a positive story open. It is useful to distinguish a negative thesis from a positive thesis. To compare: noncognitivism in metaethics is a negative view; expressivism in metaethics is a positive view; though expressivism entails noncognitivism, not vice versa (see more in the next section). As a negative thesis, MAR can be at least, in principle, supported without appealing to a positive story.

Metaethical Noncognitivism, Expressivism and MAR

We may introduce MAR in the philosophy of mind by considering the view of noncognitivism or expressivism in metaethics. Metaethical noncognitivists think that moral sentences are not representational. For example, the sentence "Lying is wrong" does not represent the moral fact that lying is wrong; the sentence "Sorting out the trash is right" does not represent the moral fact that sorting out the trash is right. Noncognitivism contrasts with commonly known as (metaethical) descriptivism. According to the latter, a moral sentence is similar to a descriptive sentence, such as "Snow is white." It is descriptive or representational just as the sentence "Snow is white" is descriptive or representational; it describes or represents a moral fact just as the sentence "Snow is white" describes or represents the fact that snow is white.

I mentioned that the view of MAR is a negative thesis, and so is noncognitivism, as the term suggests. As a negative thesis, the view of MAR in the philosophy of mind and the view of noncognitivism in metaethics converge as follows,

The Negative Thesis:

MAR: Psychological sentences are not representational.

Metaethical Noncognitivism: Moral sentences are not representational.

Although both views will be incomplete without some positive stories, the negative thesis can be separate from the positive one. And as a positive thesis, the two views diverge as follows,

The Positive Thesis:

MAR: Psychological sentences are a type of what I will call rationality sentences.

Metaethical Expressivism: Moral sentences express a certain type of mental state.

Noncognitivists in metaethics think that moral sentences express our non-cognitive or desire-like attitudes. That is why the positive view is labeled as "expressivism." For example, the sentence "Lying is wrong!" expresses the speaker's disapproval of lying; the sentence "Sorting out the trash is right" expresses the speaker's approval of sorting out the trash. Now, the label "expressivism" is not ideal. It is a technical term. To see this, consider again the descriptive sentence "Snow is white." It is clear that we can also use it to express the speaker's attitude, the belief that snow is white. So, the idea of expressivism must not merely be that those sentences in question are expressive. Rather, the key is the type of mental states expressed. Gibbard thus concludes that,

The term 'expressivism' I mean to cover any account of meanings that follows this indirect path: to explain the meaning of a term, explain what states of mind the term can be used to express [2].

In fact, the point that moral sentences and descriptive sentences express different types of mental states is often used to explain why descriptive sentences are representational but moral sentences are not. That is, the positive thesis is often used to explain the negative thesis. If that is correct, the positive thesis will be the core, and the negative thesis will be the secondary. The content of a cognitive mental state (what is believed, for example) is supposed to fit how things are in the world and thus can be true or false in that sense. But the content of a non-cognitive mental state (what is desired, for example) is not supposed to fit how things are in the world and thus cannot be true or false in that sense.

However, I think that is a mistake. The positive thesis is neither sufficient nor necessary for the negative thesis. It is not sufficient because some descriptive sentences can also be used to express some non-cognitive attitudes. For example, the descriptive sentence "You made a mistake again." can be used to express discontent, which is a type of non-cognitive attitude. What expressivists need seems to be that moral sentences can only be used to express non-cognitive attitudes. But this further step needs some other justification. On the other hand, it is not necessary because, as I mentioned, the negative thesis can be separated from the positive thesis. The negative thesis can be defended without the positive thesis.

The positive thesis of MAR is not that psychological sentences express some non-cognitive or desire-like attitudes. Instead, I argue that psychological sentences are a type of rationality sentence (See more in "My Master Argument Outline and Some Challenges"). The underlying idea is that attributing mental states is rationalizing some behaviors. And I argue that the point that psychological sentences are identified with rationality sentences explains why psychological sentences are not representational, the negative thesis of MAR. If rationality sentences are not representational, then psychological sentences are not representational.

To argue for MAR, the basic question I ask is, "What do we do when we attribute mentality to an attributee S?" Here is my quick answer: we take S as one of us, i.e., as a rational agent, and rationalize S's behaviors (verbal or non-verbal). Notice that the answer leaves open whether non-human animals can have mentality. Some might say that clearly, animals have mentality, at least some kind of mentality, for example, that animals can have sensations such as pain.

Then, in my view, when we attribute the sensations to animals, we have to attribute at least some kind of rationality to those animals².

Mental Eliminativism, Fictionalism and MAR

Let us compare MAR with other views of mind on the market, such as mental eliminativism and fictionalism. Eliminativists typically think that psychological sentences are false. That is why they can be eliminated from the mature scientific theory of mind, mature neuroscience, for instance [4]. To support eliminativism, Churchland argues that folk psychology is a theory; since it is a kind of theory, it is at least possible that the theory is false. Given how poorly folk psychology deals with some mental phenomena, it would not be surprising if it really turns out to be false. Eventually, folk psychology will be superseded by neurosciences, just as alchemy is superseded by chemistry.

Churchland's view is quite implausible. Churchland says that compared to the achievements of physical sciences, folk psychology achieves little. It has not changed for thousands of years. And without presupposing some other mental states that also need explanation, the prediction power of folk psychology is none. While those points Churchland made are correct, they can also be used to show that folk psychology is fundamentally or categorically different from physical sciences. Even if folk psychology is a kind of theory (in a broad sense), it can be radically different from scientific theories, so different that it is no wonder that it achieves so little from the perspective of a scientific theory. Churchland makes a category mistake from the perspective of MAR. We should distinguish between true/false (TF) in a robust and deflationary sense. If psychological sentences are not representational, then they cannot be TF in a robust sense, although they can be TF in a deflationary sense. Clearly, Churchland uses TF in a robust sense. In that sense, psychological sentences cannot be in the business of TF.

In fact, the idea that psychological sentences can be eliminated from mature neuroscience is equivocal between the view that psychological sentences are false and the view

² Rationality is not an all-or-nothing thing. It is a matter of degree. The notion of rationality should be understood very broadly. In that sense, attributing to an animal a mental state would attribute to it some sort of rationality. Consider a case in which I say, "The cat is hungry." What I say would mean that provided the circumstances, it makes sense for the cat to do such and such, e.g., searching for some food. There will be some rational (in the sense of making sense) relation between the mental state I attribute (feeling hungry in this case) and what the cat would do. Further, when we attribute a mental state to a cat, we more or less project our own situation on to the cat. What we mean is that if we were in a situation as that of the cat, we would feel hungry. If that is the case, we implicitly attribute some sort of rationality to the cat, provided that we ourselves are rational creatures.

that TF is not applicable (NA) to psychological sentences. If NA is the right view, psychological sentences can also be eliminated from a mature theory (indeed, any theory) because, by definition, a theory can be TF.

That is, MAR can accept the point of eliminativism: psychological sentences can be eliminated from mature neuroscience. It is just that even mature neuroscience cannot explain or exhaust all mental phenomena.

Let's turn to mental fictionalism. While, as far as I know, all fictionalists do not endorse MAR, I think that the gap between MAR and mental fictionalism is much smaller than they thought. While the point may not give MAR much credit, this is at least attractive for those who are sympathetic to mental fictionalism.

According to Demeter et al., Mental fictionalists commonly agree with three aspects: Factualism (the view that psychological sentences can be TF), Utility (the view that psychological sentences are useful and thus cannot be eliminated), and Storytelling (the view that psychological sentences can be employed as a kind of storytelling) [5].

MAR shares the latter two aspects with mental fictionalism. First, mental fictionalism also differs from mental eliminativism in that mental fictionalism does not commit us to the existence or the non-existence of mental states, and thus avoids ontological commitments. The basic idea of mental fictionalism is that psychological sentences should be understood as a kind of storytelling. As Demeter et al. point out, mental fictionalism is supposed to be a compromise position between mental eliminativists and their opponents. It might be the case that the story is about something that exists in the world. But it might not be. Mental fictionalism does not make a choice between them. Second, mental fictionalists also endorse the indispensability of psychological sentences. In our social practices, attributing mental states and thus using psychological sentences can be necessary for explaining and predicting our behaviors.

MAR apparently differs from mental fictionalism in Factualism. However, fictionalists cannot really have their cake and eat it at the same time. With the distinction between TF and NA in hand, we know that, strictly speaking, "TF in a fiction" is incoherent. Things in a fiction are not a part of the world. Indeed, in our daily practices, sometimes when we say that it is a fiction, we mean that it does not exist. Now my point is not that we should not say "TF in a fiction" but that we should be aware of the distinction between the robust sense of TF and the deflationary sense of TF. In a deflationary sense, it would be OK to say "TF in a fiction." However, it would be clear that, in that case, fictionalism is not really a middle way between mental eliminativists and their opponents. Similarly,

the disagreement between MAR and mental fictionalism on Factualism is merely apparent. Fictionalists do not think that psychological sentences represent facts in the robust sense. So, strictly speaking, they do not think that TF is applicable to psychological sentences, either.

My Master Argument Outline and Some Challenges

How may the view of MAR be defended? I propose the following master argument for MAR³:

- Psychological discourse is a kind of rationality discourse.
- Rationality discourse is not representational.
- Therefore, psychological discourse is not representational.

For the first premise, the key idea is that to attribute a mental state is to rationalize some behaviors. The idea is implicit in the writings of Donald Davidson [7] and Daniel Dennett [8,9]. How can we cash it out? I propose the following view:

S is in a mental state M \equiv There are some circumstances C such that S is in C, and it is rational for anyone in C to do certain something.

I assume that attributions of mentality are essential to our understanding of psychological sentences. Some might think that whether S thinks (for instance) that global warming is a problem or not seems to be independent of whether there is an attributor who attributes the thought to S. But then we assume that mental states are ontologically on a par with physical states. For a physical state such as the atomic mass of iron being 56, it seems to be irrelevant whether there is an attributor who attributes the property to iron or not. However, that is the very contention between MAR and its opponents. I propose suspending the assumption and seeing whether there is some independent justification for it instead. In fact, there are some reasons for us to doubt it. If our concern is about psychological sentences, it is very unlikely that to analyze the meaning of a sentence, we can simply adopt the way in which we write the sentence on a board and ask about the meaning of it. In that way, it seems that the sentence can be totally detached from its context and that we can analyze the meaning of it just like we can analyze the properties of iron by having iron in front of us. An alternative way to analyze the meaning of a sentence, the Wittgensteinian way, is to look at how the sentence is used. Similarly, I propose that when we analyze mentality

³ Gibbard [6] uses a similar one to argue for his view that moral discourse is not representational.

or mind, we do not merely write psychological sentences on board and ask about their meanings of them. Instead, we examine how those psychological sentences are used and thus how mental states are attributed by those psychological sentences. The hope is that we will understand mentality better by understanding attributions of mentality better.

For the second premise, the key idea is the uncodifiability of rationality thesis. To codify something is to give a code of rules or a rulebook for it. Rationality is codified if and only if being rational requires following that code for every situation. Or, more precisely:

Rationality can be codified if and only if for every possible circumstance C and possible action A that could be performed in C, there are some explicit rules stating whether doing A in C is the rational thing to do.

The uncodifiability thesis denies such explicit rules governing rationality. Suppose that is the case. How can we entail premise two? The following argument is proposed:

- If a rationality sentence is representational, then rationality can be codified.
- Rationality cannot be codified. (The Uncodifiability Thesis)
- A rationality sentence is not representational.

For the first premise, my thought is that if representationalists accept the truth-conditional semantics (which is presupposed in the representationalists' model), then there will be some rules determining the meaning of a rationality sentence. Then, rationality can be codified. Of course, those arguments here are just mentioned and should be fully developed. For now, let us turn to the biggest challenge to my view.

The key idea underlying the equivalence between a mentality sentence and a rationality sentence (the first premise in the Master Argument) is that attributing a mental state is rationalizing some behaviors. At first glance, "attribute" or "rationalize" is also mentally loaded. How can we understand the mental state of attributing or rationalizing? Either we understand it by another attributor, or we do not. If we understand it by another attributor, then we will have an infinite regress problem because the other attributor would also have the mental state of attributing or rationalizing. If we do not, then notions as attributing or rationalizing would have to be understood as not depending on the interpretation or as some intrinsic properties of the attributor, so to speak. Both options seem unsatisfactory.

The problem is not peculiar to my view. To take Dennett's view of the intentional stance as an example, for Dennett, a

system is an intentional system if and only if we can adopt an intentional stance towards it. That is, an interpreter is required for us to understand an intentional system. But then, how can we understand the mental states of the interpreter? Adopting an intentional stance is itself an intentional state.

Similarly, as another example, fictionalists think that we should treat mental states as a fiction, i.e., to adopt a fictional stance. But adopting a fictional stance is itself an intentional state. The problem in the literature on fictionalism is sometimes called "the Cognitive Suicide Problem" [10] or "the Cognitive Collapse Problem" [11]. I will simply label it as the Question-Begging Problem. The problem is that the existence of mental states (i.e., adopting, treating, attributing, or rationalizing) seems to be presupposed when I say that certain stances are adopted, some mental states are attributed, or some behaviors are rationalized. And that is question-begging.

I will first talk about fictionalists Meg Wallace and Adam Toon's solutions to the Problem and then argue that I can avoid it in some similar way⁴. Wallace does not think that fictionalism really has the problem [13]. As she characterizes the problem, the issue is that the distinction required by fictionalism between fictional talk and non-fictional talk relies on mental attitudes but that, strictly speaking, there are no such mental attitudes (because they are fictional). Wallace's answer is that fictionalists do not say that there is absolutely nothing going on when we are fictionalizing. According to her, the point is just that the mental state of fictionalizing does not uniquely pick out any brain activity that is, in fact, occurring.

While I think that talking about brain activity can be misleading because a fictionalist does not think that brain activity is in any way essential, I take it that Wallace's idea is that the distinction between fictional talk and non-fictional talk can be made without appealing to a mental fact. The point that there is no uniquely corresponding brain activity for fictionalizing is supposed to show that no mental fact like that can be used to explain fictionalizing. Fictionalists do not need to deny that there may be some brain activities. But as merely physical facts, they are part of the circumstances the person is in.

⁴ It is not clear how Dennett would address the issue. But when Robert Brandom talks about Dennett's issue of so called "stance stance", that is, to adopt an intentional stance is also a stance. Brandom provides his own answer. Brandom thinks that we should not presuppose that acts such as adopting an intentional stance need to be explained by explicit mental states as beliefs and desires. Otherwise, to explain those explicit mental states, some other explicit mental states would be needed, and repeat the same process to the infinite. Instead, Brandom suggests that we should understand adopting an intentional stance as merely *implicitly* relying on mental states as beliefs and desires [12].

Neither does Toon think that fictionalism really has the problem. He thinks that fictionalism does not say that psychological sentences do not represent anything. Toon says,

[F]ictionalism can acknowledge the existence of external, public representations with content. [...] [T]hese representations play [the role] within norm-governed social practices. [...] In the case of folk psychology, in pretending to describe inner representations, we are actually saying something about behaviour. Does the notion of pretence itself pre-suppose that of mental states? Does pretending that *p* is true involve imagining *p*, for example, or believing that *p* isn't true? Not necessarily. [...] Instead of focusing on its supposed connections with mental states like imagination or disbelief, the best way to understand pretence is as a public, rule-governed activity found in distinctive social practices, like putting on plays [14].

Sharing Wallace's idea, Toon makes clear that we do not need to understand those mental terms in question as representing some mental states but merely as some public, rule-governed activities.

I submit that, for MAR, something similar should be made in the face of the Question-Begging Problem. To understand the act of rationalizing, we do not have to appeal to or presuppose some mental states. The point is not that the act is not mental or is merely some physical movement. It is that the act can be understood without representing any mental states or facts. The act of rationalizing is also a type of social practice that is governed by some norms. For example, an actor can act with a blank mind, but the act can still be interpreted by audiences as an act of rationalizing.

To support MAR, I do not have the commitment to the existence of those mental states as "rationalizing" or "attributing." When I say that to attribute a mental state is to rationalize some behaviors or that a rationality sentence expresses the mental state of acceptance of some norms, we should not understand those mental states in the robust sense. If the negative thesis (namely, that psychological sentences are not representational) is correct, then we should also understand those mental states in the framework of MAR. The problem is really how we should understand the mental states of rationalizing, accepting, adopting, and attributing. And they are explained by MAR.

References

1. Huang R (2023) Mental Anti-Representationalism: Prospects for a Noncognitivist Account of Psychological Discourse. Dissertation, University of Florida, Gainesville, Florida.
2. Lewis D (1978) Truth in Fiction. *American Philosophical Quarterly* 15(1): 37-46.
3. Gibbard A (2003) *Thinking How to Live*. Harvard University Press, Cambridge, pp: 7.
4. Churchland PM (1981) Eliminative Materialism and the Propositional Attitudes. *The Journal of Philosophy* 78(2): 67-90.
5. Demeter T, Parent T, Toon A (2022) *Mental Fictionalism: Philosophical Explorations*. Routledge, New York, pp: 3-4.
6. Gibbard A (1990) *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press, Cambridge.
7. Davidson D (1991) Three Varieties of Knowledge. In: Davidson D (Ed.), *Subjective, Intersubjective, Objective*. Oxford University Press, Oxford.
8. Dennett DC (1978) *Brainstorms: Philosophical Essays on Mind and Psychology*. MIT Press, Cambridge.
9. Dennett DC (1987) *The Intentional Stance*. MIT Press, Cambridge.
10. Wallace M (2007/2022) Mental Fictionalism. In: Demeter T, et al. (Eds.), *Mental Fictionalism: Philosophical Explorations*. Routledge, New York, pp: 34&45.
11. Toon A (2022) Fictionalism and Intentionality. In: Demeter T, et al. (Eds.), *Mental Fictionalism: Philosophical Explorations*. Routledge, New York, pp: 60.
12. Brandom R (1994) *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press, Cambridge, Massachusetts, pp: 60-61.
13. Wallace M (2007/2022) Mental Fictionalism. In: Demeter T, et al. (Eds.), *Mental Fictionalism: Philosophical Explorations*. Routledge, New York, pp: 45-46.
14. Toon A (2022) Fictionalism and Intentionality. In: Demeter T, et al. (Eds.), *Mental Fictionalism: Philosophical Explorations*. Routledge, New York, pp: 61-62.