Beliefs as Self-Verifying Fictions*

Angela Mendelovici

Penultimate Draft

Abstract

In slogan form, the thesis of this paper is that beliefs are self-verifying fictions: We make them up, but in so doing, they come to exist, and so the fiction of belief is in fact true. This picture of belief emerges from a combination of three independently motivated views: (1) a phenomenal intentionalist picture of intentionality, on which phenomenal consciousness is the basis of intentionality; (2) what I will call a "self-ascriptivist" picture of derived representation, on which non-fundamental representational features are a matter of our ascribing contents to ourselves or our mental states or contents; and (3) a representationalist picture of the attitudes, on which the attitude components of mental states (e.g., the "belief" bit of a belief that P) are represented contents. This paper outlines and motivates the view of beliefs as self-verifying fictions, compares the view to alternative views of belief, and contrasts beliefs on the resulting picture to other belief-like mental states.

Keywords: belief, phenomenal intentionalism, attitude representationalism, fictionalism

*Forthcoming in Eric Schwitzgebel and Jonathan Jong (eds.), *What Is Belief?* Oxford University Press.

1 Introduction

We intuitively think of ourselves as rational epistemic agents with beliefs, desires, and other mental states. These mental states, together with our actions, make us assessable for rationality, correctness, and moral goodness—they're central to our status as thinking, reasoning, and acting agents. For example, you might take yourself to have a desire to write a paper and some beliefs about what it would take to do so. To the extent to which these beliefs are supported by your evidence, you are epistemically rational. To the extent to which you act in a way that is likely to satisfy your desire given your beliefs, you are practically rational. All this depends at least in part on your having the relevant mental states.

Not all of these mental states that we intuitively take ourselves to have are present at all times in our conscious awareness. At best, only a small number are at any given time, and even when we are consciously aware of a mental state, its entire content might not be clearly within view. For example, you might believe that the argument from illusion for the existence of sense data is a terrible argument, but you needn't be consciously aware of this belief in order to count as having it. And even if at some time you are consciously aware of such a belief, you might not be consciously aware of it in its entirety: you might be aware of believing that the argument named "The Argument from Illusion" is a terrible argument without being aware of the full content of the putative belief. In short, we seem to in some sense carry our full system of mental states with us even when we are not consciously aware of it in its entirety.

This lack of full awareness of the totality of our mental states leaves room for skepticism: It is compatible with our experience at a time that we lack many of the beliefs, desires, and other mental states that we intuitively take ourselves to have and on which our status as epistemic agents depends. While it is implausible to deny the existence of mental states of which we're consciously aware, it is an open theoretical possibility that we fail to have any mental states beyond our awareness, that we are subject to a kind of refrigerator light illusion

whereby beliefs and other mental states are only there when we look for them.

This paper is about a specific kind of mental state: belief. While I don't think that we are subject to a refrigerator light illusion in the case of belief, I do think that there is no more in the fridge than what we populate it with ourselves. The view I will propose draws on (1) a phenomenal intentionalist picture of intentionality, on which phenomenal consciousness is the basis of intentionality; (2) a self-ascriptivist picture of derived representation, on which non-fundamental representational features are a matter of our ascribing contents to ourselves or our mental states or contents; and (3) a representationalist picture of the attitudes, on which the attitude components of mental states (e.g., the "belief" bit of a belief that P) are represented contents. On the resulting view, very roughly, beliefs are self-verifying fictions—we make them up, but in so doing, they come to exist, and so the fiction of belief is in fact true.

2 The target: belief

Let us begin by fixing reference on our target: belief. Following what I take to be common philosophical usage, let us say that a subject has a *belief that P* when they have a state that

(1) represents the propositional content P,

(2) involves an acceptance-like attitude towards P, and

(3) is a cognitive state.

According to (1), beliefs are representational by definition. They are "of" or "about" something; they "say" something, which is their *content*. I will clarify what this amounts to shortly, but for now it suffices to say that beliefs are representational in the sense of "representation" that captures our perspective on the world, how things seem us.

Also according to (1), the contents of beliefs are propositional by definition. A *propositional content* is a content that has a propositional form, such as <grass

is green>, which is a matter of saying that something is the case. Propositional contents might turn out to be abstract propositions (e.g., abstract sets of possible worlds), but they might instead turn out to be items of some other ontological category that have propositional forms.[1]

According to (2), beliefs must, by definition, involve an acceptance-like attitude. It is commonly thought that the attitude aspect of mental states is distinct from the content aspect, allowing us to bear different attitudes to the same content (e.g., we can believe that it's snowing or desire that it's snowing) and the same attitude towards different contents (e.g., we can believe that it's snowing or believe that it's raining). An *acceptance-like attitude* is an attitude towards a propositional content that takes that content to be true. Acceptance-like attitudes can be constrasted with other attitude types, such as desire-like attitudes, which do not commit to the truth of a proposition but rather specify a preference, want, or desire. As we will soon see, there are multiple attitude types that count as acceptance-like, any of which suffice for meeting the second definitional criterion of belief.

According to (3), beliefs are by definition cognitive states. *Cognitive states,* like beliefs, desires, and thoughts, can be contrasted with *sensory states*, like perceptual states, pains, bodily sensations, emotions, and moods. Since there might be mental states with a propositional content and an acceptance-like attitude that do not qualify as beliefs (e.g., on some views, perceptual states), our definition of "belief" includes the requirement that beliefs are cognitive states.[2]

A few further clarifications are in order: First, belief states are states of subjects, though subjects might count as having beliefs in virtue of having particular states, events, or other items within them or their subsystems. For

[1]For this ontologically non-committal way understanding of propositional contents, see Mendelovici (2018b).

[2]The term "cognitive state" is sometimes used in opposition to "conative state" to mark a distinction between non-sensory mental states involving an acceptance-like attitude and those involving a desire-like attitude. On the usage employed in this paper, in contrast, the term "cognitive states" picks out any non-sensory state. Thanks to an anonymous reviewer for requesting this clarification.

example, a subject might believe that grass is green partly in virtue of having an internal representational state that represents that grass is green.

Second, the definition of "belief" tells us what it would take for something to qualify as a belief (at least on my usage of "belief"), but it does not give us an account of the deep nature of beliefs. It is the job of a theory of belief to specify the nature of whatever can or does satisfy the definition (if anything).

Third, a distinction is sometimes made between occurrent and standing mental states, where *occurrent* mental states are "happenings"; they are states that are "active" or "undergone" at the time that a subject counts as having them. Thoughts, judgments, and perceptual experiences are examples of occurrent mental states. *Standing* mental states are mental states that a subject can count as having even when they are not "active" or "undergone". Beliefs are sometimes thought to come in both occurrent and standing varieties. For example, if you ask yourself whether you think that the Acropolis is in Athens, you will likely cause yourself to have an occurrent belief with the content <the Acropolis is in Athens>. But five minutes ago you (presumably) had a standing belief, but not an occurrent belief, with that content. (Whether we want to say that the standing belief and the occurrent belief are the *same* belief depends on our identity conditions for beliefs. For our purposes, it will be useful to count occurrent and standing beliefs with the same content as distinct states, but nothing hangs on this choice.)[3]

While we can at least sometimes introspectively notice that we have occurrent beliefs, there are two ways in which beliefs can escape our introspective awareness: First, standing beliefs are never introspectively accessible, even if occurrent beliefs with the same content sometimes are. As a result, the claim that we have standing beliefs is not directly supported by our everyday, introspective access to our

[3]The distinction between occurrent and standing states is related to a distinction that is sometimes drawn between occurrent and "dispositional" states, where occurrent states are understood as I am using the term and dispositional states dispositions to give rise to relevant occurrent states. While it might turn out that the notion of a standing state is co-extensive with that of a dispositional state, it could also turn out that there are standing states that are not dispositional and vice versa. Thanks to an annonymous reviewer for suggesting this clarification.

own mental states. As far as our introspective evidence is concerned, we might be subject to a kind of refrigerator light illusion—a belief that P occurs to us whenever we look for it, but when we're not looking for it, there is no such belief. So, a commitment to standing beliefs will have to be motivated by more than mere introspection and intuition that is ultimately based on introspection.

Second, the precise contents of even occurrent beliefs can escape our introspective capacities. Suppose you have an occurrent belief that the argument from illusion is a terrible argument. Even though you might be introspectively aware of having this belief, its content—what it "says"—might not be fully within view. You might be merely introspectively aware of believing, say, that some argument based on misperception for the existence of sense data is a terrible argument without being aware of its premises and conclusion or of any of its uniquely identifying features. In some cases, what you intuitively take to be your belief's full content might be available to you upon further inspection, but in other cases it might not (e.g., in the case of broad belief contents, which are determined by a subject's environment). Again, as far as our introspective evidence is concerned, we may fail to have the fully fleshed out belief contents we intuitively take ourselves to have and a commitment to such fully fleshed out contents would have to be motivated by more than mere introspection and intuition that is ultimately based on introspection.

3 The proposal

The view I will propose is based on three general views of mental states and attitudes: phenomenal intentionalism, self-ascriptivism about derived representation, and representationalism about attitudes. In this section, I will describe these views and how they contribute to my proposed picture of belief.

3.1 Phenomenal intentionalism

Phenomenal intentionalism is a theory of intentionality. As a first pass, we can understand intentionality as the "ofness" or "aboutness" of mental states. My preferred way of defining "intentionality" more precisely is by ostension: We notice that we have occurrent thoughts, occurrent perceptual states, and other occurrent mental states that we are tempted to describe as being "of", "about", or "directed at" something or as "saying" something. These mental states are *intentional states*, and the feature we notice and are tempted to describe using representational vocabulary is *intentionality.* Although the paradigm cases of intentionality that we use to pick out the phenomenon of interest are introspectively accessible and such that we are tempted to describe them representationally, there is no requirement that other intentional states have this feature. It could turn out that there are unconscious, introspectively unavailable, or apparently non-representational intentional states, so long as they have the same feature that we notice in paradigm cases.[4]

According to *phenomenal intentionalism*, the having of at least the most basic kind of intentional state is nothing over and above the having of phenomenal states, where *phenomenal states* are states of there being something it is like to be something (Nagel 1974).[5] In other words, phenomenal intentionalism claims that at least the most basic kind of intentionality is *phenomenal intentionality*, intentionality that is nothing over and above phenomenal consciousness. For example, a perceptual experience might represent the content <blue triangle>, perhaps as part of a richer content, simply by having a bluish-trianglish phenomental character. In this case, <blue triangle> is (or is part of) the perceptual experience's *phenomenal content.*

Why accept phenomenal intentionalism? One kind of argument for the

[4] For a defense of this way of defining "intentionality", see Mendelovici (2018a, Chapter 1).

[5] Proponents of phenomenal intentionalism include Strawson (1994), Horgan and Tienson (2002), Loar (2003), Pitt (2004), Farkas (2008), Kriegel (2003), Kriegel (2011), Bourget (2010), Mendelovici and Bourget (2014), and Mendelovici (2018a). For overviews, see Kriegel (2013), Bourget and Mendelovici (2016), and Mendelovici and Bourget (2020).

view aims to establish that phenomenal consciousness (or at least phenomenal consciousness of the right kind) is metaphysically sufficient for intentionality. For example, it seems that simply in virtue of having a bluish phenomenal experience one is already representing <blueness>. It seems inconceivable that one could have a bluish phenomenal experience without thereby already representing <blueness>.[6] Thought experiments involving disembodied souls or brain in vats, which might have phenomenal experiences despite lacking any interesting causal connection to things in the external world, evoke similar intuitions: there is some way that disembodied souls and brains in vats represent the world simply in virtue of their phenomenal experiences.[7]

Another kind of argument for phenomenal intentionalism aims to establish the view's ability to make correct predictions about particular cases. For example, phenomenal intentionalism can handle cases of intentional states that represent something other than the features of the world that reliably cause them (e.g., color experiences, which are reliably caused by physical surface reflectance properties of objects but seem to represent sui generis, primitive colors).[8] And phenomenal intentionalism ascribes contents determinately in cases where other views inappropriately ascribe contents indeterminately.[9]

Phenomenal intentionalism faces challenges in accounting for putative intentional states that seem to go beyond a subject's phenomenal states. For example, unconscious states, like standing states and (presumably) the subpersonal states posited by cognitive science, have no accompanying phenomenal characters to capture the contents we want to ascribe them. Many occurrent thoughts also seem to have contents that cannot be fully captured by any accommpanying phenomenal characters. For example, any broad contents we might represent arguably go beyond any phenomenal contents a subject can plausibly be said to represent. For instance, a pair of intrinsic duplicates embedded in differ-

[6] See Mendelovici (2018a, §5.2.2); Horgan and Tienson (2002); and Siewert (1998).
[7] See Horgan et al. (2004); Loar (2003); and Kriegel (2011).
[8] See Mendelovici (2018a, §5.2.1).
[9] See Graham et al. (2007); Horgan and Graham (2012); and Searle (1990).

ent environments might represent different broad contents (say, <H_2O> and <XYZ>[10]) but have the same phenomenal states. If so, not all of their contents are phenomenally-determined. But even narrow (i.e., not broad) contents can fail to be captured by a subject's phenomenal states. For example, a subject thinking that the mental supervenes on the physical might fail to have a phenomenal state capturing their precise understanding of supervenience or other nuances of their thought's putative content.

Many phenomenal intentionalists aim to account for many of the alleged contents that go beyond a subject's phenomenal states by saying that they're derivatively represented, where *derived representation* is representation that is ultimately at least partly determined by *original intentionality*, the most fundamental kind of intentionality.[11] Phenomenal intentionalism is committed to the claim that all original intentionality is phenomenal intentionality, but it can allow that there are cases of derived representation so long as they ultimately derive from phenomenal intentionality in some way. In the next subsection, I will outline my preferred picture of derived representation.[12]

How does phenomenal intentionalism apply to the case of belief? At least some occurrent beliefs are phenomenally conscious: there is something it is like to have them. But their phenomenal features might not fully capture the contents we intuitively take them to have. As we saw earlier, much of the content of occurrent belief lies beyond our immediate introspective awareness, and, presumably, is not phenomenally conscious. For example, an occurrent belief that the mental supervenes on the physical might, perhaps, involve a phenomenal experience of

[10]In versions of Putnam's (1975) well-known Twin Earth thought experiment, we are encouraged to accept that the Earthling Oscar represents <H_2O> when he says "Water is wet", while his Twin Earthling doppelganger represents <XYZ>.

[11]It is an open question whether derived representation is a species of intentionality or, instead, a distinct kind of phenomenon. In order to avoid appearing to prejudge this issue, I call it "derived representation" rather than "derived intentionality".

[12]Kriegel (2011) claims that, roughly, a subject's derived contents are those that would be attributed by the originally intentional states of an ideal observer applying Dennettian (1987) intentional systems theory to the subject. Pautz (2021) offers a similar view on which we derivatively represent the contents we would be interpreted as having on something close to a Lewisian (1974) interpretationism. Bourget (2010) offers a view of derived representation invoking multiple derivation mechanisms: compositionality, functional connections, descriptive reference, and deference.

the word "supervenes", some phenomenally conscious mental imagery of one thing above another, a phenomenally conscious but gisty overall grasp of some notion of dependence, or some other such vague, gisty, schematic, or partial content. While such a content might be phenomenally represented by the belief, it is impoverished compared to the full-fledged content we intuitively want to ascribe to the belief, which presumably includes a specification of the subject's understanding of supervenience. This richer, full-fledged content will have to be derivatively represented if it is represented at all.

Standing beliefs also presumably go beyond what we phenomenally represent; they have no phenomenal contents, since there is nothing it is like to have them. The phenomenal intentionalist will have to say that all the contents of standing beliefs are derivatively represented. Let us turn, then, to the second part of the overall view, the picture of derived representation.

3.2 Self-ascriptivism

Elsewhere, I've argued for a self-ascriptivist picture of derived mental representation, on which derived mental representation is a matter of ascribing contents to our own mental items or even to ourselves (Mendelovici 2018a, Chs. 7–8; 2019; 2020; forthcoming). This picture allows us to say that beliefs have derived contents that go beyond their phenomenal contents, allowing us to capture the belief contents that we pre-theoretically take ourselves to represent.

According to *self-ascriptivism*, a subject S's mental item M (e.g., S's phenomenal content, internal state, or S themselves) derivatively represents a content C (for S) just in case and because S ascribes C to M—in short, just in case and because S *self-ascribes* C. Ascribing a content is a matter of our assigning a content to something, accepting one thing as meaning or standing for another, or otherwise taking one thing to stand for or mean something else.

Why should we think that self-ascriptions are relevant to derived mental representation? The main motivation for the view comes from consideration of

how derived representation works in non-mental cases. A stop sign represents <stop!> because we assign it this content; the word "fox" represents foxes because we (perhaps implicitly) accept that it does; stipulating "Let the salt shaker stand for the getaway car" makes the salt shaker represent the getaway car in virtue of our stipulation. In these and other non-mental cases of derived representation, ascribing a content to something *makes* that thing have that content for the ascriber (though not necessarily for anyone else). In the same way, ascribing a content to one's *mental* item makes that mental item represent that content for the ascriber. In short, ascriptions are metaphysically sufficient for derived representation, and, arguably, nothing else is. So, if there is such thing as derived mental representation, it must be a matter of our ascribing contents to ourselves or our mental items—in short, it must be a matter of our self-ascriptions.

That is what it would take for us to derivatively mentally represent. I want to suggest that we do in fact satisfy the conditions for derived mental representation: we self-ascribe contents. There are various ways in which we do so: One is by being disposed to take our phenomenally-represented contents to mean, stand for, be placeholders for, or cash out into further contents. For example, take the occurrent belief that the mental supervenes on the physical. Suppose that in having this belief, corresponding to your concept of supervenience, you entertain an impoverished phenomenal content, perhaps consisting of some mental and verbal imagery and a gisty grasp of a relation of dependence. This is the phenomenal content of your concept SUPERVENIENCE. Although in having the occurrent belief you are only entertaining this impoverished content, you in some sense take yourself to mean something more by it. If you ask yourself to clarify what you mean by your impoverished content <supervenience>, you might produce a further elucidation <by <X supervenes on Y> I mean <the relation between X and Y such that there there can be no difference in X without a difference in Y>>. Call such thoughts *cashing out thoughts*, since they specify that one content cashes out into another. The content that <supervenience>

cashes out into might itself cash out into further contents, which might themselves cash out into further contents, eventually yielding a fully cashed out content that captures your best understanding of <supervenience>. In virtue of these dispositions to have cashing out thoughts, you count as ascribing this fully cashed out content to your phenomenal content <supervenience>. And so, your phenomenal content <supervenience> counts as derivatively representing this fully cashed out content. Importantly, it is sufficient for self-ascription and, hence, derived representation that we have the relevant dispositions—we needn't in fact have the relevant cashing out thoughts. Similarly, in order for a stop sign to represent <stop!> we needn't be occurrently thinking that it does at any given moment; it is enough that we tacitly or implicitly accept that it does, which might amount to nothing more than a disposition to accept that it does.

On this picture, thought is symbolic: In thinking, we are usually only using our impoverished phenomenal contents, contents that fall far short of the fully cashed out contents that we in some sense target. But these phenomenal contents derivatively represent these further contents and we can, at least sometimes, retrieve them if needed. Our phenomenal contents, then, serve as "stand-ins" for our fully cashed out contents. This is more efficient than entertaining our fully cashed out contents, and it has the added benefit of allowing our phenomenal contents to form their own inferential connections independent of those of their corresponding derived contents.[13]

We can distinguish two types of ascriptions and, correspondingly, two types of derived representation. *Direct* ascriptions are ascriptions that ascribe a content by presenting the content itself within the ascription. The example of <supervenience> given above is an example of direct ascription, since the relevant cashing out thought ascribes the content <the relation between X and Y such that there can be no difference in X without a difference in Y> to <X supervenes on Y>. Direct self-ascriptions result in a kind of mental shorthand,

[13]This picture of thought as involving impoverished "stand-ins" for richer contents has been developed in different ways by Eliasmith (2013); Barsalou (1993, 1999); Prinz (2002); Wickelgren (1992).

though they do not augment our representational powers. While, in principle, we are able to phenomenally represent the ascribed content, we introduce a new phenomenal content to stand for it. This might be convenient, but it does not allow us to represent contents that we cannot antecedently phenomenally represent.

Unlike direct ascriptions, indirect ascriptions can augment our representational powers, allowing us to represent more than we antecedently could. An *indirect* ascription ascribes a content by referring to it without presenting it. Suppose that you are disposed to have the following cashing out thought: <by <X supervenes on Y> I mean the relation between X and Y such that there can be no difference in X without a difference in Y>. In this case, you are ascribing the relation of supervenience *itself*—not just a description or characterization of it—to your phenomenal content <X supervenes on Y>. Of course, in order for you to be able to do this, you need to be able to refer to this relation using your phenomenal contents.[14] But, granting that this is possible, we can then ascribe to our phenomenal contents items that are not themselves phenomenal contents, like the relation of supervenience itself. This allows us to derivatively represent singular contents, which involve particular worldly items. If intrinsic duplicates can refer to different items with internal states that are intrinsically alike, this also allows us to derivatively represent broad contents.

So far, we have a sketch of an account of the contents of occurrent beliefs: Occurrent beliefs have original contents, which are phenomenally represented. These contents are impoverished, but we take them to cash out into further contents. Our concepts' fully cashed out contents capture our best understanding of our targets. When the relevant self-ascriptions are direct, we end up with occurrent beliefs that derivatively represent rich and complex phenomenal contents that it is possible for us to phenomenally represent (at least in principle, though perhaps not in practice due to their complexity). When the relevant

[14]Accounting for reference is not trivial for phenomenal intentionalism. For discussion, see Mendelovici 2018a, Ch. 9 and Appendix H; 2023; forthcoming; Ott2016-OTTPIA-3,BourgetForthcoming-BOURVA; Bourget and Mendelovici forthcoming.

self-ascriptions are indirect, we end up with occurrent beliefs that derivatively represent the referents of possible phenomenal contents; in this case, we derivatively represent contents that we might not be able to phenomenally represent, such as broad contents and singular contents involving particular worldly objects. Some beliefs might contain concepts exhibiting both direct and indirect derived representation, resulting in mixed derived contents composed of both possible phenomenal contents and the worldly objects, kinds, or other referents of phenomenal concepts.

The contents of standing beliefs, which have no phenomenal features, can also be understood in terms of derived mental representation. In the case of standing beliefs, we ascribe contents to ourselves as a whole rather than to some phenomenal content we entertain. We might compare this to the way in which we might assign a theme, moral, or overall message to a book, story, or a work of art: The content is assigned to the work as a whole—*globally*, we might say—rather than to any of its words, sentences, or parts. And the thing to which the content is assigned does not represent the content in the sense of being used in place of or "standing for" the content but instead has the content as part of its overall significance or meaning. In the same way, one might ascribe a content to oneself as a whole. We have a general sense that what is in our mind goes beyond what is in our consciousness at any given time. When needed, contents that we were not previously aware of can come to mind and we accept them as our own. These contents that we are disposed to accept as our own upon entertaining count as being self-ascribed. For example, you might be disposed to think that the Eiffel Tower is in Paris. You have the disposition to accept this content as your own upon entertaining it, so you thereby count as self-ascribing and hence derivatively representing this content.

When the standing state contents that we are disposed to self-ascribe are themselves such that we ascribe further contents to them, perhaps by being disposed to take them to cash out into further contents, the standing state

contents inherit these further contents. Let's go back to the example in which you are disposed to accept yourself as having the content <the Eiffel Tower is in Paris>, which might be a possible impoverished phenomenal content. This content, or its constituent contents, might be such that you are disposed to take them to cash out into further contents. For example, the fully cashed out content corresponding to <Eiffel Tower> might be a rich descriptive characterization that refers to the Eiffel Tower, or it might be the referent of such a characterization—the Eiffel Tower itself—depending on whether this fully cashed out content is directly or indirectly ascribed. The end result is that our standing states can be ascribed rich possible phenomenal contents or the referents of such contents.

It is noteworthy that on this overall picture of derived mental representation, it can be indeterminate which contents we self-ascribe, and so it can be indeterminate which contents we derivatively represent. This is because we can derivatively represent a content by having the right dispositions to take ourselves or our phenomenal contents to have it or cash out into it upon sufficient reflection, but we have not precisely specified what the relevant conditions of manifestation are for such dispositions because we have not specified precisely what counts as sufficient reflection. This is a feature, not a bug: Except in the case of phenomenally represented contents, it often *is* indeterminate precisely which contents we are targeting. In some circumstances, you might cash out the phenomenal content <bachelor> as <unmarried man>, while in others you might cash it out as <man available for marriage>. There might be no fact of the matter as to which best captures your understanding or which you are in some sense targeting when you think <John is a bachelor>. If so, then your phenomenal content <bachelor> has a derived content that is indeterminate between the two options. This seems to be the right answer. In the same way, the contents of standing states can be indeterminate.[15]

[15]This indeterminacy is not problematic. We do not have strong evidence for the relevant contents being fully determinate; we only have strong evidence for the occurrent contents that we're immediately aware of being determinate, and these are our phenomenal contents, which *are* fully determinate. Also, no theoretical weight is being put on the notion of derived contents. Derived contents are not supposed to ground any fully determinate facts or play other roles

One might worry that the picture of derived representation makes representing a non-phenomenal content too demanding in some respects and too easy in others. In order to represent a non-phenomenal content, we must be able to either phenomenally represent it in a self-ascription (perhaps in a piecemeal fashion) or to phenomenally represent a content that picks it out. This seems too demanding—we need to be capable of specifying the targeted content—and also too easy—all we have to do to represent a content is to ascribe it to ourselves; we need not meet any further conditions. While I agree that derivatively representing a content is both demanding and easy in the relevant respects, I think this is the right level of demandingness and ease. We can also see why the level of ease and demandingness is appropriate by considering again the nature of derived representation: once we appreciate that derived representation requires ascriptions, we can appreciate that once the ascriptive work is done that is all that's needed for derived representation. This is just how derived representation works.

We can also independently motivate the claim that the account makes the representation of non-phenomenal contents appropriately easy and demanding: If we could represent non-phenomenal contents without being in some way able to specify them, we could be epistemically blameworthy or praiseworthy for beliefs and other mental states with contents that we in no sense recognize as our own. We could end up having false or inconsistent beliefs that are utterly foreign to us and that we might even reject or deny having. We could end up having true beliefs through no effort on our part that rationalize our actions even though we in no sense endorse those beliefs. In short, the picture of derived representation intuitively gives us the right amount of control of the epistemically-relevant portion of our minds: it allows us to represent all and only the contents we either phenomenally represent or are able to specify and accept as our own. In this way, the account preserves a picture of the epistemically-relevant portion of our

that would require them to be determinate. They merely serve to capture the contents we take ourselves to have that go beyond our phenomenal contents and that are relevant to assessing ourselves as rational, thinking beings.

minds on which our minds are autonomous and in control of their contents.[16]

Importantly, the claim is not that we succeed at derivatively representing precise contents in all cases. Instead, the claim is that insofar as we manage to represent contents that go beyond our conscious minds, we do so by self-ascribing them. These are the contents that we carry with us in our personal minds, the contents that make us subjects of epistemic and practical rationality. I also believe that we do in fact succeed in derivatively representing many contents, but which specific contents we derivatively represent is a further empirical question.

I've sketched a picture on which we have occurrent states with impoverished phenomenal contents and rich derived contents, and standing states with rich derived contents. When these states have propositional contents, they satisfy condition (1) of our definition of "belief". In order to qualify as a belief, any such state must also involve an acceptance-like attitude and be cognitive. The next subsections will sketch a representational picture of attitudes and the cognition-sensation distinction, which will support the claim that there exist mental states that satisfy not only the first but also the second and third conditions.

3.3 Representationalism about attitudes

So far, I've suggested that a combination of phenomenal intentionalism and self-ascriptivism can allow us to capture belief contents. But in order for a mental state of representating a content to qualify as a belief—rather than a desire, intention, fear, wish, or other kind of mental state—it must come with an acceptance-like attitude. It is commonly thought that the attitude aspect of mental states is distinct from the content aspect, allowing us to bear different attitudes to the same content and the same attitude towards different contents. This makes sense of how two subjects can believe different things and one subject can desire what another believes.

[16] This does not preclude, of course, the possibility of other aspects of our minds that are not under our control in the relevant ways. See Section 4.1. I will argue, however, that these aspects of our minds are not representational.

There are different views of attitudes. They might be a matter of the functional roles of representational states, associated phenomenal states, or different primitive relations that subjects bear to contents. The view I want to recommend is *attitude representationalism*, which takes attitudes to be a matter of represented contents. More specifically, attitudes are represented properties that are ascribed to what we intuitively take to be the contents of our mental states. For example, a belief that P might be represent the content <P is a fact>, <P is the case>, or <P is true>, while a desire that P might represent <P is good>, <P is desirable>, or <P is rewarding>.[17,18]

Representationalism about attitudes faces many challenges, many of which are discussed elsewhere, but it also has many theoretical and explanatory virtues. For instance, as Pearce (2016) argues, a representationalist picture of attitudes is best placed to account for the differing roles of belief versus desire in determining the rational features of practical inferences. And as Montague (2022) argues, representationalism has the resources needed to distinguish between different types of representational states, making additional ingredients unnecessary for distinguishing different intentional states.

It is a further question what is the "default" attitude—which attitude does a case of merely representing a propositional content have? One view is that the default attitude is mere entertaining. On this view, representing the content <P> amounts to entertaining that P, which involves no commitment to the truth of P. A problem with this view is that if the default attitude is entertaining, it's unclear how thinking <P is true> can amount to anything more than merely entertaining that P is true, which would involve no commitment to P's truth. In order for attitude representationalism to succeed, it looks like the default attitude will have to be some kind of attitude of acceptance, some kind of taking as true. When we represent <P>, we by default accept that P is true. On this

[17] Representationalist views of the attitudes have been defended by Pearce (2016), Montague (2022), and Mendelovici (2018a, Appendix E; 2020).

[18] Our concern here is with the propositional attitude of belief, but the view can easily be extended to cover other attitudes, including non-propositional attitudes.

view, then, entertaining will have to involve a cancellation of the commitment to truth. For example, perhaps entertaining that P amounts to representing the content <P, which may not be true>.[19]

There is also a question about how the attitude-specific contents are represented. On the combination of phenomenal intentionalism and self-ascriptivism, they might be phenomenally represented or derivatively represented. Whether there are attitude-determining contents that are phenomenally represented depends on the phenomenal features of mental states with various attitudes, while whether there are attitude-determining contents that are derivatively represented depends on whether we self-ascribe any such attitude-determining contents.

One possible view is that there is a unique phenomenal content corresponding to each attitude. For example, desires might have a unique phenomenal character that determines a particular phenomenal content, like <good> or <rewarding>.[20] Alternatively, there might be multiple distinct phenomenal characters and corresponding phenomenal contents that constitute desire. If, instead, attitude-determining contents are merely derivatively represented, an occurrent state would merely involve the entertaining of a non-attitude-determining content, but upon reflection, we might ascribe to ourselves an attitude-determining content towards that content. The ascribed attitude-determining content might be directly or indirectly derivatively represented.

My favored view is that representational states have both phenomenal contents that characterize the general attitude type and derived contents that capture further nuances in attitude type. For example, a desire might have a phenomenal content like <P is good>, and we might be disposed to take <good> to cash out into <good for me>, <good in itself>, <pleasant>, or some other further content. Which further content we are disposed to take <good> to cash out into might depend on the circumstances, allowing states with the same phenomenal contents to have different nuanced derived contents. Note

[19] Thanks to Adam Pautz for pressing this question.

[20] See Pearce (2016) for a very interesting account of the content of desires on which they represent their object as rewarding.

that even if acceptance is the default attitude, both phenomenal and derived contents can provide further nuance—we can accept with certainty, tentatively accept, predict, and so on, and this might be a matter of either our phenomenal or our derived contents. When the attitude-determining content is at least partly derived, there can be indeterminacy in the precise attitude we take towards a proposition, and even in whether the attitude qualifies as a belief or not. For instance, it can be indeterminate whether we believe, anticipate, or fear a future outcome.[21]

On the resulting overall view, then, our attitudes are largely self-ascribed, just like many of the rest of our contents. If acceptance is the default attitude, then belief need not involve further attitude-determining contents, though it might.[22]

3.4 Cognition and sensation

As stated earlier, *cognitive states,* like beliefs, desires, and thoughts, can be contrasted with *sensory states*, like perceptual states, pains, bodily sensations, emotions, and moods. Beliefs, by definition, are cognitive states.

There are different possible views of what it takes for a state to be a cognitive state. I want to again suggest a representationalist picture, on which differences between phenomenal and derived contents distinguish between cognitive and sensory states: Sensory states have rich phenomenal characters with correspond-

[21]Cashing out thoughts and other mental states in virtue of which we count as self-ascribing contents also involve attitudes, which might be stipulative, meaning-conferring, or meaning-accepting attitudes. To avoid regress, the attitude-determining content in virtue of which ascriptive mental states qualify as ascriptive mental states would have to either be the default attitude or phenomenally represented—it should not be derivatively represented (though further attitude-determining contents that are not required in order for an ascriptive mental state to count as an ascriptive mental state might be derivatively represented). It is plausible that ascriptive mental states do have such non-derivatively-represented attitude contents, since we are consciously aware of our taking one thing to stand for another or taking ourselves to have a belief or other mental state.

[22]The suggestion is not that differences in phenomenal and derived content are the *only* differences between mental states with different attitudes—perhaps attitudes have characteristic functional roles or other features that distinguish them. The suggestion, rather, is that the phenomenon that we introspectively and intuitively appreciate of taking different attitudes towards contents, where these different attitudes play a role in our being rationally assessable beings, is best explained by differences in content.

ingly rich phenomenal contents. For example, your perceptual experience as of a unique (i.e., pure) red wall at arm's length away has a rich phenomenal character corresponding to the represented redness, which grounds (or perhaps is identical to) your phenomenal content of redness. In contrast, your belief that there is a unique red wall at arm's length away lacks this rich phenomenal character. Though your belief may represent the same content, including the same content of unique redness, its representation of this content is captured by its derived content: perhaps you have an impoverished phenomenal content corresponding to some verbal imagery, but you represent the fuller content thanks to your dispositions to have cashing out thoughts involving this content (perhaps you are disposed to take the impoverished redness-related phenomenal content to cash out into the way things appear to be when you have experiences of what you call "pure red"[23]). In short, a key difference between cognitive and sensory states is that sensory states are richer in phenomenal character and hence richer in phenomenal content. Cognitive states might represent some of the same contents as sensory states, but these contents are largely derived. Since richness of phenomenal contents is a matter of degree, there may be no non-arbitrary dividing line between cognitive and sensory states.[24,25]

3.5 Summary

This completes the picture of belief. At bottom, having a belief is a matter of representing particular contents. To believe that P, one must represent a content that is or includes as a part the propositional content P, which must be entirely or largely derivatively represented, and one must have an acceptance-like attitude

[23]See also Mendelovici (2018a, p. 148) for an account of how we can account for dispositional concepts in terms of derived content.

[24]I am ignoring the (theoretical) possibility that there are sensory states that do not have phenomenal contents. I don't think there are such states (see Mendelovici 2018a, Appendix B; 2013b; 2013a), but if there are, we might say that they qualify as sensory states in virtue of their rich phenomenal characters or their failure to have representational contents.

[25]As in the case of the attitudes (see fn. 22), there might be other characteristics of sensory and cognitive states that distinguish them from one another. My claim is not that there are no other such characteristics but rather that the phenomenon that we can introspectively and intuitively appreciate of there being a difference between sensory and cognitive states is best explained by differences in content.

towards P, which might be a matter of merely representing that P or a matter of representing that P is the case, true, accepted, etc. In the case of standing beliefs, which have no phenomenal contents, all the contents characterizing a belief are derived. In the case of occurrent beliefs, some belief contents are phenomenal, but most are derived.

Since beliefs are fully or largely a matter of derived mental representation and derived mental representation is a matter of self-ascription, beliefs are largely self-ascribed: we have them because we ascribe them to ourselves. This makes beliefs somewhat of a fiction: We make them up. When we intuitively take ourselves to have a belief, we manifest the very dispositions that make it the case that we have the belief. Had we not had the relevant dispositions, we would not have had the belief. But the fiction is self-verifying—it makes itself true. In self-ascribing beliefs, we thereby count as having them, for real.

4 Arguments and comparison with alternative views

Why should we accept this view? The main line of argument for the overall view is a combination of the arguments for phenomenal intentionalism, self-ascriptivism, and representationalism about attitudes and the sensation/cognition distinction, which I've briefly sketched or pointed to above. If we accept these views, there is nothing else that beliefs could be. For instance, if we accept the arguments for phenomenal intentionalism and self-ascriptivism, then the only way that we could represent a belief's content is by either phenomenally representing it or self-ascribing it. Nothing else would be metaphysically sufficient for representing a belief's content.

Rather than rehash the arguments for these views,[26] I will consider where the proposed view parts ways with three main alternative views of belief—the

[26] I have argued for them elsewhere, especially Mendelovici 2018a, 2020; Mendelovici and Bourget 2020; Mendelovici 2019.

internal state view, dispositionalism, and interpretationism—and why I prefer the proposed view.

According to the *internal state view*, beliefs are internal states (which might be physically or functionally individuated) that represent a particular propositional content and involve an attitude of belief, which might be a matter of their functional role. A well-known example of this kind of view is Jerry Fodor's (1975; 1987) language of thought picture, on which, roughly, a subject has a belief if they have a syntactically individuated internal state that represents a propositional content and plays a particular functional role (which can be glossed as having the internal state in their "belief box").[27]

My proposed picture is compatible with some aspects of the internal state view. It's compatible with the claim that we have distinct internal states corresponding to many or all of our standing and occurrent beliefs. It's also compatible with the claim that these internal states play certain characteristic functional roles that roughly correspond to the roles we take beliefs, desires, and other attitudes to play. These are claims about cognitive architecture and processing.

But my picture does not take having such internal states to be either necessary or sufficient for belief—it's compatible with our having beliefs that have no corresponding internal states, and it's compatible with our having such internal states that don't figure into our beliefs (or desires or other personal-level mental states) but that are merely subpersonal or non-representational states.[28] This seems to be the right answer, since we are inclined to attribute beliefs to ourselves

[27]See also Millikan (1984); Dretske (1988); Quilty-Dunn and Mandelbaum (2018). This view is sometimes called "representationalism", since it takes beliefs to be inner representations (see Schwitzgebel 2006).

[28]Derived mental representation can be realized in many ways, including in ways that only require us to have particular dispositions. Even if the relevant dispositions are sometimes grounded in our having various internal states, we could have them without the relevant states. For example, perhaps many of the contents ascribed by our cashing out thoughts are generated on the fly. Indeed, it seems unlikely that we have distinct cognitive structures corresponding to all our fully cashed out contents, partly because our fully cashed out contents, which in many cases will specify necessary and sufficient conditions for concept application, presumably do not play much of a psychological role in categorization, inference, and behavior (Smith and Medin 1981; Murphy 2004). See Mendelovici (2018a, §7.2.3; 2019, fn. 5) for discussion.

and subjects irrespective of internal machinery (see Dennett 1987; Schwitzgebel 2013). The main reason for preferring my proposal, however, is that beliefs are contentful states and (if the arguments for phenomenal intentionalism and self-ascriptivism are correct) the only way to represent a content is by phenomenally representing it or self-ascribing it. Thus, whatever picture the internal state view provides, if it does not result in phenomenal contents or self-ascriptions, it is not a picture of a representational state, and so it is not a picture of belief.[29]

Another view of belief is *dispositionalism*, the view that what it is to have a belief that P is to have a certain set of dispositions characteristic of having the belief that P (Braithwaite 1932; Marcus 1990; Schwitzgebel 2002). On this view, beliefs are what are sometimes called "traits"—they are psychological dispositions or sets of dispositions. There are different views of what the relevant dispositions might be. For Schwitzgebel (2002), who offers what is arguably the most promising version of this account, the relevant dispositions include behavioral dispositions, including dispositions to make certain utterances; dispositions to make inferences or enter certain other cognitive states; and dispositions to have certain phenomenal experiences, such as experiences of surprise when confronted with conflicting evidence. To the extent to which a subject has the dispositions characteristic of believing that P, the subject believes that P.

A virtue of this view, which it shares with my proposal, is that it allows for "in-between" cases of belief, cases in which we have some but not all of the characteristic features of having a particular belief (Schwitzgebel 2001). A crucial difference between dispositionalism and my proposal is that dispositionalism allows that we have beliefs that we in no sense recognize in ourselves, even beliefs that we deny having. Insofar as we have reason to think that we can have such beliefs, even if they are mere in-between cases, we should favor dispositionalism.

I am not moved by this consideration for this reason: The conditions that the dispositionalist takes to be metaphysically sufficient for belief are not in

[29]See also Mendelovici (2018a, chapter 8) for a different line of argument against taking the internal states of the internal state view to be representational in some other way.

fact metaphysically sufficient for representing a content. A subject might have the dispositions that the dispositionalist takes to be sufficient for belief but fail to represent a content because they neither phenomenally represent nor self-ascribe the relevant derived contents. So, while the dispositional account might offer a characterization or improvement of one everyday concept of belief (see §4.1), it does not provide us with an account of belief as a representational state. (This line of argument, of course, depends crucially on the arguments for taking ascriptions to be necessary for representing a content that is beyond our consciousness, which were gestured to in §3.2.)

Another view of belief is *interpretivism*, on which having a belief is a matter of being interpretable as having the belief given certain constraints on interpretation (Lewis 1974; Dennett 1987, 1991; Davidson 1984).[30] Typically, the idea is that the core of our folk psychological theory of minds and behavior posits beliefs, desires, and other mental states that interact with one another and generate behavior in accordance with the dictates of rationality. A subject's beliefs and desires are those that make best sense of their behavior on the assumption that they are rational. Like dispositionalism, interpretivism allows for "in-between" cases of belief and doesn't require a particular internal machinery in order for a subject to qualify as having beliefs. Unlike dispositionalism, interpretivism holds that what makes a subject have a particular belief is the fact that they are fruitfully interpreted in a particular way, not the fact that they have particular dispositions.

Like interpretivism, the proposal involving self-ascriptivism takes having a belief to be largely a matter of interpretation. However, for interpretivism, either the interpreter is a hypothetical ideal agent (as on Kriegel 2011) or what matters for having a propositional attitude is whether a subject is interpret*able* from the perspective of a particular theory (as on Dennett 1987, 1991). In contrast, on the proposal involving self-ascriptivism, the interpreter is the subject herself.

[30]Some theorists offer interpretivist pictures of derived representation but take original representation to be a matter of phenomenal intentionality or a kind of tracking relation (Kriegel 2011; Pautz 2013, 2021; Williams 2019).

This points to an important consideration in favor of the proposal invoking self-ascriptivism over nearby interpretivist views: As suggested above, ascribing content C to X makes X derivatively represent C—but only *to the ascriber*. For example, stipulating that my pen is to stand for the person with the longest hair on Earth at the time of my writing this sentence makes my pen derivatively represent that person to me (and to any audience that goes along with my stipulation), but it does not make my pen itself, other people, cats, butterflies, or anything else derivatively represent that person by my pen. So, the fact that something other than the subject of a belief themselves—e.g., a hypothetical ideal agent or a theory as applied to a subject—would ascribe a content to that subject is not metaphysically sufficient for *the subject* to have a content. In short, like the inner state view and dispositionalism, while interpretivism might provide us with one everyday or theoretically interesting belief-like notion (see §4.1), it fails to describe metaphysically sufficient conditions for belief as a representational state.[31,32]

4.1 States, traits, and representational attitudes

Even though the internal state view, dispositionalism, and interpretivism are not true of beliefs in the sense specified in Section 2, there may indeed exist the relevant internal states, dispositions, and interpretational posits. These items would not qualify as beliefs in my sense because they are not representational—they do not have phenomenal or derived contents—but they might nonetheless be interesting states or posits and they might even satisfy alternative definitions of "belief" on which "beliefs" need not be representational states.

For example, perhaps there is an everyday notion of "belief" that picks out a trait (a psychological disposition or set of dispositions) that need not be or involve representational states. Examples of traits include kindness, fearfulness,

[31]See also Mendelovici (2018a, chapter 8) for related arguments against interpretivist pictures of derived representation.

[32]Another advantage of the proposal involving self-ascriptivism over interpretivism is that my proposal can easily allow subjects to be irrational.

gullibility, and rudeness. In some cases, having a trait involves having or being disposed to have representational states or other mental states (e.g., gullibility), but in other cases it need not (e.g., rudeness). The suggestion here is there may be an interesting way of defining "belief" on which it picks out a particular kind of trait, which we might call a *trait-belief*. Having a trait-belief might amount to having the dispositions Schwitzgebel takes to be constitutive of belief. One important difference between a trait-belief and a belief, though, is that beliefs must be representational, while trait-beliefs need not be.

Distinguishing between beliefs and trait-beliefs allows us to make sense of cases where we intuitively want to say that a subject lacks a representation of P but behaves in a P-directed way or in a way that would make sense if P were true. This might be the best way to handle cases of implicit bias, in which subjects act and think in accordance with prejudices and stereotypes that they may not explicitly endorse. For example, a subject with an implicit bias against people belonging to group X might have a tendency to overlook people belonging to group X in job searches and judge them to be less worthy in various respects than otherwise similar individuals not belonging to group X, despite explicitly avowing egalitarian attitudes. Such an implicit bias might fail to meet the criteria for being a belief in some relevant propositions but might nonetheless qualify as a trait-belief in those propositions. This picture is similar to Schwitzgebel's (2001) account of implicit biases as cases of "in-between" beliefs, which is based on his dispositionalist picture of belief. The key difference between this suggestion and Schwitzgebel's is that this suggestion does not require that implicit biases be representational: they need not involve representing a content. The relevant dispositions might be mere dispositions to behavior, further thought, emotion, etc., without any of these dispositions constituting the representing of a content. Trait-beliefs might, nonetheless, be fruitfully described by citing a content that the subject acts as if is true, e.g., that individuals belonging to group X have various negative features.[33]

[33]Such an account of implicit bias would in no sense absolve subjects of moral or epistemic

In the same way, we can accept a useful sense of "belief" that picks out a functional or physical internal state, similar to the kind of states the internal states view identifies with beliefs, but that need not have representational features. Like trait-beliefs, such *internal-state-beliefs* might exist and play interesting explanatory roles in the internal processing underlying cognition and behavior even if they do not, strictly speaking, have representational contents.[34] And we can similarly accept a sense of "belief" ascribing the folk psychological states we are interpretable as having that the interpretationist wants to identify with beliefs. Again, these *folk-psychological-beliefs* might fail to be representational. What it is to have such states might be nothing more than to be interpretable in the right way as having them, irrespective of any accompanying representation of a content. Still, ascribing these folk-psychological-beliefs might be useful for certain purposes, like predicting subjects' behaviors at a high level of description in everyday contexts.

While beliefs are largely self-ascribed, trait-beliefs, internal-state-beliefs, and folk-psychological-beliefs are not. As a result, we have less control over them and it is more difficult to find out about them through reflection alone. Additionally, though belief requires the capacity to represent phenomenal contents, trait-beliefs, internal-state-beliefs, and folk-psychological-beliefs need not. This allows non-conscious systems, like (presumably) insects, subpersonal systems of conscious beings, and artificial intelligences to have trait-beliefs, internal-state-beliefs, and folk-psychological-beliefs even if they cannot have genuine beliefs.

---

responsibility for their implicit biases. We are responsible for more than just our explicit beliefs. For instance, we're responsible for our actions, conscious thoughts, and patterns of thoughts and action (all of which might be constitutive of or the result of implicit bias). Insofar as we know or are in a position to know that we have or might have implicit biases or other harmful, unfair, or irrational tendencies to think, feel, or act, we are responsible for improving ourselves for our own sake and for the sake of others.

[34]Such internal states might nonetheless be helpfully characterized in terms of propositions or other content-like entities in virtue of their tracking certain items in their environment or playing certain functional roles, e.g., roles in inferences involving genuinely representational states. Though this would not amount to their representing in my sense, there is a deflationary sense of "representation" on which we can speak of such states as representing (Mendelovici 2018a, Chapter 8).

5 Conclusion

There is a notion of a belief as a cognitive state of bearing an acceptance-like attitude towards a proposition. I've proposed, very roughly, that beliefs in this sense are mere fictions in that they are mere products of our own self-attributions. But in virtue of having these fictions, they are real. While their psychological reality is fairly shallow, this does not mean they are wholly uninteresting or unimportant. They form part of the self-verifying fiction of the folk psychological mind, a bundle of self-ascribed beliefs, desires, and other propositional attitudes that go beyond what's in our immediate consciousness but that nonetheless forms part of our conception of ourselves as thinking, representing, and rationally-assessable beings. They're fictions, they're true, and they matter because they matter to us.[35]

References

Barsalou, L. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In Collins, A., Gathercole, S., Conway, M., and Morris, P., editors, *Theories of Memory*. Lawrence Erlbaum Associates, Hillsdale, N.J.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22:577–660.

Bourget, D. (2010). Consciousness is underived intentionality. *Noûs*, 44(1):32–58.

Bourget, D. and Mendelovici, A. (2016). Phenomenal intentionality. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Bourget, D. and Mendelovici, A. (forthcoming). Is intentionality a relation? a dialogue. *Argumenta*.

Braithwaite, R. B. (1932). The nature of believing. *Proceedings of the Aristotelian Society*, 33(n/a):129–146.

Davidson, D. (1984). *Inquiries Into Truth and Interpretation*. Oxford University Press.

Dennett, D. C. (1987). *The Intentional Stance*. MIT Press.

[35]Many thanks to David Bourget, Eric Schwitzgebel, Jonathan Jong, and an anonymous reviewer for extremely helpful comments and discussion.

Dennett, D. C. (1991). Real patterns. *Journal of Philosophy*, 88(1):27??51.

Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. MIT Press.

Eliasmith, C. (2013). *How to build a brain: A neural architecture for biological cognition*. Oxford University Press.

Farkas, K. (2008). Phenomenal intentionality without compromise. *The Monist*, 91(2):273–293.

Fodor, J. A. (1975). *The Language of Thought*. Harvard University Press, Cambridge, MA.

Fodor, J. A. (1987). *Psychosemantics*. MIT Press, Cambridge.

Graham, G., Horgan, T. E., and Tienson, J. L. (2007). Consciousness and intentionality. In *The Blackwell Companion to Consciousness*, pages 468–484. Blackwell.

Horgan, T. and Graham, G. (2012). Phenomenal intentionality and content determinacy. In Schantz, R., editor, *Prospects for Meaning*. De Gruyter.

Horgan, T. and Tienson, J. (2002). The intentionality of phenomenology and the phenomenology of intentionality. In Chalmers, D. J., editor, *Philosophy of Mind: Classical and Contemporary Readings*, pages 520–533. Oxford University Press, Oxford.

Horgan, T. E., Tienson, J. L., and Graham, G. (2004). Phenomenal intentionality and the brain in a vat. In Schantz, R., editor, *The Externalist Challenge*. De Gruyter.

Kriegel, U. (2003). Is intentionality dependent upon consciousness? *Philosophical Studies*, 116:271–307.

Kriegel, U. (2011). *The Sources of Intentionality*. Oxford University Press.

Kriegel, U. (2013). The phenomenal intentionality research program. In Kriegel, U., editor, *Phenomenal Intentionality*, pages 1–26. Oxford University Press.

Lewis, D. (1974). Radical interpretation. *Synthese*, 23.

Loar, B. (2003). Phenomenal intentionality as the basis of mental content. *Reflections and Replies: Essays on the Philosophy of Tyler Burge*.

Marcus, R. B. (1990). Some revisionary proposals about belief and believing. *Philosophy and Phenomenological Research*, 50(n/a):133–153.

Mendelovici, A. (2013a). Intentionalism about moods. *Thought: A Journal of Philosophy*, 2(1):126–136.

Mendelovici, A. (2013b). Pure intentionalism about moods and emotions. In *Current Controversies in Philosophy of Mind*, pages 135–157. Routledge, New York, NY.

Mendelovici, A. (2018a). *The Phenomenal Basis of Intentionality.* Oxford University Press.

Mendelovici, A. (2018b). Propositionalism without propositions, objectualism without objects. In Grzankowski, A. and Montague, M., editors, *Non-proposiitonal intentionality.* Oxford University Press, Oxford, UK.

Mendelovici, A. (2019). Immediate and reflective senses. In Shottenkirk, D., Curado, M., and Gouveia, S., editors, *Perception, Cognition, and Aesthetics*, pages 187–209. New York: Routledge.

Mendelovici, A. (2020). Propositional attitudes as self-ascriptions. In Oliveira, L. R. G. and Corcoran, K., editors, *Common Sense Metaphysics: Themes From the Philosophy of Lynne Rudder Baker*, pages 54–74. Oxford, UK: Routledge.

Mendelovici, A. (2023). Truth and content in sensory experience. In Kriegel, U., editor, *Oxford Studies in Philosophy of Mind Volume 3*, pages 318–338. Oxford, UK: Oxford University Press.

Mendelovici, A. (forthcominga). Attenuated representationalism. *Analysis.*

Mendelovici, A. (forthcomingb). Singular experiences (with and without objects). In French, R., Brogaard, B., and Bueno, O., editors, *The Roles of Representations in Visual Perception.* Springer.

Mendelovici, A. and Bourget, D. (2014). Naturalizing intentionality: Tracking theories versus phenomenal intentionality theories. *Philosophy Compass.*

Mendelovici, A. and Bourget, D. (2020). Consciousness and intentionality. In Kriegel, U., editor, *The Oxford Handbook of the Philosophy of Consciousness*, pages 560–585. New York, USA: Oxford University Press.

Millikan, R. G. (1984). *Language, Thought and Other Biological Categories.* MIT Press, Cambridge, MA.

Montague, M. (2022). Rethinking the Attitudes. In *Oxford Studies in Philosophy of Mind Volume 2.* Oxford University Press.

Murphy, G. L. (2004). *The Big Book of Concepts.* MIT Bradford, Cambridge.

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, 83(4):435–450.

Pautz, A. (2013). Does phenomenology ground mental content? In Kriegel, U., editor, *Phenomenal Intentionality*, pages 194–234. Oxford.

Pautz, A. (2021). Consciousness meets lewisian interpretation theory: A multistage account of intentionality. In Kriegel, U., editor, *Oxford Studies in Philosophy of Mind, Vol. 1.*

Pearce, S. (2016). A pure representationalist account of the attitudes.

Pitt, D. (2004). The phenomenology of cognition or what is it like to think that *P*? *Philosophy and Phenomenological Research*, 69(1):1–36.

Prinz, J. (2002). *Furnishing the Mind: Concepts and their Perceptual Basis.* MIT Bradford, Cambridge, MA.

Putnam, H. (1975). *The Meaning of "Meaning"*, pages 215–271. Cambridge University Press, Cambridge.

Quilty-Dunn, J. and Mandelbaum, E. (2018). Against dispositionalism: Belief in cognitive science. *Philosophical Studies*, 175(9):2353–2372.

Schwitzgebel, E. (2001). In-between believing. *Philosophical Quarterly*, 51(202):76–82.

Schwitzgebel, E. (2002). A phenomenal, dispositional account of belief. *Noûs*, 36(2):249–75.

Schwitzgebel, E. (2006). Belief. In *Stanford Encyclopedia of Philosophy.*

Schwitzgebel, E. (2013). A dispositional approach to the attitudes. In Nottelmann, N., editor, *New Essays on Belief*, pages 75–99. Palgrave.

Searle, J. (1990). Consciousness, explanatory inversion and cognitive science. *Behavioral and Brain Sciences*, 13:585–642.

Siewert, C. (1998). *The Significance of Consciousness.* Princeton University Press, Princeton.

Smith, E. E. and Medin, D. L. (1981). *Categories and Concepts.* Harvard University Press, Cambridge.

Strawson, G. (1994). *Mental Reality.* MIT Press, Cambridge.

Wickelgren, W. A. (1992). Webs, cell assemblies, and chunking in neural nets. *Concepts in Neuroscience*, 3(1):1–53.

Williams, J. R. G. (2019). *The Metaphysics of Representation.* Oxford: Oxford University Press.