

DOSSIER

A Dilemma for Naturalistic Theories of Intentionality

Um dilema para teorias naturalistas da intencionalidade

Michael J. Hegarty¹

ABSTRACT

I argue that a dilemma arises for naturalistic philosophers of mind in the naturalised semantics tradition. Giving a naturalistic account of the mind is a pressing problem. Brentano's Thesis — that a state is mental if, and only if, that state has underived representational content — provides an attractive route to naturalising the mental. If true, Brentano's Thesis means that naturalising representation is sufficient for naturalising the mental. But a naturalist who accepts Brentano's Thesis thus commits to an eliminativism about the category of the mental. This is because naturalistic theories of representation are reductive, and so over-generalise by applying to patently non-mental states. According to these theories, it has been argued, phenomena like tree rings and saliva come out as representational. Only proposing further Naturalistic conditions on representation could avoid the eliminativist conclusion. But this shows that Naturalists have made only limited progress towards naturalising the mental. And if a Naturalist rejects Brentano's Thesis, then she gives up on a clear link between representation and mentality. Hence, it is incumbent on the Naturalist to propose another, naturalistically acceptable, mark of the mental. This, again, shows that Naturalists have made only limited progress on the issue of naturalising the mental.

Keywords: Intentionality, representation, physicalism, eliminativism, Brentano, materialism, naturalism.

RESUMO

Eu argumento que surge um dilema para os filósofos naturalistas da mente na tradição semântica naturalizada e, assim, dar uma explicação naturalista da mente é uma questão urgente. A Tese de Brentano - de que um estado é mental se, e somente se, esse estado tem conteúdo representacional não derivado - fornece uma rota atraente para naturalizar o mental. Se for verdadeira, a Tese de Brentano significa que naturalizar a representação é suficiente para naturalizar o mental. Mas um naturalista que aceita a Tese de Brentano compromete-se, portanto, com um eliminativismo sobre a categoria do mental. Isso ocorre porque as teorias naturalistas da representação são redutivas e, portanto, generalizam excessivamente, aplicando-se a estados patentemente não mentais. De acordo com essas teorias, argumentou-se, fenômenos como anéis de árvores e saliva contam como representacionais. Somente a proposta de outras condições naturalistas sobre a representação poderia evitar a conclusão eliminativista. Mas isso mostra que os naturalistas fizeram ape-

¹ University of Connecticut,
Department of Philosophy, 344
Mansfield Road, Unit 1054, 06269-
1054, Connecticut, USA, Email:
michael.hegarty@uconn.edu.

nas um progresso limitado no sentido de naturalizar o mental. E se um naturalista rejeita a tese de Brentano, então ele abandona uma ligação clara entre representação e mentalidade. Consequentemente, cabe ao naturalista propor uma outra marca do mental, naturalisticamente aceitável. Isso, mais uma vez, mostra que os naturalistas fizeram apenas um progresso limitado na questão da naturalização do mental.

Palavras-chave: Intencionalidade, representação, fisicalismo, eliminativismo, Brentano, materialismo, naturalismo.

Introduction

An enduring problem in analytic philosophy of mind is how to locate the mental in the natural world. How are we to do full justice to, for example, phenomenal consciousness or the intentionality of thought within the confines of a materialist ontology? It has seemed that reducing mental facts to physical facts is the surest way to avoid the spectre of Cartesian Dualism.² But reductive materialism about the mental faces charges of being inadequate to the phenomena.³ Although significant work has been devoted to how phenomenal consciousness might be accommodated in a materialist account of the mental, it has recently been suggested that naturalising intentionality presents an even more difficult problem for materialism (Lycan, 2009, p. 553). In this paper I will argue that reductive naturalists attempting to naturalise the mental via naturalising intentionality encounter a dilemma.

Franz Brentano proposed that intentionality was the mark of the mental. This has since become known as Brentano's Thesis. In contemporary analytic philosophy of mind, Brentano's Thesis is commonly understood as offering necessary and sufficient conditions on mentality: a state is mental if, and only if, that state has underived representational content (Textor, 2019, p. 56). Brentano's Thesis is attractive because it promises to provide a way to separate the mental from the non-mental, and thus offers a clear path to naturalising the mental. In the form just stated, Brentano's Thesis implies that if representation is naturalised, then the intimidating task of naturalising the mental is complete. But, I will argue, a dilemma seems to present itself to the reductive naturalist at this point. If she accepts Brentano's Thesis then, I will argue, the naturalist ends up embracing eliminative materialism. If she rejects Brentano's Thesis, then, in having naturalised representation alone, she has made only limited progress on naturalising the mental. This is not a fatal dilemma, but it does suggest that naturalistic work on mental representation has made less progress on the issue of naturalising the mental than may be generally thought.

In Section 1 I clarify Brentano's Thesis and show how reductive naturalists of the kind I am interested in generally understand it. In Section 2 I sketch the broad naturalistic approach that I will focus on: the naturalised semantics tradition. Then I will explain the problem such naturalists face in naturalising intentionality and how they typically approach this task. In Section 3 I state the dilemma and explain how naturalists risk committing to eliminativism about the category of the mental if they adopt Brentano's Thesis as it is commonly formulated. That is the first horn of the dilemma. The other horn rejects Brentano's Thesis, but this then suggests that naturalists have made only limited progress towards naturalising the mental. In Section 4 I consider ways the naturalist might respond to the dilemma. In Section 5 I sketch some possible directions forward.

Before I begin, I should make some clarifications. I shall be assuming a representational theory of mind without defence — as do all the naturalists I consider here. That is, I will assume that intentional mental states come to be about things because they stand in relation to mental representations.⁴ Also, it is commonplace to distinguish two questions regarding mental representations. What makes a certain state a representational state? And: in virtue of what does a particular representational state come to represent what it does?⁵ Some theorists propose to answer both questions within the same framework (e.g. Millikan, 1984), while others advocate offering separate theories for each (e.g. Ramsey, 2007 and Fodor, 1990). I will speak generally of *naturalistic theories of representation*, and will not complicate matters by explaining a given theorist's stance on these issues. I do not think this affects my overall point.

1

The first step towards naturalising the mental is to clearly separate the mental from the non-mental. Franz Brentano proposed a way to do just that. In *Psychology from an Empirical Standpoint* he wrote:

Every mental phenomenon is characterized by... the intentional (or mental) inexistence

² For doubts about the viability of non-reductive physicalism see e.g. Schneider (2012).

³ See, e.g. Chalmers (1996) for an argument that materialism is threatened by an 'explanatory gap' in how consciousness could depend on the non-conscious. See also Lycan (2009), Kriegel (2013) and Pautz (2013) for criticisms of reductive approaches to intentionality.

⁴ See, e.g. Fodor, 1990, p. 16. See also Egan (2014) who glosses representationalism about the mind as "(...) the view that the human mind is an *information-using* system, and that human cognitive capacities are representational capacities" (2014, p. 115).

⁵ See, in particular, Ramsey (2007) for further discussion.

tence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning a thing), or immanent objectivity... This intentional inexistence is characteristic exclusively of mental phenomena. No physical phenomenon exhibits anything like it (1995, p. 88-89).

Brentano here seems to be pointing out some puzzling aspects of mental states (to put things in contemporary terms), and suggesting that these aspects are common to all and only mental states. In contemporary philosophy the remarks above are almost universally taken to commit Brentano to Brentano's Thesis: intentionality is the (distinctive and exclusive) mark of the mental. In other words, Brentano gives necessary and sufficient conditions for some states's being a *mental* state. It is also broadly agreed that 'intentionality' is the property something has of being *about* something other than itself. This aboutness has come to be understood as representation: a mental state like a perception of a red triangle can be said to be about the triangle because it possesses a certain representational content. A state is intentional just in case that state has representational content.

It is not a trivial matter to proceed from Brentano's original remarks on 'intentional inexistence' to identifying intentional states with those that have representational content. Proper attention to this may, however, provide the naturalist with a way out of the dilemma I will set up. So it is worth briefly unpacking how a naturalist might come to this commitment.

Ruth Millikan (2000) offers one path from Brentano's remarks to the thought that all intentional states are states with representational content. Millikan begins by noting how Brentano "equated intentionality with the capacity to bear a real relation to something nonexistent" (Millikan, 2000, p. 83) And she claims that Brentano was "surely mistaken" in thinking that this marks *only* the mental. She says that purposes are also capable of going unfulfilled, and hence can stand in relation to something nonexistent. And this suggests, she thinks, that focusing on natural purposes can provide a route to naturalising the phenomenon Brentano seeks to capture by pointing to this peculiar relation to nonexistent things. But, she notes, many systems do have natural purposes and yet fail to be *about* anything. The purpose of a stomach is to digest food, but stomachs are not about anything:

One thing that the example of natural purposes shows is that although being subject to Brentano's relation may be necessary to intentionality, if we take intentionality to imply aboutness it is not sufficient. Aboutness is associated with a purpose only when the purpose is explicitly represented. On the other hand, for there to be an explicit representation of a purpose, there must first be a purpose to

represent. The naturalist challenge here is to show, first, that the phenomenon of natural purposiveness can fulfill this second requirement, that a natural purpose may, equally naturally, be a represented purpose. Second, it needs to show how the same kind of analysis can be used also to naturalize intentionality in cases where facts are represented rather than purposes or ends. It must be shown, for example, how this way of naturalizing Brentano's relation can apply to the intentionality of beliefs, and of sentences that state facts. It needs to be shown in every case that what creates the appearance of Brentano's relation is merely an underlying natural purpose (Millikan, 2000, p. 84).

Millikan gives a detailed account of representation in terms of 'proper function' in *Language, Thought and Other Biological Categories*, thus claiming to fulfil her 'second requirement' that a natural purpose be a represented purpose (1984). This is one way to reason from Brentano's remarks to the view that intentionality just is representation. I will simply assume that intentionality is representation for now (as the naturalists I consider here do), and return later to question the assumption. Hence, a first pass at stating Brentano's Thesis (BT) more precisely is this:

(BT) A state, S, is mental if, and only if, S has representational content.

(BT) is supposed to provide necessary and sufficient conditions on mentality. All and only states with representational content are mental states. But, as is well known, there are obvious counterexamples to (BT). Sentences, maps and pictures are all representations yet surely do not count as mental states. These counterexamples appear to show that a state's having representational content is not sufficient for it to be mental. As a result, it is common to qualify (BT) by restricting it to 'underived' representational content. I will follow Neander (2017) in defining 'state with derived (intentional/representational) content' as a state with intentional content derived from some source of independently existing intentionality. Maps, sentences, and other artefacts derive their aboutness, it is commonly supposed, from human mental states with preexisting aboutness that either "is not derived from anything or it is derived from the nonintentional facts and properties of the world" (Neander, 2017, p. 9). Applying this qualification gives us the following version of Brentano's Thesis:

(BT*) A state, S, is mental if, and only if, S has underived representational content.

In this form, Brentano's Thesis is attractive because it

seems to provide necessary and sufficient conditions on mentality.⁶ It is common, but not universal, for philosophers to adopt (BT*) (See e. g. Dretske, 1995 and Lycan, 2008, p. 234)

(BT*) forges a link between representation and mentality, and thus promises a path towards naturalising the mind. If (BT*) is true, then naturalising representation would suffice for naturalising the mental. In the next section I will explain how the kind of naturalist I am considering here seeks to naturalise representation. This naturalist belongs to what is known as the ‘naturalised semantics’ tradition.

2

The naturalised semantics tradition begins with the view that representations have semantic and intentional properties that are essential to explaining how a representation is about its represented object. There is no well-defined single notion of naturalism (Neander, 2017, p. 3) But it is enough for my purposes to note that naturalists in the naturalised semantics tradition support a materialist ontology and hold that naturalistically acceptable theories and explanations appeal only to concepts and frameworks that are acceptable to our best scientific theories.⁷ Given the strictures of this naturalism, it is far from obvious how to locate the apparent semantic and intentional properties of representations in the natural order. Jerry Fodor (1984) gave what has become the canonical statement of the issue:

The worry about representation is... that the semantic/intentional properties of things will fail to supervene upon their physical properties. What is required to relieve the worry is therefore, at a minimum, the framing of naturalistic conditions for representation. [...] [W]hat we want at a minimum is something of the form ‘R represents S’ is true iff C where the vocabulary in which condition C is couched contains neither intentional nor semantical expressions (1984, p. 232).

In response to this challenge, philosophers in the naturalised semantics tradition sought to give conditions under which some state could be a representation in terms featuring no unreduced semantic, intentional, or psychological notions. Of course, much hangs on what we mean by ‘concepts and frameworks that are acceptable to our best scientific theories.’ Typically, Naturalists have relied on notions culled from the physical and biological sciences to naturalise representation.

Dretske (1981; 1986; 1988; 1995) relies on the notions of indicator function and information. Fodor (1987; 1990) gives a purely causal theory of content. Millikan (1984; 1989) explains representation in terms of historically normal proper function by drawing on natural selection. It is now commonly held that these naturalists attempt to give reductive explanations of representation (Burge 2010; Pautz 2013; Kriegel 2013; Morgan 2018). Henceforth, I will use ‘Naturalist’ and ‘Naturalism’ as labels to refer to this reductive project of naturalising representation and its practitioners. Naturalists thus endorse Fodor’s slogan that “If aboutness is real, it must be really something else” (1987, p. 97)

My question is: Can a Naturalist about the mind accept Brentano’s Thesis? Naturalists understand Brentano’s Thesis as (BT*). I have already mentioned how (BT*) is attractive because, if it is true, it appears to make the Naturalist’s project to naturalise the mind clear and tractable. However, in the next section I explain how this project produces a dilemma for the Naturalist. There are difficulties for the Naturalist both if they do accept (BT*) and if they reject it.

3

The dilemma arises because Naturalist theories of representation (as is well-known) result in giving representational status to all manner of patently non-mental phenomena. For example, Allen and Hauser (1992) complain that Millikan’s theory attributes representational content to certain interactions between acacia trees. Further, Kim Sterelny (1995) argues that Millikan’s theory entails that saliva ends up representing food. Alex Morgan (2014) criticises William Ramsey’s (2007) account for entailing that the internal states of some plants are representations. And finally, Morgan (2018) accuses Tyler Burge’s account of representation (2010) for also applying to the same internal states of plants.

The dilemma is this. A Naturalist can either accept or reject (BT*). If she accepts it, then she ends up committing to eliminativism about the mental. If she rejects it, then she has apparently taken only a small step towards naturalising the mental.

The first horn of the dilemma arises because, if (BT*) is true, then a fully naturalised account of representation must apply only to those states that are mental. This is just what (BT*) says in forging a link between the representational and the mental. But, if the sorts of complaints due to Sterelny, Morgan and others about various Naturalist theories are correct, then the category of the mental turns out to be far broader than is plausible. Surely, this line of thought goes, no internal states

⁶ Some philosophers also allege that there are counterexamples to the necessity of underived representational content to mentality. States like pains, fatigue, and generalised anxiety appear not to be about anything, yet are surely mental. I note this but will not consider it further here. See e.g. Crane (1998) for discussion.

⁷ For a representative example, Millikan characterises her project as “(...) a naturalist account of the specialness (or various kinds of specialness) of the mapping relation between a representation and its represented... an account that places this specialness in the realm of natural science, that is, of physics, physiology, biology, and evolutionary theory” (1984, p. 87).

of plants (to take Morgan's example) are really *mental* states. A Naturalist endorsing (BT*) must accept that, if their Naturalistic theory is true, then the category of the mental just turns out to be much broader than we may have thought.

But this is surely a form of eliminativism about the category of the mental. On our ordinary understanding of mentality, plants (and artefacts like thermometers, or natural phenomena like smoke⁸) just do not qualify. If the Naturalist disputes this (as they surely must if they hold (BT*) true and their account of representation applies to saliva, trees and so on), then this is a radical reimagining of what mentality is. And this is to admit the category of the mental is not the well-behaved one that we might have thought it to be. If a Naturalistic theory of representation like Millikan (1984) or Dretske (1981) is true, and (BT*) is true too, then it follows that mentality extends beyond what have traditionally been regarded as mental phenomena like perceptions, pains, beliefs and desires. A Naturalist might defend this reimagining of the category of the mental by questioning that there ever was a good reason to unify perceptions, pains etc. as exclusive to the mental domain in the first place. But this is precisely characteristic of eliminativism about the mental. It is reminiscent of Richard Rorty's eliminativist position: "The attempt to hitch pains and beliefs together seems ad hoc — they don't seem to have anything in common except our refusal to call them 'physical'" (Rorty, 1979, p. 22).

Further, we can note that such eliminativism carries a potential heavy cost. If our ordinary category of the mental is thrown out by the Naturalist, then what is the subject matter of psychology? It is incumbent on the Naturalist to say something about what the sciences of the mind are doing if the category of the mental — redrawn Naturalistically — includes things previously regarded as obviously non-mental. Tyler Burge raises essentially this criticism of (reductive) Naturalism. His main target is Millikan's biosemantics, which tries to naturalise representation by reducing representational function to biological function. Against this strategy, Burge argues that representation has a different explanatory role in psychology than biological function has in biology. The former, he thinks, must be given in terms of accuracy or veridicality conditions, whereas biological function is understood in terms of practical success or failure. This is a "root mismatch" between "representational error and failure of biological function," he thinks (Burge, 2010, p. 301) And it is this root mismatch which explains why we have psychological explanation at all. We lose this if representational function gets reduced to biological function (Burge, 2010, p. 300ff).

One response the Naturalist might make is to highlight that she seeks to offer a revisionary theory of representation. Hence, that the theory applies to things we ordinarily regard

as non-mental is a consequence we may expect. It is this revisionary nature that leads to the charge of eliminativism in the first place. And this revisionary project is entirely deliberate. For example, Millikan's self-stated core aim in her 1984 is to develop a "general theory of signs" (1984, p. 12). Mental representations are but one manifestation or application of this general theory. She intends to have her theory apply to a broad class of what she calls 'intentional icons' including: animal signals like bee dances and beaver tail slaps, linguistic devices like imperatives and indexicals, as well as mental states like beliefs.

She is not alone. Fodor observed that theories (like Dretske's) seeking to ground intentionality in information "might license the literal ascription of (underived) intentionality to thermometers, thermostats, and the like" (Fodor, 1990, p. 130). But, he further noted: "I don't think that should count as a reductio" (Fodor, 1990, p. 130). And Dretske is also explicit in early work: "intentionality, rather than being a 'mark of the mental,' is a pervasive feature of all reality — mental and physical" (1994, p. 285).⁹ And later he declares "there is no need to naturalize intentionality. It is already a familiar part of our physical world" (1981, p. 471) Thus, the Naturalist may suggest, we ought to (as good scientists do) let our best theories guide our conception of the world. If our best theory of representation attributes content to states that we think of intuitively as non-mental, then so be it. It is rational to revise our intuitions.

But this response makes no headway against the objection of eliminativism. It merely serves to justify why eliminativism about the commonsense category of the mental may be warranted in the light of Naturalistic theories of representation. Thus, it is not available to any Naturalist unhappy with eliminativism. And, furthermore, we are still owed an explanation of what the subject matter of psychology might be.

A stronger response the Naturalist can give is that the sorts of complaints that Sterelny, Allen and Hauser, and Morgan raise are fundamentally misguided for the following reason. In order to say that a Naturalist theory gets it wrong by telling us that plants have internal representations (where the having of internal representations is taken to indicate mentality via (BT*)) we need some independent grip on what things are *really* representations. Critics of Naturalism typically advance considerations from explanatory utility to show how smoke, tree rings, saliva, and Circadian clocks in plants fail to meet the requirements on being *real* representations. Indeed, William Ramsey's *Representation Reconsidered* is largely an analysis of what sort of role a state must play in order to qualify as *really representational*.

The thought is grounded in a certain approach to mental representation derived from philosophy of science. Mental representations, on this approach, are posits in theories that aim to explain behaviour and cognitive capacities.¹⁰ Crudely, repre-

⁸ Dretske (1981) is clear that these are intentional according to his informational semantics.

⁹ This remark would suggest that Dretske denies (bt*). I shall have more to say on this later.

¹⁰ Ramsey (2007) follows this strategy and attributes it originally to Stich (1983) and Cummins (1989). Shea (2018) is another adherent.

sentations are posited to explain the sort of flexible behaviour seen only in relatively sophisticated minded creatures. We do not need to posit representations to explain, say, the lawful behaviour of iron filings in response to a magnetic field. But we do need them (let us suppose) to explain a human agent's ability to (for example) plan out a walking route in advance. Hence, any theory that claims a certain state or system is representational must be mistaken if the explanation of 'behaviour' of that system gains nothing by supposing that the state is a representation, or that the system is representational.

A classic example is the magnetotactic bacterium prominently discussed in Dretske (1986). These water-dwelling bacteria have internal states that are responsive to magnetic fields. In their natural habitat, the bacteria move in response to the magnetic fields in the direction of oxygen-deficient water (which is beneficial for their survival). Both Dretske's theory (1986) and Millikan (1989) attribute representational status to these internal states. But, one might argue, there is *no explanatory benefit* to thinking of these states as representational. A cause-effect explanation (like in the case of the iron filings) would do just as well. Hence, there is justification for resisting Naturalist theories based on their being overly liberal in attributing representational status. Representations are mere idle cogs in the explanation of salivary response, or the magnetotactic bacteria, and other cases.

So, a certain view on the explanatory role of *representation* must be *presupposed* in order to criticise Naturalistic theories for attributing representational status too liberally. But, Naturalists like Millikan in particular will argue that this presupposition is not allowed. The presupposition relies on an analysis of the notion of *representation*. And Naturalistic theories like those of Millikan and Dretske try to give naturalistically acceptable accounts of just that notion. But they do not try to give *conceptual analyses* of the notion. The account of 'representation' in Millikan, Dretske and others is something akin to a hypothesis developed to give a general explanation of a diverse set of apparently similar phenomena. This is why Millikan aspires to give a 'general theory of signs'. She sees the common factor of relation to a nonexistent content — as identified by Brentano in the realm of the mental — as present in other phenomena like language and animal communication. She then attempts to give a unifying explanation of all of the diverse instances exhibiting this phenomenon. This is not *a priori* conceptual analysis. In fact, she strongly rejects this as a methodology. She puts things explicitly in later work, declaring that her intention is not conceptual analysis "(...) but rather of producing a notion that will organize certain natural phenomena for us in a way that casts light on

the apparently paradoxical nature of intentionality" (Millikan, 2000, p. 84-85). Thus, criticisms of Naturalism through the explanatory idleness of representations do not get off the ground according to the Naturalists themselves.

But this still does not escape the dilemma. The Naturalist approach just sketched allows them to undercut the accusations of insufficiency from, for example, Morgan. But the deeper point remains: what becomes of mentality, the familiar category? The Naturalist has said nothing to assuage the charge of eliminativism yet. So, either they embrace eliminativism and face questions about what this means for psychology, or they must admit that, yes, they believe in mentality *as such* and must accept that the Naturalist account of representation so far seems not to have completely captured it.¹¹

On the second of those last two options, the Naturalist may protest that a theory of representation could not — and was never intended to — be a theory of mentality. This is a reasonable line to take, and is endorsed by Naturalists such as Fodor (and Dretske (1988)), who sharply distinguishes a theory of content from an account of the functional roles of mental states. States with content, on Fodor's view, are only *mental* states when those states play the correct functional roles within the cognitive architecture of an organism (See, e.g. Fodor, 1987). But this seems to be effectively just a denial of (BT*) after all. This is because (BT*) identifies mentality with the possession of underived representational content, while the suggestion at hand is to add an additional condition to (BT*) specifying the functional role a state with content must play.

Of course, there is the option for the Naturalist to maintain (BT*) and attempt to add further naturalistically acceptable conditions to rule out the consequence that apparently non-mental phenomena are representational. This response would, however, show that the Naturalist does indeed need to provide further theory in order to have naturalised the mental.

Of course, a Naturalist may decide that (BT*) is mistaken and fits only uncomfortably with the Naturalist project. This option has only limited appeal in the context of the dilemma, however. Brentano's Thesis is attractive because it offers a mark of the mental: a way to separate the mental from the non-mental. As the Naturalist's task is to naturalise the mind, it is essential to have some way to separate mental from non-mental phenomena. Otherwise eliminativism threatens again.¹² But any Naturalist who wishes to resist eliminativism by denying (BT*) must thereby face the fact that they apparently made little progress in naturalising the mind.

While rejecting (BT*) may avoid eliminativism, it does so at the cost of abandoning a clear 'mark of the mental'. And any philosopher wishing to avoid eliminativism must

¹¹ It is always, of course, open to the Naturalist to deny (BT*). This may in fact be the most sensible move for her.

¹² Of course, some Naturalists will happily embrace this conclusion. William Ramsey is in fact an eliminativist, and his account of representation applies to patently non-mental entities like plants (Ramsey, 2007, p. 230). Thus, if my analysis in this paper is correct, his naturalistic account of representation must either entail eliminativism if combined with (BT*) or (the other horn of the dilemma) it cannot be sufficient for naturalising the mind. Clearly, Ramsey can happily embrace eliminativism. Moreover, theorists like Millikan who are suspicious of conceptual analysis for independent reasons may also be untroubled by the conclusion that their theory of representation entails eliminating the category of the mental. If the category is too fuzzy to be useful, maybe we are better off without it.

surely need some account of what distinguishes the mental from the non-mental. It has generally been thought that, if there is a mark of the mental, Brentano's Thesis gives us the best hope for that. Hence, a non-eliminativist Naturalist seeking to reject (BT*) must be prepared to give a naturalistically promising alternative mark of the mental. And this constitutes an admission that Naturalism about the mind has much more to do. Further, this new mark of the mental must be one that will not simply succumb to the same problems of over-generalising that the naturalistic reduction of representation faced. For otherwise the charge of eliminativism would also threaten again.

In this section I set out a dilemma for the Naturalist and considered what options the Naturalist has for getting out of it. Naturalist theories of representation broaden the notion of 'representation' as a key part of their reductive strategy to locate representation in the natural, physical world. Given this, Naturalist theories deem many intuitively non-representational phenomena representations. Thus, any Naturalist simultaneously affirming (BT*) seems to eliminate the category of the mental in the same stroke as they naturalised representation. Hence the Naturalist faces a choice: either embrace eliminativism, or accept that existing Naturalistic theories of representation are insufficient for naturalising mentality. The Naturalist does have a number of dialectical moves, such as rejecting conceptual analysis, and insisting that it is misguided to think that naturalising representation should ever naturalise mentality. But I have argued that each of these moves is merely a short-term fix and fails to avoid the dilemma. If a Naturalist rejects (BT*), then they give up a plausible mark of the mental and with it a clear path to naturalising the mental. Hence this suggests that the non-eliminativist Naturalist is in a difficult situation: either succumb to eliminativism, or accept that little progress has been made towards naturalising the mind. In the next section I will consider ways to dissolve the dilemma, and also suggest future directions for naturalising the mental.

4

I have been considering how the Naturalist may attempt to naturalise the mind. Brentano's Thesis is an attractive way to link representation and mentality, and a commonly accepted version of Brentano's Thesis is (BT*). But Naturalist theories of representation find themselves in a dilemma when attempting to naturalise the mind. Is there a way for the Naturalist to avoid this consequence?

There is, in fact, a questionable presupposition shared by both horns. Brentano's original claims about intentionality are notoriously obscure. So, as I mentioned, equating Brentano's Thesis with (BT*) is a substantive claim. Thus, in order to dissolve the dilemma, one might reject the widespread assumption that Brentanian intentionality simply is the possession of underived representational content. I explained how Millikan arrives at this interpretation. But her reasoning relied on the claim that "(...) Brentano equated intentionality with the capacity to bear a real relation to something non-existent" (Millikan, 2000, p. 83). This apparently interprets Brentano's idea of 'intentional inexistence', and it is commonplace in the analytic tradition to understand Brentano as emphasising how thoughts can be about nonexistent things. But some Brentano interpreters contend that this is a misreading. Tim Crane, for example, says that "Inexistence" expresses the idea that the object on which the mind is directed exists in the mental act itself" (Crane, 1998, p. 232).¹³

So there is reason for pause on whether understanding Brentano's 'intentional relation' as a relation to some possibly nonexistent object is correct. Thus (BT*) itself — formulated in terms of the underived representational content, and hence emphasising the semantic properties of mental representations — may be a presupposition to be rejected. Crane (1998) suggests that a more historically accurate way to think about Brentanian intentionality is as specifying that in every act something is given to the mind, in some sense (Crane, 1998, p. 243). Whether this idea can be made sufficiently precise without arriving again at something like underived representational content is a question for another paper.¹⁴ The suggestion just made amounts to this: the dilemma arises only because of a flawed and unworkable understanding of intentionality as underived representational content. So, it may be that Brentano's original observation cannot be captured by the dominant framework of conceiving of mental states as internal states with content.¹⁵

Another option that might dissolve the dilemma is to criticise the Naturalists for their particular brand of naturalism. As I mentioned, Naturalists attempt to reduce representation by giving a more 'naturalistically acceptable' account with no appeal to unreduced intentional, semantic or psychological notions. They typically rely on the physical and biological sciences for this project. As Burge points out, philosophers in the naturalising semantics tradition thus have a narrow and restrictive notion of what counts as a science (Burge, 2010, p. 296ff). In his view, it is mere dogma on the Naturalists' part to refuse to acknowledge the psychological sciences as legitimate sciences. And if disciplines like perceptual psychology are mature and

¹³ Moran (1996) makes a similar point. According to him, 'inexistence' simply refers to the Scholastic ontological notion of an accident inhering in a substance (Moran, 1996, p. 8).

¹⁴ It is also worth noting that Textor (2019) argues against understanding Brentano as offering a mark of the mental at all. Textor argues that Brentano intended his remarks to give "a piece of advice on how to become introspectively aware of the distinctive feature of mental phenomena" rather than offering an analytic definition of the mental (Textor, 2019).

¹⁵ Thanks to Dorit Bar-On for discussion on this point.

successful sciences, then there is no reason to demand that intentional or semantic notions like representation be reduced to physical or biological ones. Burge uses this strategy of broadening what *naturalism* is to say that representations are themselves primitive explanatory kinds (2010, p. 63). Broadening naturalism in this way thus opens the possibility of maintaining (BT*) while naturalising representation, without eliminating the category of the mental.¹⁶

So far I have considered two options for dissolving the dilemma. The first suggestion was that we might attempt to formulate Brentano's Thesis in different terms than (BT*). The second suggestion was to broaden 'naturalism' to include the psychological sciences, and therefore discharge naturalism from the task of having to find a non-intentional, non-psychological, non-semantic account of representation. Of course, the dilemma facing Naturalists is not fatal. Some Naturalists will happily grasp the eliminativist horn. And the other horn simply requires additional theoretical work. To close this paper, I will outline what this work might consist in.

It may be that separating mental states from non-mental states requires still further conditions beyond the mere possession of underived representational content. If this is so, then there will be states that are properly representations, and have a particular representational content, yet that are non-mental. But this need not commit us to denying that there is any real category of the mental so long as the further condition proposed is not susceptible to over-generalising. One attractive general suggestion for what this initial condition could be is that it should reference how a representational state with a particular content is used in a certain system. Only states that are used in the appropriate way by certain systems would then qualify as mental states.

Ramsey (2007) discusses this issue at length and proposes a 'use' condition to bridge the gap between a state's *merely having* a certain representational content and a state's being used *in virtue of* that content. This is also one of the driving principles in Dretske (1988). He writes: "The fact that [content-possessing mental states] have a content, the fact that they have a *semantic* character, must be relevant to the kind of effects they produce" (Dretske, 1988, p. 80) Both Dretske and Ramsey hold that a state with representational content must play a particular causal role internal to the system of which it is a part in order to count as a *mental* representation.¹⁷ Ram-

sey's use condition derives from the 'mental models' conception of mental representations found in Craik (1943), and developed by Swoyer (1991). In brief, the idea is that a system's internal state functions as a mental representation if that state is used as a surrogate or stand-in for what it represents.

However, any use condition like Ramsey's mental model conception must also be able to avoid a fatal 'homunculus' regress. A Naturalistic account of how the system uses a representational state *as* a representation of such-and-such cannot invoke any notions like 'understanding', 'taking' or 'seeing that'. If it did, then the explanation of the use condition would seem to require positing further internal representational states to explain how a system could 'understand' or 'take' a certain representational state as being a representation of such-and-such. Consequently, an infinite regress of representations appears to loom.¹⁸

To address this worry, Ramsey proposes a mechanistic account of how representations get used as surrogates by purely causal-physical systems. The key point is that causal-physical systems are, by definition, mindless. The threat of regress arises because it is notoriously difficult to say how a system could use a representation *as* a representation in virtue of its having a particular content without invoking further representations in an 'interpreting' system (Ramsey, 2007, p. 193-203). Hence, a purely mechanistic account of the use condition (if successful) would succeed in avoiding the regress because mechanistic explanations, by definition, do not make reference to any minds or interpreters.¹⁹ Unfortunately, this appears to face the very same kind of over-generalisation problem that dogged the original Naturalists in the naturalised semantics tradition. Morgan (2014) argues that Ramsey still has not succeeded in fully naturalising mental representation because Circadian clocks in plants apparently make use of representations in Ramsey's sense. Hence Ramsey's account threatens to apply too broadly and so veers towards eliminativism (though this is a consequence Ramsey is content with).

What is missing is some requirement to exclude obviously non-mental causal-physical systems from having states that qualify as representations by the use condition.²⁰ This is analogous to the problem Naturalistic theories of content face in applying too broadly because they explain having content in terms of a purely Naturalistic relation a state stands in to

¹⁶ Morgan (2018) argues that Burge's account of representation actually fails on this front because it attributes representational states to Circadian clocks in plants. I cannot evaluate Morgan's argument here. But his paper suggests that even a broadened understanding of naturalism may have some work to do to avoid the dilemma I have highlighted. If a naturalist account of representation (in the new, broader sense of naturalism) still attributes representational status to apparently non-mental states, then this new account is as susceptible to the charge of eliminativism as any of those from the naturalised semantics tradition.

¹⁷ See also Morgan (2014, p. 228ff) for discussion.

¹⁸ For attempts to show how this regress is not vicious, see Dennett (1978) and Lycan (1987).

¹⁹ See Morgan (2014) for criticism of Ramsey on this point.

²⁰ Morgan puts this general point like this in his 2018: "[S]ophisticated tracking theories [like those of Dretske, Millikan etc.] flesh out what it means for an informational state to be 'internal' by appealing to the distinctive ways in which those states are enlisted and used within a system such that a robust notion of accuracy finds purchase, but that this doesn't guarantee that informational states are integrated within the distinctive capacities of psychological systems" (2018, p. 5425).

what it represents. Ramsey faces over-generalisation because the use condition is given in strictly Naturalistic, mechanistic terms. The alternative is for the Naturalist to find another account of the use condition that is adequate to the task. Hence, we have again a stark choice for the Naturalist between eliminativism and the prospect that there is significant theoretical work still required to naturalise the mental.

Let me make some suggestive preliminary remarks about where to go next. The thought that a state may be a representation and have a determinate representational content without being a mental state suggests that we can distinguish between two kinds of internal states: consciously accessible, personal-level states and subpersonal states that are not consciously accessible.²¹ Once this is noted, it suggests that the problem of demarcating the mental is in the vicinity of the issue of distinguishing between personal and subpersonal states.²² And personal states are those that are consciously accessible to a subject, and hence can contribute to a subject's first-person perspective. Thus, one potential route forward might be to focus on what conditions a state would have to meet to be used in contributing to this perspective.²³ This is to recognise the importance of the subject in naturalising representation. But how exactly this may fit with any kind of naturalism about the mind is clearly a deep and difficult issue.²⁴

5

I have argued that a dilemma arises for Naturalists attempting to naturalise the mind. If a Naturalist accepts (BT*) as a version of Brentano's Thesis, then they appear to commit to an eliminativism about the category of the mental. This is because (BT*) is the thesis that having underived representational content is necessary and sufficient for a state's being mental. But Naturalistic theories of representation are reductive, and so over-generalise by applying to patently non-mental states. Only proposing further Naturalistic conditions on representation could avoid the eliminativist conclusion. But this shows that Naturalists have made only limited progress towards naturalising the mental. And if a Naturalist rejects (BT*), then she gives up on a clear link between representation and mentality. Hence, it is incumbent on the Naturalist to propose another, naturalistically acceptable, mark of the

mental. This, again, shows that Naturalists have made only limited progress on the issue of naturalising the mind.

References

- ALLEN, C.; HAUSER, M. 1992. Communication and cognition: is information the connection? *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, **2**: 81–91.
- BAR-ON, D. 2004. *Speaking My Mind: Expression and Self-Knowledge*. Oxford, Clarendon Press.
- BRENTANO, F. 1995. *Psychology from an Empirical Standpoint*. O. KRAUS (Ed.). New York, Routledge.
- BURGE, T. 2010. *Origins of Objectivity*. Oxford, Oxford University Press.
- CHALMERS, D. J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford, Oxford University Press.
- CRAIK, K. 1943. *The Nature of Explanation*. Cambridge, Cambridge University Press.
- CRANE, T. 1998. Intentionality as the mark of the mental. *Royal Institute of Philosophy Supplement*, **43**: 229–251.
- CUMMINS, R. 1989. *Meaning and Mental Representation*. Cambridge, Mass., MIT Press.
- DENNETT, D. C. 1978. *Brainstorms: Philosophical Essays on Mind and Psychology*. Hassocks, Harvester Press.
- DRAYSON, Z. 2012. The uses and abuses of the personal/subpersonal distinction. *Philosophical Perspectives*, **26**(1): 1–18.
- DRETSKE, F. I. 1981. *Knowledge and the Flow of Information*. Oxford, Basil Blackwell.
- DRETSKE, F. I. 1986. Misrepresentation. In: R. BOGDAN (Ed.), *Belief: Form, Content, and Function*. Oxford University Press, p. 17–36.
- DRETSKE, F. I. 1988. *Explaining Behavior: Reasons in a World of Causes*. Cambridge, Mass., MIT Press.
- DRETSKE, F. I. 1994. If You Can't Make One, You Don't Know How It Works. *Midwest Studies in Philosophy* **19**(1): 468–82. *Midwest Studies in Philosophy*
- DRETSKE, F. I. 1995. *Naturalizing the Mind*. Cambridge, Mass., MIT Press.
- EGAN, F. 2014. How to think about mental content. *Philosophical Studies*, **170**(1): 115–35.
- FODOR, J. A. 1984. Semantics, Wisconsin Style. *Synthese*, **59**(3): 231–50.
- FODOR, J. A. 1987. *Psychosemantics: The Problem of Meaning*

²¹ See Drayson (2012) for discussion on personal/subpersonal.

²² Similarly, Burge (2010) is concerned to mark the difference between mere 'sensory registration' and full-blown perception. However, he also appears to hold that what he calls 'subindividual' states can be properly representational "I believe that in principle, both particular thoughts and particular perceptual states can sometimes occur only at subindividual levels—that is, only in modular subsystems. In such cases, the representational natures of the states are still determined by causal patterns indicated by anti-individualism. But, paradigmatically, both perception and propositional thought are imputable to individuals. Unlike perception, propositional thought essentially involves an ability by individuals to engage in inference that depends on propositional form or structure" (2010, p. 104).

²³ Something like this suggestion is to be found in Morgan (2018).

²⁴ One possible route may be signalled by Bar-On (2004). She suggests that we overhaul the way we think about mental states as *internal states* of subjects. Instead, perhaps we could think of mental states as "(...) conditions the subjects are in, not states that are in the subjects" (Bar-On, 2004, p. 424, my emphasis).

- in the *Philosophy of Mind*. Cambridge, Mass., MIT Press.
- FODOR, J. A. 1990. *A Theory of Content and Other Essays*. Cambridge, Mass., MIT Press.
- KRIEGEL, U. 2013. The phenomenal intentionality research program. In: U. KRIEGEL (Ed.), *Phenomenal Intentionality*. Oxford, Oxford University Press, p. 1-26.
- LYCAN, W. G. 1987. *Consciousness*. Cambridge, Mass., MIT Press.
- LYCAN, W. G. 2008. Phenomenal intentionalities. *American Philosophical Quarterly*, **45**(3): 233–52.
- LYCAN, W. G. 2009. Giving dualism its due. *Australasian Journal of Philosophy*, **87**(4): 551–63.
- MILLIKAN, R. G. 1989. Biosemantics. *The Journal of Philosophy*, **86**(6): 281–97.
- MILLIKAN, R. G. 1984. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, Mass., MIT Press.
- MILLIKAN, R. G. 2000. Naturalising intentionality. *Proceedings of the Twentieth World Congress of Philosophy*, **9**: 83-90.
- MORGAN, A. 2018. Mindless accuracy: on the ubiquity of content in nature. *Synthese*, **195**: 5403–29.
- MORGAN, A. 2014. Representations gone mental. *Synthese*, **191**: 213–44.
- MORAN, D. 1996. The inaugural address: Brentano's thesis. *Proceedings of the Aristotelian Society Supplementary*, **LXX**: 1-27
- NEANDER, K. 2017. *A Mark of the Mental: In Defense of Informational Teleosemantics*. Cambridge, Mass., MIT Press.
- PAUTZ, A. 2013. Does phenomenology ground mental content? In: U. KRIEGEL (Ed.), *Phenomenal Intentionality*. Oxford, Oxford University Press, p. 194-227.
- RAMSEY, W. M. 2007. *Representation Reconsidered*. Cambridge, Cambridge University Press.
- RORTY, R. 1979. *Philosophy and the Mirror of Nature*. Princeton, Princeton University Press.
- SCHNEIDER, S. 2012. Why property dualists must reject substance physicalism. *Philosophical Studies*, **157**(1): 61–76.
- SHEA, N. 2018. *Representation in Cognitive Science*. Oxford, Oxford University Press.
- STERELNY, K. 1995. Basic minds. *Philosophical Perspectives*, **9**: 251–70.
- STICH, S. P. 1983. *From Folk Psychology to Cognitive Science: The Case Against Belief*. Cambridge, Mass., MIT Press.
- SWOYER, C. 1991. Structural representation and surrogate reasoning. *Synthese*, **87**(3): 449–508.
- TEXTOR, M. 2019. Brentano's empiricism and the philosophy of intentionality. *Philosophy and Phenomenological Research*, **98**(1): 50–68.

Submitted on October 07, 2020.
Accepted on December 07, 2020.