

# Consciousness and Intentionality

Angela Mendelovici and David Bourget

September 24, 2018

Philosophers traditionally recognize two key features of mental states: intentionality and phenomenal consciousness. To a first approximation, intentionality is the “aboutness” of mental states, and phenomenal consciousness is the felt, experiential, qualitative, or “what it’s like” (Nagel 1974) aspect of mental states. In the past few decades, these features have been widely assumed to be distinct and independent. But several philosophers have recently challenged this assumption, arguing that intentionality and consciousness are importantly related. This article overviews the key views on the relationship between consciousness and intentionality and describes our favored view, which is a version of the phenomenal intentionality theory, the view that the most fundamental kind of intentionality arises from phenomenal consciousness.<sup>1</sup>

---

<sup>1</sup>This article overviews many ideas that are developed in greater depth in Mendelovici 2018.

# 1 Consciousness and intentionality

Phenomenal consciousness is the felt, experiential, qualitative, or “what it’s like” (Nagel 1974) aspect of mental states. Some paradigm examples of mental states that exhibit phenomenal consciousness are sensations (e.g., pains, visual experiences) and emotional feelings (e.g., feelings of sadness or elation). For present purposes, we can define (*phenomenal*) *consciousness* ostensively as the salient feature of such states that is naturally described using terms like “what it’s like” and “experience”. Mental states, such as the emotion of joy or the perception of a rose, may have multiple properties, some of which are phenomenal and others of which are not. It is useful to have a term designating the purely phenomenal features of mental states. We will refer to these features as *phenomenal properties* and to instantiations of phenomenal properties as *phenomenal states*.

Above, we offered a gloss of intentionality as “aboutness”. This characterization, which is common in the literature, is merely a first approximation, rather than a strict definition. It is a fairly loose way of describing a phenomenon that we are able to at least sometimes notice introspectively in ourselves. The phenomenon is exemplified by thoughts, the kinds of mental states that we enjoy when we think, as well as by visual perceptual experiences. Both in thought and in visual experience, our mental states seem to “say” something, or be “about”, “of”, or “directed” at something, and it seems this requires no corresponding external entity or state of affairs. For

example, a perceptual experience might be described as being “about” a cup, and a thought might be described as “saying” that grass is green. We take *intentionality* to be this phenomenon that we notice introspectively in at least some cases and that we are tempted to describe using representational terms like “says”, “about”, “of”, and “directedness”.<sup>2</sup> A state’s (*intentional*) *content* is that which we are tempted to describe as what an intentional state “says” or is “directed at”. We will say that intentional states (*intentionally*) *represent* their intentional contents. While we are primarily concerned with intentional contents and intentional representation, we allow that there are other (arguably looser and more permissive) everyday uses of the term “represent” and that we can speak of the “contents” that are thus represented. As in the case of phenomenal consciousness, complex mental states might exhibit intentionality along with other features. We will call the purely intentional features of mental states *intentional properties* and the instantiations of intentional properties *intentional states*.<sup>3</sup>

---

<sup>2</sup>See Mendelovici 2010, 2018 and Kriegel 2011b for further development of ostensive ways of defining “intentionality”. One of us (DB) has tended to prefer a different definition of “intentionality” as a non-factive relation to propositions, which may or may not pick out the same thing as our present definition. DB’s more theoretically-loaded definition is suitable for his project in Bourget 2010a, 2010b, 2015, 2017a, 2017c, 2017d, 2017b, forthcoming a, and forthcoming b, where his aim is to shed light on consciousness in terms of non-factive relations to propositions. Part of our aim here, however, is to discover the nature of a phenomenon that we can introspectively observe in ourselves, so employing DB’s definition, or any other definition making substantive commitments with respect to the nature of intentionality, would beg the question in favor of certain views of the introspectively observed phenomenon. This is why we employ an ostensive definition for our purposes. See Mendelovici 2018 for an explicit defense of such an approach.

<sup>3</sup>The term “intentional state” is sometimes used to mean a state that has intentional properties. Since we are primarily interested in the relationship between consciousness and intentionality, and not the relationship between consciousness and other features of

Like our definition of “consciousness”, our definition of “intentionality” is ostensive. In the case of intentionality, our paradigm cases are thoughts and visual experiences. Standing propositional attitudes, such as beliefs that one counts as having even when not occurrently entertaining them, are also sometimes taken to be central cases of intentionality. However, we choose not to include standing propositional attitudes in our paradigm cases because they are not immediately observable through introspection in the same way that many thoughts and visual experiences are, and we believe that, when possible, it is preferable for ostensive notions to be grounded in the most immediately observable cases available.<sup>4</sup> Of course, how we define “intentionality” is merely a terminological choice. We will discuss this choice again when it becomes relevant below.

## 2 Three views of the relationship between consciousness and intentionality

Many mental states have both intentional properties and phenomenal properties. For example, when you see a rose, there is something it is like for you to see the rose, and your mind is seemingly directed at something, such as a

---

states that might happen to be intentional, such as their “modes” or “attitudes”, we use the term “intentional state” to pick out instantiations of intentional properties. On our usage, an instance of representing that grass is green is an intentional state, while a *belief* that grass is green is a mental state that involves the intentional state of representing that grass is green.

<sup>4</sup>See also Mendelovici 2018, §1.4.2.

rose or a possible state of affairs involving a rose. It is natural to ask what is the relationship between these two mental features. Roughly following Horgan and Tienson (2002a), we can distinguish three main views on this question: According to *representationalism*, all actual phenomenal states are nothing over and above, or *arise from*, intentional states (perhaps together with other ingredients).<sup>5</sup> According to the *phenomenal intentionality theory* (PIT), all actual intentional states, or at least all actual originally intentional states (more on this below), arise from phenomenal states.<sup>6</sup> According to *separatism*, neither kind of state arises from the other.

The notion of a set of states A arising from another set of states B is supposed to capture the intuitive idea that the states in A are nothing over and above the states in B. There are different ways in which a set of states A can arise from another set of states B: Every state in A might be identical to, grounded in, constituted by, or realized by some states in B (or some combination of such states).

Representationalism is often thought of as offering a theory of consciousness in that it tells us what consciousness arises from. According to representationalism, some intentional states, by their very nature, and perhaps together with the help of certain further ingredients, are phenomenally conscious or automatically result in phenomenal states. For example, a perceptual state representing a red square might, simply in virtue of representing a

---

<sup>5</sup>Introductions to representationalism include Lycan 2000, Bourget and Mendelovici 2014, and Seager and Bourget 2007.

<sup>6</sup>Introductions to PIT include Kriegel 2013 and Bourget and Mendelovici 2016.

red square, automatically have a “reddish” phenomenal character.

Similarly, PIT is often thought of as offering a theory of intentionality in that it tells us what intentionality arises from. According to PIT, certain phenomenal states, all by themselves, automatically give rise to intentional states. For example, a perceptual state with a “reddish squarish” phenomenal character might, all by itself, automatically result in the representation of a red square, or of there being a red square.<sup>7</sup>

Separatism denies both representationalism and PIT, maintaining that we cannot have a theory of consciousness in terms of intentionality or a theory of intentionality in terms of consciousness. The separatist might say that, although many states are both intentional and phenomenal, the intentional and the phenomenal are largely independent of one another. For example, a separatist might say that it is possible for a perceptual state to have a “reddish” phenomenal character but to represent the property of being green.<sup>8</sup>

A strong and simple form of representationalism states that all actual phenomenal states arise from intentional states alone. A strong and simple form of PIT states that all actual intentional states arise from phenomenal states alone. Most representationalists and phenomenal intentionality theorists do not hold these simple views. The main reason is that these views

---

<sup>7</sup>While PIT is primarily a view of how we represent various contents, many advocates of PIT and nearby views have also argued that the attitude components of propositional attitude states, like the belief component of the belief that grass is green, arise from phenomenal consciousness (Horgan and Tienson 2002a, Pitt 2004, Jorba 2016, Mendelovici 2018, Appendix E).

<sup>8</sup>See especially Block 1990 and 1996 for arguments for such a separatist view.

face challenges with intentional states that are not accompanied by any phenomenal states, such as the standing propositional attitudes that one has on a continuous basis (even when sleeping dreamlessly) and intentional states involved in early visual or linguistic processing that we are not aware of having. Given the reasonable assumption that such states can have the same contents as states that are accompanied by phenomenal consciousness, the simple version of representationalism faces a challenge, since these cases seem to show that phenomenal consciousness is not just a matter of having particular intentional states. Intentional states without accompanying phenomenal states also challenge the simple version of PIT because they seem to show that not all actual intentional states arise from phenomenal states.

These challenges have helped motivate weakened versions of representationalism and PIT. The simple version of representationalism described above is sometimes called *pure representationalism*, since it claims that all phenomenal states arise from intentional states *alone*. According to pure representationalism, all that matters for phenomenal consciousness is intentionality. The weakening of this view that is thought to avoid the above-mentioned problems is *impure representationalism*, which claims that all actual phenomenal states arise from intentional states *combined with other ingredients*, which might include functional roles, ways of representing, or intentional modes similar to the attitude components of propositional attitudes.<sup>9</sup> Impure representationalism can deal with the problem cases mentioned above by

---

<sup>9</sup>See Chalmers 2004 for the distinction between pure and impure representationalism.

denying that standing propositional attitudes and other non-phenomenally conscious states have the extra ingredients required for being phenomenally conscious.<sup>10</sup>

In the case of PIT, a different distinction is typically made. Let us call intentionality that arises from phenomenal consciousness alone *phenomenal intentionality*. A *phenomenal intentional state* is an intentional state that arises from one or more phenomenal states alone, and a phenomenal intentional state's content is its *phenomenal content*. The simple version of PIT mentioned above takes all intentionality to be phenomenal intentionality. We will refer to it as *strong PIT*. *Moderate PIT* is a weakening of this view according to which *some* intentional states are phenomenal intentional states, and all other intentional states derive in some way from phenomenal intentional states.<sup>11</sup> While moderate PIT recognizes non-phenomenal intentionality, it nevertheless maintains that phenomenal consciousness is the *source* of all intentionality (Kriegel 2011b, 2013).

Moderate PIT can be equivalently understood by use of a distinction

---

<sup>10</sup>Some versions of impure representationalism take the relevant extra ingredients to merely determine *whether* a phenomenal state arises given the presence of a particular intentional state—*which* phenomenal state it is that arises is determined by the content of the corresponding intentional state. This kind of impure representationalism is sometimes labeled *intermodal representationalism*, following Byrne (2001). *Intramodal representationalism*, in contrast, is a version of impure representationalism that takes the extra ingredients to help determine not only whether a phenomenal state arises given the presence of a particular intentional state, but also *which* phenomenal state it is that arises (see Lycan 1987). Bourget (2010a, 2010b, 2015, 2017b and 2017d, forthcoming) argues for intermodal representationalism and against intramodal representationalism.

<sup>11</sup>See Bourget and Mendelovici 2016 and Mendelovici 2018 for the distinction between strong PIT and moderate PIT. In Mendelovici and Bourget 2014, we use the terms “extreme PIT” and “strong PIT” to mark this distinction.



that is sometimes drawn between original and derived intentionality. *Derived intentionality* is intentionality that derives from other actual or merely possible instances of intentionality, while *original intentionality* is intentionality that is not derived. For example, it is sometimes thought that linguistic intentionality is a kind of derived intentionality in that the intentionality of linguistic expressions derives from the original intentionality of mental states. Moderate PIT, then, is the view that all original intentionality is phenomenal intentionality and any other intentionality is (ultimately) derived from phenomenal intentionality.<sup>12</sup>

Impure representationalism and moderate PIT weaken the simple versions of representationalism and PIT, respectively, but in different ways.<sup>13</sup> Impure representationalism denies that all actual phenomenal states arise from intentional states *alone*, allowing that ingredients apart from intentionality matter for phenomenal consciousness. Moderate PIT, in contrast, rejects the requirement that *all* actual intentional states arise from phenomenal states alone, allowing that some intentional states do not arise from phenomenal states, so long as they are instances of derived intentionality.

Why does the representationalist deny the “alone” part of the simple ver-

---

<sup>12</sup>Proponents of moderate PIT, or something close to it, include Bourget (2010a), Farkas (2008a,b), Horgan and Tienson (2002b), Horgan et al. (2004), Kriegel (2003, 2011a,b), Loar (2003a), Searle (1992), Mendelovici (2010, 2018), Mendelovici and Bourget (2014), Montague (2016), Pitt (2004, 2009, 2011), Pautz (2013), Siewert (1998), Smithies (2011, 2013a,b, 2014). See Section 4 of this article and Mendelovici 2018 for a defense of strong PIT. Bourget (2017c, 2018) also defends a view in the spirit of strong PIT.

<sup>13</sup>There are also weakened versions of PIT that weaken PIT by taking intentionality to arise from phenomenal consciousness together with other ingredients. Farkas (2013) and Masrour (2013) defend such views.

sion of her view while the advocate of PIT denies the “all” part of the simple version of her view? Recall that the representationalist aims to account for all phenomenal states, which involves specifying the conditions under which we have particular phenomenal states. Since intentional states do not uniquely determine phenomenal states, she cannot do so by invoking intentional states *alone*; she must invoke extra ingredients apart from intentionality. So, the simple version of representationalism is most naturally weakened to impure representationalism.

In contrast, the advocate of PIT aims to account for intentional states, which involves specifying the conditions under which we have any given intentional state. But, since phenomenal states are thought not to be necessary for all intentional states, she at best can only use phenomenal states alone to specify the conditions under which we have a subset of intentional states—these are the states with phenomenal intentionality. The intentional states that do not correspond to phenomenal states must be accounted for in some other way. This motivates weakening the simple version of PIT to moderate PIT, which takes some intentional states to be a matter of phenomenal consciousness alone and others to have merely derived intentionality.

The above points show that although the weakenings of the simple versions of representationalism and PIT are superficially quite different with respect to their methods of weakening, there is a deep agreement between the two strategies in that they both aim to accommodate intentional states that do not correspond to phenomenal states.

Before moving on, it is worth noting that, as we've defined the views, some, but not all, forms of representationalism and PIT are compatible with each other. For example, since identity is not asymmetric, some versions of representationalism and PIT taking the relevant arising relations to be identity relations are compatible with each other.<sup>14</sup> In contrast, versions taking the relevant arising relations to be grounding relations are not compatible with each other, since grounding is an asymmetric relation, so intentionality cannot ground consciousness while consciousness grounds intentionality.<sup>15</sup>

Since another paper in this volume is focused on representationalism, the rest of this paper will focus on PIT.

### 3 Motivating PIT

This section describes what we take to be a central motivation for accepting PIT as a theory of intentionality. The next section explores challenges to PIT and develops our favored version of PIT in response, which, we will see, is a version of strong PIT.<sup>16</sup>

As mentioned above, PIT can be understood as a theory of intentionality,

---

<sup>14</sup>We believe that identity versions of both views are true and have defended representationalism elsewhere. See Mendelovici 2018, Ch. 6 for discussion of why the compatibility of representationalism and PIT does not necessarily threaten the claim that the views provide theories of consciousness and intentionality, respectively.

<sup>15</sup>We explore other aspects of the relationship between representationalism and PIT in Bourget and Mendelovici 2016.

<sup>16</sup>See Bourget and Mendelovici 2016 and Kriegel 2013 for extensive discussions of a broad range of motivations, and Mendelovici 2018 for a more detailed treatment of the arguments presented here.

a theory that tells us what intentionally really *is*, metaphysically speaking. It is not a *naturalistic* theory in the traditional sense of a theory couched in physical-functional language, but it is nonetheless an attempt to explain intentionality, i.e., to describe its nature. Arguably, much of the interest in PIT stems from dissatisfaction with alternative theories of intentionality. In our view, one of the most important motivations for PIT is that its main competitors face unforgivable problems concerning empirical adequacy, while PIT does not.<sup>17</sup>

PIT's two main competitors are tracking theories and functional role theories. *Tracking theories* of intentionality maintain that original intentionality arises from tracking, which is detecting, carrying (or having the function of carrying) information about, or otherwise appropriately corresponding to items in the environment, such as particular objects, properties, or states of affairs. The tracking relations that have been thought to explain intentionality are supposed to be entirely reducible to physical features in the fashion championed by such authors as Dretske (1988, 1993), Fodor (1990a,b), and Millikan (1984).

*Functional role theories* maintain that original intentionality arises from functional roles, where the functional role of an inner representation (understood as some kind of token in the head) is the sum-total of the causal

---

<sup>17</sup>See also Kriegel 2013 and Mendelovici and Bourget 2014 for a critical assessment of PIT in comparison to alternative theories of intentionality, particularly the tracking theory. In Mendelovici and Bourget 2014, we also argue that PIT is naturalistic in the sense of “naturalism” that matters most.

relations that it is disposed to enter into with other inner representations, input stimuli (e.g., retinal stimulation), and outputs (e.g., bodily movements). On most versions of the view, it is only a subset of these causal relations that is required and relevant for intentionality, usually those causal relations corresponding to “correct” inferences. A hybrid theory taking original intentionality to arise from a combination of functional roles and tracking relations is also possible, and is sometimes called a *long-arm functional role theory* (see Harman 1987).

Tracking and functional role theories of intentionality have received considerable attention over the past few decades. For some time, it appeared that “naturalizing” intentionality by accounting for it in terms of tracking or functional roles was one of the most important goals in philosophy of mind. But this research program has lost momentum. Over time, it has become clear that offering an empirically adequate theory of intentionality in terms of tracking or functional role (let alone one that is genuinely explanatory) is very challenging.

Many challenges to the empirical adequacy of tracking theories have been lodged, most of which work against some tracking theories but not others. Rather than provide a general overview of these challenges here, we will focus on one kind of challenge to empirical adequacy that tracking theorists themselves have hardly discussed. The problem is that there are large classes of cases in which what a representation represents doesn’t match anything it can plausibly be said to match. Such cases are *mismatch cases*, and the

problem that they pose for tracking theories is the *mismatch problem*. Below we will describe one mismatch case, that of perceptual color representations.

Intuitively, when we visually represent the color red, we represent the vivid, striking, and warm quality that many of us are familiar with. Let us stipulate that this is what we mean by “redness”. If our visual states representing redness have their contents in virtue of what they track, they have to represent properties available to be tracked, which, on most tracking theories, are properties that are or have been instantiated in the actual world.<sup>18</sup> According to our best scientific understanding of apparently colored objects, the best candidate properties that are available for perceptual experiences of redness to track are properties such as the property *being disposed to primarily reflect electromagnetic radiation of wavelengths of around 650nm*. Call this property *EM650*. It is not very important here what is the best candidate physical basis of color, so we will use EM650 as an example without further discussion of other options. The problem for tracking theories is that redness, the property that we visually represent, and EM650, the property that our visual states track, seem to be entirely different properties. One is categorical, vivid, striking, and warm. The other is dispositional and has to do with electromagnetic radiation and wavelengths. The two properties differ

---

<sup>18</sup>There are tracking theories that allow us to track properties that have never been instantiated, such as Fodor’s asymmetric dependence view. However, the asymmetric dependence theory requires lawful connections between tracked properties and inner representations to obtain and be relatively strong, which is a condition that is not plausibly met in the kinds of cases we will discuss, so it does not help the tracking theorist avoid the mismatch problem. See Mendelovici 2013b, 2016, 2018, Appendix A for more details.

in their higher-order properties, so, by Leibniz's law of the indiscernability of identicals, they are distinct properties. The same problem arises in other cases, such as experiences of hotness and coldness, sweetness, moral or other kinds of value, and thoughts about many of these same contents.<sup>19</sup>

There are many objections one might make to the mismatch problem. We will only discuss one, which we think might seem particularly compelling<sup>20</sup>: One might object that apparent differences between EM650 and redness are illusory. One might draw an analogy with the case of the apparent distinctness of physical and mental properties. The mental and the physical, one might say, seem different, but, it might be argued, this is compatible with mental properties being identical to physical properties. It is just that we represent mental properties in a special way, perhaps using a special "mode of presentation", which makes them seem distinct from physical properties. Perhaps, similarly, EM650 and redness are one and the same property, but we do not realize this because we represent it in two different ways. But note that there is an important difference between the argument from the mismatch problem against tracking theories and the well-known arguments against physicalism. The arguments against physicalism rest on the observation that we cannot a priori infer mental facts from physical facts.<sup>21</sup> In contrast, the argument from the mismatch problem rests on the observation

---

<sup>19</sup>The mismatch problem for tracking theories is developed in detail in Mendelovici 2018, Ch. 3.

<sup>20</sup>See Mendelovici 2018, Ch. 3 for more objections and replies.

<sup>21</sup>See, e.g., Chalmers 1996.

that redness and EM650 have distinct higher-order properties. While a lack of inferability might perhaps be explained in terms of ways of representing (as opposed to a real difference in properties), differences in higher-order properties between two properties *entail* non-identity (by Leibniz's law). On the face of it, the typical physicalist reply is not applicable.<sup>22</sup>

One might try to apply the reply at the level of higher-order properties. One might say that properties such as those of being vivid, striking, and warm are, despite appearances, physical features of electromagnetic properties, and having to do with wavelengths and electromagnetic radiation are features of redness. One might also say that redness in fact lacks some of the properties that it seems to have, such as being categorical. By identifying certain higher-order properties that we attribute to EM650 and redness and rejecting others, one might attempt to undermine the argument from Leibniz's law. Of course, one *can* simply make the relevant claims, but, we maintain, there is little motivation independent of saving the tracking theory to think that they are true. It is *always* possible to save a theory by positing errors of judgment and illusions of non-identity like this *without independent evidence*. Absent independent reasons to think we are making such errors in this case, the reply is unconvincing.

---

<sup>22</sup>There are arguments for dualism that take the same form as the above argument, such as Descartes' argument that the mind and the body must be distinct because the former is unextended while the latter is extended. However, the best physicalist response to such arguments is not to postulate some kind of illusion of non-identity, but instead to deny that the mind (or the body) has (or lacks) the relevant property. This kind of reply seems more plausible in the case of the mind than in the case of properties such as colors.



Let us now turn to the functional role theory. The idea behind this theory is that the overall pattern of functional relations between mental representations (and perhaps their components) determines their intentional contents. The problem with this is easiest to see if we adopt the framework of the language of thought (Fodor 1975). Let us say that our mental representations are formulas in some inner language  $L$ . Causal connections that our inner formulas and their constituent symbols stand in to other formulas and symbols are supposed to determine their intentional contents. Let us model the contents of formulas and their constituent symbols as intensions, which are functions from possible worlds to entities or set-theoretic constructions out of entities (truth values, objects, sets of objects, etc.). Let us assume that the intension of a formula is determined by the intensions of its constituent symbols and their logical arrangement. Causal role is supposed to determine contents through such constraints as this:

*Representations  $A$  and  $B$  represent contents  $C_A$  and  $C_B$ , respectively, in virtue of their functional role only if it is the case that  $A$  causes  $B$  iff  $C_A$  entails  $C_B$ .*

Whatever the exact content-determining rules that one might want to specify, the causal role account proceeds by mapping causal relations between inner representations (broadly understood) to logical relations between their contents: the causal relations between a certain set of representations determine what logical relations (e.g., entailment) obtain between the contents of representations of this set. The logical relations are then supposed to

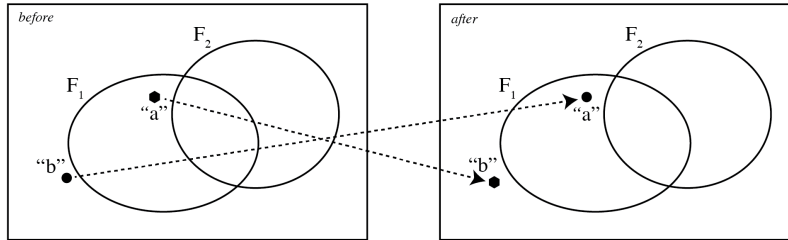


Figure 1: Constructing alternative interpretations for mental symbols

determine the specific contents.

One problem is that logical relations are not sufficient for determining contents.<sup>23</sup> This can be shown using a method similar to that used in Putnam’s model-theoretic arguments (Putnam 1981, Appendix). Let us suppose that there is at least one interpretation  $I_1$  of the symbols in  $L$  that is consistent with their causal roles. We can think of  $I_1$  as assigning intensions to all the non-logical symbols in  $L$ . Statements in  $L$  get their intensions compositionally. Assume that some predicate  $F_1$  in  $L$  is non-trivial at some world  $w_1$ , in that  $F_1$  is true of some objects at  $w_1$  (e.g., object  $a$ ) and false of other objects at  $w_1$  (e.g., object  $b$ ). Now picture the set of all objects in  $w_1$  laid on a surface such as the left rectangle in Figure 1. Imagine the extensions of all names and predicates in  $L$  at  $w_1$  specified by  $I_1$  being marked as points (for names) and shapes (for predicates) on this surface, as illustrated in the left rectangle in Figure 1 for names “ $a$ ” and “ $b$ ” and predicates  $F_1$  and  $F_2$  (each symbol is shown next to its extension).<sup>24</sup> Now take two objects such

<sup>23</sup>See Bourget MS for the argument that follows.

<sup>24</sup>The extensions of monadic predicates could be non-contiguous shapes without this affecting the argument. A different drawing convention would have to be used for non-

that one is in the extension of  $F_1$  at  $w_1$  and the other is not, for example, the extensions of “a” and “b”, a and b. If we swap the places of a and b while leaving all labels in place, we obtain new extensions for “a”, “b”, and  $F_1$ , and any other symbol that has a or b in its extension. The result of swapping a and b is illustrated in the right rectangle of Figure 1. The new extension of “a” is b, the new extension of “b” is a, and the new extension of predicates are the same as on  $I_1$  except that a and b are swapped. This swapping procedure does not change the truth value of any statement at  $w_1$ . We can thus use this procedure to construct an alternative to  $I_1$  that makes all the same statements true at any world and preserves all logical relationships between statements: Let  $I_2$  be the interpretation that assigns to each expression  $e$  in  $L$  the intension whose values are defined as follows: at  $w_1$ , it is the value at  $w_1$  of the intension assigned to  $e$  by  $I_1$ , except that a and b are swapped as described above; at any other world, it is the value at that world of the intension assigned to  $e$  by  $I_1$ . By construction, the intensions assigned by  $I_2$  yield the same truth-values as the intensions assigned by  $I_1$  for all sentences of  $L$  and all worlds (including  $w_1$ ). So, all broadly logical relations such as entailment between the sentences of  $L$  are exactly the same on  $J$  and on  $I_1$ . However,  $I_1$  and  $I_2$  are clearly different interpretations, since they assign different intensions and extensions to “a”, “b”, and  $F_1$  (and potentially many other mental symbols). This shows that determining logical relations be-

---

monadic properties, whose extensions are sets of n-tuples; we are setting those aside for simplicity.

tween contents is not sufficient for determining contents. Our Putnam-style procedure constructs minimally differing extensions and intensions, but it is easy to see that we can also construct massively different extensions and intensions, since many worlds contain large numbers of objects, which can be swapped. Even if some of our contents in fact are a little bit indeterminate, it is implausible that they are massively indeterminate in this way.<sup>25</sup>

It might be thought that the long-arm functional role theory avoids the underdetermination worry for functionalism, since it takes functional roles *and* tracking relations to be relevant to content determination. For example, it can say that certain color representations get determinate color contents from tracking relations, while other color representations get their contents

---

<sup>25</sup>One might suggest that the functional role theory could determine probabilistic relations rather than logical relations and that it would be harder to generate deviant interpretations consistent with probabilistic relations (Fine (1977) discusses a view in the ballpark, but he does not offer it as a theory of content). It may be that our swapping procedure would not preserve probabilistic relations, but such a theory of mental content would rely on the existence of objective, mind-independent conditional probabilities between arbitrary propositions, and we are skeptical that there are such probabilities (for what it's worth, Fine talks only about subjective probabilities).

Another possible response is that interpretations should be limited to functions that assign “natural” intensions, which would presumably be intensions that assign “natural” extensions. This is Lewis’ “reference magnetism” reply to Putnam. The resulting view would not be a version of the functional role theory, but a view aiming to account for intentionality in terms of both functional roles and reference magnetism. We are not convinced that facts about natural kinds can play the required roles, since it is unclear why natural kinds should constrain interpretation in the required way—in other words, it is unclear why natural kinds should be “magnetic”. See Chalmers 2012 (extended edition, 20th excursus) for more objections to reference magnetism.

Other determinacy worries for conceptual role theories arise from “Kripkenstein”-style considerations (see Kripke 1982). BonJour (1998, 176-7), Graham et al. (2007), Searle (1990), Strawson (2008), and Kriegel (2011b) also raise worries concerning content determinacy for tracking and functional role theories of intentionality. See also Pautz 2013 for critical discussion.

from their relations to these color representations. Tracking relations provide a representational system with determinate contents, which are then transformed and passed around to other representations. However, insofar as tracking relations play a role in determining content, long-arm functional role theories inherit the tracking theory's mismatch problem, since the content allegedly provided by tracking relations is sometimes the wrong content.<sup>26</sup>

The underdetermination problem and the mismatch problem show that the functional role theory and the tracking theory cannot attribute content correctly, that they are empirically inadequate. The mismatch problem shows that the tracking theory makes the wrong predictions in mismatch cases, while the underdetermination problem shows that the functional role theory cannot give the right answer in all cases of non-trivial predicates for in-principle reasons. These are strong reasons to reject the views.<sup>27</sup>

We believe that PIT can attribute content correctly in all cases. Since it does not take content to be determined by logical relations between representations, it does not face the functional role theory's in-principle worries with correct content attribution. It can also yield the right answer in the tracking theory's mismatch cases. Some versions of PIT simply identify intentional states with phenomenal states, while others take phenomenal states to realize, constitute, or ground intentional states. Either way, PIT has the

---

<sup>26</sup>See Mendelovici 2018, Ch. 4.

<sup>27</sup>In Mendelovici and Bourget 2014 and Mendelovici 2018, Ch. 6, we argue that failure of empirical adequacy cannot be made up for by having other virtues, such as that of being naturalistic.

resources to ascribe the right content in the case of experiences of redness, since they involve a phenomenal character that matches the content we want to ascribe. This phenomenal character might simply be identical to the content, or it might realize, constitute, or in some other way ground it. Similar claims can be made about other mismatch cases for the tracking theory.

The preceding does not conclusively show that PIT ascribes correct, or even determinate, content in all cases. Many objections to PIT concern cases where it appears not to attribute content correctly. We turn to such cases in the next section. If what we say there is correct, then PIT is arguably empirically adequate, which provides significant support for the view, particularly when its main competitors are not.<sup>28</sup>

## 4 Challenges to PIT

In the previous section, we outlined what we take to be some important motivations for PIT. We now turn to various challenging cases for the view. As we will see, different ways of handling these cases result in different versions of PIT. We will argue for an approach that results in a version of strong PIT, which claims that all intentionality is phenomenal intentionality.

There are four main kinds of challenging cases we will consider: conscious

---

<sup>28</sup>Empirical adequacy, of course, is not enough to show that a view is viable. In order for PIT to succeed, phenomenal consciousness has to be metaphysically sufficient to give rise to intentionality. Another line of argument for PIT aims to establish that while tracking and functional relations do not have the power to give rise to intentionality, phenomenal consciousness does (see Mendelovici 2018, Chs. 4 and 5).

thoughts with complex or abstract contents that don't seem to correspond to phenomenal states, intentional states with wide contents, standing propositional attitudes, and nonconscious representations of the kind described by cognitive science. We will overview each kind of case in turn before providing a sketch of how proponents of PIT might deal with them.

**Thoughts.** It seems that we have all sorts of complex or abstract thoughts, some of which represent entities such as political systems, norms of behavior, unobservable particles, and highly abstract mathematical entities. When we have such thoughts, it might not seem that we have correspondingly complex or abstract phenomenal states. Unlike in the case of experiences of redness, where the feel of an experience seems to match what is represented, in the case of complex or abstract thoughts, there seem to be no phenomenal states matching what is represented. This casts doubt on strong PIT, which requires that all intentionality is phenomenal intentionality. Moderate PIT, which takes all intentionality to be phenomenal intentionality or derived from phenomenal intentionality, is not committed to every intentional state corresponding to a matching phenomenal state. But the case of complex and abstract thoughts also casts doubt on moderate PIT, since it is unclear that such thoughts are related to phenomenal consciousness at all.

**Wide intentional states.** A problem also arises with wide intentional states, which are states whose contents at least partly depend on factors external to the individual whose states they are. If Twin Earth intuitions are

right (see Putnam 1975), then Oscar's thought that water is wet represents the content  $\langle \text{H}_2\text{O is wet} \rangle$ . But, on the plausible assumption that phenomenal states are internally determined, Oscar has no phenomenal state matching  $\langle \text{H}_2\text{O} \rangle$ . Similarly, wide intentional states involving singular contents, such as the thought you might have with the content  $\langle \text{Justin Trudeau is currently in Europe} \rangle$ , do not come with a phenomenology uniquely matching their singular contents. Here, too, the problematic cases directly challenge strong PIT, since wide contents clearly do not seem to be phenomenal contents, but the cases also challenge moderate PIT, since it is not clear how such contents might be related to phenomenally intentional states.

**Standing propositional attitudes.** Standing propositional attitudes are beliefs, desires, and other propositional attitudes that we count as having even when we are not occurrently undergoing them. The problem with standing propositional attitudes is that there is nothing that it is like to have them. For example, there is nothing it is like to believe that monkeys like bananas, at least when not occurrently entertaining this belief. So, it does not seem that standing propositional attitudes arise from phenomenal states, which makes them problematic for strong PIT. It is also unclear how they might be related to consciousness at all, which makes them problematic for moderate PIT.

**Nonconscious representational states.** Cognitive science describes all sorts of representational states that seem to be intentional but independent



of any phenomenal states we might have. For example, representations occurring in early visual processing and our tacit knowledge of grammar seem to have no echo in phenomenal consciousness, yet one might hold that they are intentional. Again, this would directly contradict strong PIT and cast doubt on moderate PIT.

In addition to the above challenges, it might be claimed that phenomenal states that might be thought to lack intentional content, such as feelings of elation and headaches, are a challenge for PIT. Such cases pose a challenge to representationalism, but they are not a problem for most versions of PIT, which only claim that all intentional states arise from phenomenal states, not that every phenomenal state gives rise to an intentional state. For this reason, such cases do not concern us here.<sup>29</sup>

Note that standing states and the nonconscious representational states posited by cognitive science are precisely the kinds of states that motivate impure representationalism over pure representationalism and moderate PIT over strong PIT, as we saw in Section 2. In what follows, we will reconsider these motivations for moderate PIT, eventually arguing that strong PIT is in fact correct.<sup>30</sup>

---

<sup>29</sup>See Block 1996 and Kind 2003, 2013 for worries with representationalism based on the cases of moods, pain, and other such cases, and Harman 1990, Tye 1995, 2008, Dretske 1995, Byrne 2001, Crane 2003, Bain 2003, Seager and Bourget 2007, Mendelovici 2013a, 2014, Bourget and Mendelovici 2014, Bourget 2015, 2017b, and Smithies forthcoming for defenses of representationalist treatments of such cases. See Mendelovici 2018, Appendix B for further discussion of the relevance of such cases to PIT.

<sup>30</sup>Similar arguments can also show that pure representationalism is defensible. See Mendelovici 2010, 2013a, and 2014 and Bourget 2010a,b.

For any problematic state, there are three main strategies that a proponent of moderate PIT might adopt. *Inflationism* claims that the problematic state, despite appearances, has rich phenomenal character from which its content arises. *Eliminativism* denies that the problematic state exists or has any intentionality at all. *Derivativism* claims that, while the problematic intentional state does not arise solely from phenomenal consciousness, it derives from intentional states that do. The first two strategies, but not the third, are open both to the proponent of strong PIT and to the proponent of moderate PIT.

We will now consider how each kind of strategy can be applied to the problematic cases, focusing on our favored strategies.

#### 4.1 Standing propositional attitudes

In the case of standing propositional attitudes, inflationism seems to be a nonstarter: there is clearly no phenomenology associated with most of our standing beliefs and other standing propositional attitudes. A more promising strategy is an eliminativist strategy that flat-out denies the existence of anything answering to the notion of a standing propositional attitude. While we are sympathetic to this strategy, we think a more nuanced eliminativist strategy is preferable. We will turn to it after considering a related derivativist strategy.

One promising derivativist strategy takes propositional attitudes to be dispositions to have certain related occurrent beliefs, occurrent desires, or

other thoughts, whose contents are either phenomenal contents or derived from phenomenal contents. On this view, which we might call *derivativist dispositionalism*, propositional attitudes and their contents derive from dispositions to have occurrent thoughts, whose contents are either phenomenal contents or derived from phenomenal contents. For example, your belief that monkeys like bananas might amount to a set of complex dispositions to have occurrent beliefs to the effect that monkeys like bananas (or perhaps occurrent beliefs that are entailed by such occurrent beliefs) whenever relevant.<sup>31</sup>

There is also an eliminativist version of the dispositionalist strategy, which we find preferable (though there is another view we are also partial to, which we will describe in Section 4.3). This *eliminativist dispositionalism* accepts that we have dispositions to have various occurrent thoughts and that these play many of the roles we associate with standing propositional attitudes. Unlike the derivativist dispositionalist, however, the eliminativist dispositionalist denies that the relevant dispositional states qualify as intentional states. Recall that we defined intentionality ostensibly by pointing to paradigm cases in thought and visual experience. Since a disposition to do X is different in nature from doing X, the relevant dispositions are different in nature from our paradigm cases, so our definition plausibly excludes

---

<sup>31</sup>Searle's (1989, 1990) potentialism is arguably best understood as a form of derivativism about standing propositional attitudes. Searle takes standing states that are potentially conscious to derive their intentionality from the phenomenal states they are disposed to cause. Kriegel's *interpretivism* (2011a,b) also provides a derivativist view of standing propositional attitudes, taking non-phenomenally conscious intentional states to be derived from the phenomenal intentional states of an ideal observer applying intentional systems theory to subjects based on their phenomenal intentional states and behavior.

them.<sup>32</sup> Of course, whether or not the relevant dispositional states qualify as intentional states depends on how we define “intentionality”. If we were to count propositional attitudes as paradigm cases of intentionality, then eliminativist dispositionalism would likely end up being classified as a derivativist dispositionalism.

Unlike the eliminativist strategy that flat-out denies the existence of anything answering to the notion of a standing propositional attitude, both derivativist and eliminativist dispositionalism attempt to be somewhat accommodating to our prior views of standing propositional attitudes, accepting that we have states playing the roles of standing beliefs, standing desires, and other standing states, even though their nature is merely dispositional, and perhaps not even genuinely intentional. In order for these strategies to succeed at accommodating standing propositional attitudes, however, PIT needs to be able to accommodate occurrent thoughts with the relevant contents, which might include contents that are complex or wide. We turn to these challenging cases below, starting with the case of wide thoughts.

## 4.2 Wide thoughts

In the case of occurrent thoughts with wide contents, inflationism, again, seems to be a nonstarter. Take for example the occurrent thought that monkeys like bananas. The wide content of this thought relates creatures

---

<sup>32</sup>See also Strawson 1994, p. 167, Mendelovici 2018, Ch. 8, and Pitt MS for applications of the eliminativist strategy in the case of standing propositional attitudes.

with a certain kind of DNA or evolutionary history to bananas. It seems implausible that there is a phenomenal character of thought that captures this specific kind of DNA or evolutionary history.

One kind of eliminativist strategy might simply deny that there are any wide contents, perhaps suggesting that we are mistaking referents for wide contents (see Farkas 2008a) or that our intuitions supporting wide contents are mistaken and all we really have are narrow contents (Pitt 1999, 2011). The view we will ultimately defend is also eliminativist, but it is eliminativist in a slightly more accommodating way.<sup>33</sup>

The derivativist strategy is quite plausible and widely endorsed among advocates of PIT. One natural version of this strategy takes thoughts to have both wide and narrow contents, with the wide contents deriving from the narrow contents. These narrow contents are phenomenal contents (or at least derived contents that are derived from phenomenal contents).<sup>34</sup> For example, the thought that water is wet might have a descriptive narrow content like <the clear watery stuff around here is wet>, which, in certain contexts, determines derived wide contents like <H<sub>2</sub>O is wet>. We will call this strategy the *derivativist descriptivist strategy* for wide thoughts, since it takes wide contents to be derived from broadly-speaking descriptive narrow contents.<sup>35</sup>

---

<sup>33</sup>See Siewert 1998, Kriegel 2007, Farkas 2008a, Pitt 1999, 2011, and Mendelovici 2010, 2018 for applications of the eliminativist strategy to wide states.

<sup>34</sup>Such views are defended by Horgan and Tienson 2002a, Horgan et al. (2004), Loar (2003), Bourget (2010a), and Chalmers (2010), among others. Mendelovici (2010, 2018) defends an eliminativist version of this kind of view.

<sup>35</sup>This strategy requires a broadly descriptivist view such as that defended by Jack-

As in the case of derivativist dispositionalism, there is also an eliminativist version of the descriptivist strategy. According to *eliminativist descriptivism*, thoughts have narrow descriptive contents which determine wide contents, but these wide contents are not *intentionally* represented by thoughts. While we might represent them on some loose sense of “represent”, our relation to them does not qualify as an instance of intentionality. If our paradigm cases of intentionality are all cases of phenomenal intentionality, it is quite likely that the representation of wide contents is of a different nature than our paradigm cases, and so does not qualify as a kind of intentionality. Again, whether a content qualifies as intentional turns largely on how we define “intentionality”.

Both derivativist and eliminativist descriptivism require that our thoughts represent narrow descriptive contents that determine the desired wide contents, but it is unclear that PIT can accommodate the required descriptive contents. The problem is that many descriptive contents would have to be quite nuanced and complex in order to fix on the desired wide contents, and it is not clear that such contents are phenomenal contents or somehow derived from phenomenal contents. For example, in order to fix on the natural kind *monkey*, we might need causal or metalinguistic descriptive contents like <the species around here that causes such-and-such superficial effects on ob-

---

son (1998). A view in a similar spirit is developed by Chalmers (2002a). Of course, descriptivism faces well-known objections (e.g., from Kripke 1980). We think these objections have been adequately addressed by descriptivists (see, e.g., Jackson 1998, 2003a,b, Chalmers 2002b, 2012), and focus here on objections to descriptivism that are special to PIT’s application of the view.

servers> or <the species called “monkey” around here>. But, since it does not seem that we have phenomenal states matching such narrow contents every time we think about monkeys, it is not clear how PIT can accommodate such descriptive contents. This issue for the descriptivist strategy is of a piece with the general problem of thoughts with abstract or complex contents, to which we now turn.

### 4.3 Complex and abstract thoughts

Many phenomenal intentionality theorists have applied an inflationist strategy in the case of thoughts with complex or abstract contents, arguing that they have a sufficiently rich phenomenology to account for their rich contents.<sup>36</sup> Proponents of rich cognitive phenomenology have attempted to bring out this phenomenology in various ways. One way is through the use of phenomenal contrast cases. For example, you might be asked to compare your phenomenology when hearing the words “birds fly” with that of someone who does not know what the word “bird” means. This person might have some auditory phenomenology corresponding to the word “bird”, but she seems to be missing something that you have. This something is the rich phenomenology of thought corresponding to the (perhaps narrow) idea of a bird.<sup>37</sup>

For our part, we are not entirely convinced of the inflationist strategy.

---

<sup>36</sup>See, e.g., Strawson (1994, 2008, 2011), Siewert (1998, 2011), Horgan and Tienson (2002a), Horgan et al. (2004), Chudnoff (2015), and Pitt (2009, 2011).

<sup>37</sup>For such arguments, see especially Strawson 1994, Siewert 1998, Horgan and Tienson 2002a, Chudnoff 2015, and Pitt 2009. Koksvik (2015) questions the methodology of phenomenal contrast arguments.

We agree that phenomenal contrast cases show that there is *something* in consciousness when one is thinking about monkeys, birds, or flying, but this something is not the full idea, even the full *narrow* idea, of a monkey, bird, or of flying. It seems to us that the contents that are determined by the phenomenology of conscious thoughts are gisty, partial, or schematic compared to the full narrow contents that we might want to attribute to these thoughts, which might include descriptive contents of the sort described above, or even just rough characterizations like <a winged feathery animal that lays eggs and flies>. Bourget (2017c, 2018) argues that the phenomenal contents associated with abstract or complex thoughts are often largely symbolic, representing words without precise meanings.

An alternative approach to complex thoughts is derivativist: Although the occurrent thoughts we typically have don't have phenomenal properties that capture the full narrow contents that we want to attribute to them, they are inferentially or otherwise connected with complex or abstract thoughts whose phenomenal properties determine these fuller contents. One might say that typical occurrent thoughts have the complex narrow contents they have in virtue of bearing such connections to more complex thoughts. There are different views on what the relevant connections are. We will focus on a view that takes the relevant connections to be determined by our dispositions to take ourselves to mean one content by another. According to this view, which we will call *derivativist self-ascriptivism*, we derivatively represent a content by having a disposition to ascribe it to ourselves or our own mental



states in certain circumstances. In order for such an approach to succeed, we must further say that the relevant contents of these self-ascriptions are phenomenally represented.<sup>38</sup>

To see how this view works, suppose, for example, that you are talking about physicalism with a colleague. In the course of this discussion, you say, “At least we can agree that phenomenal properties supervene on physical properties.” Suppose that your colleague asks what you mean by “supervene”. You might pause for a brief moment before producing an elucidation of supervenience such as this:

**Supervenience** Properties of class A supervene on properties of class B if and only if any possible worlds that are alike in their B properties are like in their A properties.

It seems clear that, prior to pausing and reflecting on the matter, you did not have an occurrent grasp of supervenience as that relation that Supervenience spells out. For instance, when you had the first thought about supervenience, you didn’t have in your consciousness anything like the idea of a possible world. Still, we are inclined to say that Supervenience was involved in the content of your thought because, on reflection, you ascribe

---

<sup>38</sup>This kind of self-ascriptivist strategy is developed in detail in Mendelovici 2018 (see also Mendelovici forthcoming), though it is given an eliminativist spin (see below). Bourget 2018 also defends a view along these lines. Pautz (2013) offers an alternative derivativist strategy for complex thoughts, the *consciousness-based best systems theory*, which allows states to derive intentionality from their functional relations with other states with phenomenal intentionality. Kriegel’s (2011a and 2011b) interpretivism might also be applied to the case of thought (see n. 31). See also Loar, 2003a,b and Bourget 2010a for other derivativist views of thought content.

this content to your thought. According to derivativist self-ascriptivism, such self-ascriptions have phenomenal intentionality and your thought about supervenience derivatively represents the full definition of supervenience in virtue of your disposition to have such phenomenally conscious self-ascriptive thoughts.

A few points of clarification are in order. First, in order to derivatively represent a particular content, you merely need to be *disposed* to form the relevant self-ascriptions; it does not matter whether you ever actually do. Second, the view can be refined by specifying the relevant dispositions' conditions of manifestation. The more idealized these conditions are, the less likely we are to ever have the relevant self-ascriptive thoughts, but the more likely these thoughts are to reflect our all-things-considered best understanding relating to our thoughts and concepts.<sup>39</sup> Third, any given self-ascriptive thought need only *partially* specify the full unpacking of any given thought. On the resulting view, thoughts with relatively impoverished phenomenal contents can manage to derivatively represent rich and complex contents, contents that we might never be able to entertain in a single conscious thought.

As in the case of dispositionalist and descriptivist strategies, there is

---

<sup>39</sup>Thanks to an anonymous reviewer, who worried that self-ascriptivism does not allow us to make mistaken judgments as to what we represent, for prompting this clarification. The worry is avoided because mistakes are possible outside of the relevant dispositions' conditions of manifestation. For example, we might say that the relevant circumstances are "good" ones in which subjects are awake, alert, and given ample opportunity to reflect. When subjects fail to meet these conditions, their self-ascriptive thoughts do not determine their derived contents and so they can be mistaken about those contents. See also Mendelovici 2010, §10.3.1 and 2018, §7.3 for further discussion.

an eliminativist version of the self-ascriptivist strategy, which is the view that we endorse. *Eliminativist self-ascriptivism* accepts the derivativist self-ascriptivist's story about dispositions to self-ascribe complex contents but denies that we *intentionally* represent these contents. To illustrate and motivate the eliminativist version of self-ascriptivism, note first that there are really two kinds of content at play in situations such as that of your discussion with your colleague above. When you first used the word "supervenience", you did not grasp the full meaning of this term for you as spelled out in Supervenience, but your mind was not completely empty. There was something before your mind, something that you grasped mentally as you were speaking. Plausibly, you had a gisty sense of what supervenience is. So it seems that your first, fleeting thought about Supervenience has two contents: a gisty content, which you initially grasped, and the full content spelled out by Supervenience, which you only grasped on reflection. We can call the first the *immediate content* of the thought, and the second its *reflective content*. Immediate and reflective contents can coincide, but they can also diverge, as seems to be the case in the present example. According to self-ascriptivism, the first kind of content is phenomenal content, while the second kind of content is derived content.

We defined "intentionality" by pointing to introspectively accessible paradigm cases. Now, it seems that introspection does not reveal anything about reflective contents (though reflection often does). So, our paradigm cases of intentionality are cases of the representation of immediate contents. In or-

der to count as intentional, reflective contents would have to be the same in nature as immediate contents. However, there are important differences between immediate contents and reflective contents.<sup>40</sup> For one, reflective contents arise from dispositional connections between mental states, whereas immediate contents arguably arise from the intrinsic phenomenal properties of thoughts. For these kinds of reasons, we don't take the representation of reflective contents such as Supervenience to be instances of intentionality. Our disposition to self-ascribe complex and abstract contents might create a vast illusion of our intentionally representing such contents, whereas in fact the *intentional* contents of thoughts are limited to their somewhat impoverished immediate contents. We do, however, have dispositions to self-ascribe more complex or abstract contents, but, so long as we are not occurrently entertaining them, we do not *intentionally* represent them.

Recall that the dispositionalist and descriptivist strategies mentioned above pass the buck to a theory of complex thought content. We are now in a position to see how the self-ascriptivist view of complex thought content can plug into and complete the dispositionalist and descriptivist views. The narrow descriptions required by the descriptivist strategy are a matter of our dispositions to self-ascribe descriptive contents to our occurrent thoughts.

---

<sup>40</sup>Mendelovici (2018, §7.5) argues that one of the most important differences between the having of immediate contents and the having of reflective contents is that the latter is relative to a self-ascriber, the subject who has the dispositions to ascribe the contents to herself, whereas the former is not. Bourget (2018) argues that phenomenal intentionality and non-phenomenal intentionality play different causal and rational roles, the former allowing for genuinely content-responsive thinking and the latter allowing only for an emulation of content-responsive thinking.

These descriptions, together with how the world is, fix the wide content of those thoughts. Standing propositional attitudes are a matter of dispositions to have occurrent thoughts with descriptive and/or wide contents. On this picture, one problematic kind of content or state is built up out of another, based on a foundation of dispositions to have self-ascriptions whose relevant contents are purely phenomenal. We call this picture the *scaffolding view*. One can have a derivativist, eliminativist, or mixed version of the scaffolding view, depending on whether one takes any of the relevant non-phenomenal contents to be genuinely intentionally represented or not. As we've already noted, we take each kind of non-phenomenal representation to be different in kind from genuine intentionality, so we prefer the eliminativist version of the scaffolding view to the alternative derivativist and mixed versions.<sup>41</sup>

While we find the scaffolding view attractive, we believe that self-ascriptivism alone can accommodate all the same sorts of states. Self-ascriptivism can directly account for wide intentional states and standing propositional attitudes, since, in both cases, we have dispositions to self-ascribe the relevant states or contents: In the case of the wide thought that water is wet, you are disposed to self-ascribe the content  $\langle \text{H}_2\text{O is wet} \rangle$  in that you are disposed to take yourself to be thinking that whatever happens to be the clear watery stuff around here (i.e., whatever is the referent of your water description) is wet.  $\text{H}_2\text{O}$  is what happens to be the clear watery stuff around here,

---

<sup>41</sup>For versions of the scaffolding view, see Horgan and Tienson 2002a, Bourget 2010a, and Mendelovici 2010.

so you are disposed to self-ascribe this content to yourself. In effect, we represent wide contents by being disposed to have self-ascriptions that *use* rather than mention narrow descriptive contents. Similarly, self-ascriptivism can accommodate standing propositional attitudes: we self-ascribe standing propositional attitudes in that we are disposed to take ourselves to have them.<sup>42,43</sup>

#### 4.4 Nonconscious representational states

Let us now turn to the case of nonconscious representational states posited by cognitive science. An inflationist strategy might claim that at least some such states are in fact phenomenally conscious and have phenomenal intentionality, even though we are not aware of this. Just as you are not aware of your neighbor's phenomenal states, your brain might house phenomenal states that you are not aware of. Whether these states are *your* states or the states of some other subject depends in part on how we understand subjects of experience, but it is irrelevant for the inflationist's main point, which is that the relevant states might very well have phenomenal intentionality that we are unaware of. While this might be plausible for some of the relevant nonconscious representational states (e.g., blindsight states), it is doubtful that all the relevant states involve hidden phenomenal characters. In general,

---

<sup>42</sup>See Mendelovici 2018, Chs. 8–9.

<sup>43</sup>The scaffolding view is compatible with self-ascriptivism across the board; there might be more than one way in which we come to have standing propositional attitudes or wide thoughts.

there are too many kinds of nonconscious representational states representing too many introspectively inaccessible contents for it to be plausible that they are all phenomenal.<sup>44</sup>

The derivativist strategy, which claims that the relevant states are derived from phenomenal intentional states (or from states that are eventually derived from phenomenal intentional states) might stand a better chance of accommodating all of the nonconscious representational states posited by cognitive science. For instance, Kriegel's interpretivism (2011a, 2011b) takes nonconscious intentionality to be derived from the phenomenal intentionality of an ideal interpreter who uses intentional systems theory to ascribe intentionality to nonconscious mental states (see also n. 31). Since this ideal interpreter is motivated by some of the same considerations as cognitive scientists, her content attributions are likely to match up with those of cognitive science.<sup>45</sup>

The main motivation for a derivativist strategy is a desire to be conciliatory with what we might take to be the standard view of the relevant nonconscious states. However, we prefer an eliminativist strategy, which we believe is at least as conciliatory with the standard view. The notions of representation operative in cognitive science are arguably either based on tracking or computational or other functional roles, or presumed to pick

---

<sup>44</sup>Pitt (2009), Bourget (2010a, 2017c), and Mendelovici (2018, Ch. 8) argue for an inflationist strategy along these lines for at least some cases.

<sup>45</sup>Bourget (2010a) suggests a derivativist strategy for certain nonconscious occurrent representational states. See also Horgan et al. (2004).

out something that is nothing over tracking relations and functional roles. Although we don't think tracking or functional roles can account for intentionality as we've defined it, we accept that internal states track things and have various functional roles, and that these are important features of these states that can serve many explanatory purposes. We also accept that there might be useful notions of representation that are based on such features. So we can agree with most of the claims characterizing the standard view. The only potential disagreement concerns whether the nonconscious representation posited by cognitive science is the same kind of thing as intentionality in our sense, which would require that it be the same kind of thing as the conscious intentionality we can introspect. In the previous section, we briefly overviewed reasons for thinking that the conscious intentionality we can introspect is not a matter of tracking or functional roles. If these arguments are sound, then intentionality (in our sense) is not the same kind of thing as the representation exhibited by the nonconscious representational states posited by cognitive science. The key point here is that our disagreement with the standard view concerns the nature of conscious intentionality, not the nature of the nonconscious representational states posited by cognitive science, making the eliminativist strategy quite conciliatory when it comes to the nature of the nonconscious representational states posited by cognitive science.<sup>46</sup>

---

<sup>46</sup>See also Horgan et al. 2004 and Mendelovici and Bourget 2014, Bourget and Mendelovici 2016, Bourget 2010a, 2017c, 2018, and Mendelovici 2018, Ch. 8 for arguments for the claim that the eliminativist strategy with respect to nonconscious occurrent representational states is in line with the standard view of such states. Mendelovici (2018, Ch. 8)



## 5 Concluding remarks

In this paper, we have outlined some possible views on the relationship between phenomenal consciousness and intentionality. Our focus has been on our preferred view, PIT, suggesting that one of the strongest arguments for PIT is based on the empirical inadequacy of its main competitors. We have argued that PIT can avoid the problems facing its competitors, but it too faces some challenges. We have considered four central kinds of challenging cases for PIT and three strategies that can be applied to each case. For each kind of challenging case, there are several attractive options, yielding a plethora of plausible versions of PIT. We have argued for a largely eliminativist position in all cases, which results in a version of strong PIT.

## Acknowledgements

Thanks to Uriah Kriegel and two anonymous reviewers for this volume for very helpful comments on earlier drafts of this paper.

## References

Bain, D. (2003). Intentionalism and pain. *Philosophical Quarterly*, 53(213):502–523.

Block, N. (1990). Inverted earth. *Philosophical Perspectives*, 4(n/a):53–79.

---

also argues that derivativism is less in line with the standard view than eliminativism, and that this may be a reason to prefer it.

- Block, N. (1996). Mental paint and mental latex. *Philosophical Issues*, 7:19–49.
- BonJour, L. (1998). *In Defense of Pure Reason*. Cambridge University Press.
- Bourget, D. (2010a). Consciousness is underived intentionality. *Noûs*, 44(1):32–58.
- Bourget, D. (2010b). *The Representational Theory of Consciousness*. PhD thesis, Australian National University.
- Bourget, D. (2015). Representationalism, perceptual distortion and the limits of phenomenal concepts. *Canadian Journal of Philosophy*, 45(1):16–36.
- Bourget, D. (2017a). Intensional perceptual ascriptions. *Erkenntnis*, 82(3):513–530.
- Bourget, D. (2017b). Representationalism and sensory modalities: an argument for intermodal representationalism. *American Philosophical Quarterly*, 53:251–267.
- Bourget, D. (2017c). The role of consciousness in grasping and understanding. *Philosophy and Phenomenological Research*, 95(2):285–318.
- Bourget, D. (2017d). Why are some phenomenal experiences 'vivid' and others 'faint'? representationalism, imagery, and cognitive phenomenology. *Australasian Journal of Philosophy*, 95(4):673–687.
- Bourget, D. (2018). The rational role of experience. *Inquiry*, 61(5-6):475–493.
- Bourget, D. (forthcoming). Implications of intensional perceptual ascriptions for relationalism, disjunctivism, and representationalism about perceptual experience. *Erkenntnis*.
- Bourget, D. (MS). Phenomenal intentionality, underdetermination, and functional role.
- Bourget, D. and Mendelovici, A. (2014). Tracking representationalism. In Bailey, A., editor, *Philosophy of Mind: The Key Thinkers*, pages 209–235. Continuum.

- Bourget, D. and Mendelovici, A. (2016). Phenomenal intentionality. In *Stanford Encyclopedia of Philosophy*.
- Byrne, A. (2001). Intentionalism defended. *Philosophical Review*, 110(2):199–240.
- Chalmers, D. (2012). *Constructing the World*. Oxford University Press.
- Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chalmers, D. J. (2002a). The components of content. In Chalmers, D. J., editor, *Philosophy of Mind: Classical and Contemporary Readings*. Oxford University Press.
- Chalmers, D. J. (2002b). On sense and intension. *Philosophical Perspectives*, 16(s16):135–82.
- Chalmers, D. J. (2004). The representational character of experience. In Leiter, B., editor, *The Future for Philosophy*, pages 153–181. Oxford University Press.
- Chalmers, D. J. (2010). *The Character of Consciousness*. Oxford University Press.
- Chudnoff, E. (2015). *Cognitive Phenomenology*. Routledge.
- Crane, T. (2003). The intentional structure of consciousness. In Smith, Q. and Jokic, A., editors, *Consciousness: New Philosophical Perspectives*, pages 33–56. Oxford University Press.
- Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. MIT Press.
- Dretske, F. (1993). The nature of thought. *Philosophical Studies*, 70(2):185–99.
- Dretske, F. (1995). *Naturalizing the Mind*. MIT Press.
- Farkas, K. (2008a). Phenomenal intentionality without compromise. *The Monist*, 91(2):273–93.

- Farkas, K. (2008b). *The Subject's Point of View*. Oxford University Press.
- Farkas, K. (2013). Constructing a world for the senses. In Kriegel, U., editor, *Phenomenal Intentionality*, page 99. OUP.
- Fine, K. (1977). Logic, meaning, and conceptual role. *Journal of Philosophy*, 74(7):378–409.
- Fodor, J. A. (1975). *The Language of Thought*. Harvard University Press.
- Fodor, J. A. (1990a). A theory of content i. In Fodor, J. A., editor, *A Theory of Content*. MIT Press.
- Fodor, J. A. (1990b). A theory of content ii. In Fodor, J. A., editor, *A Theory of Content*. MIT Press.
- Graham, G., Horgan, T. E., and Tienson, J. L. (2007). Consciousness and intentionality. In Velmans, M. and Schneider, S., editors, *The Blackwell Companion to Consciousness*, pages 468–484. Blackwell.
- Harman, G. (1987). (nonsolipsistic) conceptual role semantics. In LePore, E., editor, *Notre Dame Journal of Formal Logic*, pages 242–256. Academic Press.
- Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives*, 4(n/a):31–52.
- Horgan, T. and Tienson, J. (2002a). The intentionality of phenomenology and the phenomenology of intentionality. In Chalmers, D. J., editor, *Philosophy of Mind: Classical and Contemporary Readings*, pages 520–533. OUP Usa.
- Horgan, T. and Tienson, J. (2002b). The phenomenology of intentionality and the intentionality of phenomenology. In Chalmers, D. J., editor, *Philosophy of Mind: Classical and Contemporary Readings*, pages 520–533. Oxford University Press.
- 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*. Walter De Gruyter.
- Jackson, F. (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford University Press.

- Jackson, F. (2003a). Narrow content and representation—or twin earth revisited. *Proceedings and Addresses of the American Philosophical Association*, 77(2):55–70.
- Jackson, F. (2003b). Representation and narrow belief. *Philosophical Issues*, 13(1):99–112.
- Jorba, M. (2016). Attitudinal cognitive phenomenology and the horizon of possibilities. In Gutland, T. B. C., editor, *The Phenomenology of Thinking. Philosophical Investigations into the Character of Cognitive Experiences.*, pages 77–96. Routledge.
- Kind, A. (2003). What’s so transparent about transparency? *Philosophical Studies*, 115(3):225–244.
- Kind, A. (2013). The case against representationalism about moods. In Kriegel, U., editor, *Current Controversies in Philosophy of Mind*.
- Koksvik, O. (2015). Phenomenal contrast: A critique. *American Philosophical Quarterly*, 52:321–334.
- Kriegel, U. (2003). Is intentionality dependent upon consciousness? *Philosophical Studies*, 116(3):271–307.
- Kriegel, U. (2007). Intentional inexistence and phenomenal intentionality. *Philosophical Perspectives*, 21(1):307–340.
- Kriegel, U. (2011a). Cognitive phenomenology as the basis of unconscious content. In Bayne, T. and Montague, M., editors, *Cognitive Phenomenology*, pages 79–102. Oxford University Press.
- Kriegel, U. (2011b). *The Sources of Intentionality*. Oxford University Press.
- Kriegel, U. (2013). The phenomenal intentionality research program. In Kriegel, U., editor, *Phenomenal Intentionality*. Oxford University Press.
- Kripke, S. A. (1980). *Naming and Necessity*. Harvard University Press.
- Kripke, S. A. (1982). *Wittgenstein on Rules and Private Language*. Harvard University Press.

- Loar, B. (2003a). Phenomenal intentionality as the basis of mental content. In Hahn, M. and Ramberg, B., editors, *Reflections and Replies: Essays on the Philosophy of Tyler Burge*, pages 229–258. MIT Press.
- Loar, B. (2003b). Transparent experience and the availability of qualia. In Smith, Q. and Jovic, A., editors, *Consciousness: New Philosophical Perspectives*. Oxford University Press.
- Lycan, W. G. (1987). *Consciousness*. MIT Press.
- Lycan, W. G. (2000). Representational theories of consciousness. In Zalta, E. N., editor, *Stanford Encyclopedia of Philosophy*, pages 66–69. Stanford: Metaphysics Research Lab.
- Masrour, F. (2013). Phenomenal objectivity and phenomenal intentionality: In defense of a kantian account. In Kriegel, U., editor, *Phenomenal Intentionality*, page 116. OUP.
- Mendelovici, A. (2010). *Mental Representation and Closely Conflated Topics*. PhD thesis, Princeton University.
- Mendelovici, A. (2013a). Intentionalism about moods. *Thought: A Journal of Philosophy*, 2(1):126–136.
- Mendelovici, A. (2013b). Reliable misrepresentation and tracking theories of mental representation. *Philosophical Studies*, 165(2):421–443.
- Mendelovici, A. (2014). Pure intentionalism about moods and emotions. In Kriegel, U., editor, *Current Controversies in Philosophy of Mind*, pages 135–157. Routledge.
- Mendelovici, A. (2016). Why tracking theories should allow for clean cases of reliable misrepresentation. *Disputatio*, 8(42):57–92.
- Mendelovici, A. (2018). *The Phenomenal Basis of Intentionality*. Oxford University Press.
- Mendelovici, A. (forthcoming). Immediate and reflective senses. In Gouveia, S., Curado, M., and Shottenkirk, D., editors, *Perception, Cognition, and Aesthetics*. New York: Routledge.

- Mendelovici, A. and Bourget, D. (2014). Naturalizing intentionality: Tracking theories versus phenomenal intentionality theories. *Philosophy Compass*, 9(5):325–337.
- Millikan, R. G. (1984). *Language, Thought and Other Biological Categories*. MIT Press.
- Montague, M. (2016). *The Given: Experience and its Content*. Oxford University Press Uk.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83(October):435–50.
- Pautz, A. (2013). Does phenomenology ground mental content? In Kriegel, U., editor, *Phenomenal Intentionality*, pages 194–234. Oxford.
- Pitt, D. (1999). In defense of definitions. *Philosophical Psychology*, 12(2):139–156.
- Pitt, D. (2004). The phenomenology of cognition, or, what is it like to think that *P*? *Philosophy and Phenomenological Research*, 69(1):1–36.
- Pitt, D. (2009). Intentional psychologism. *Philosophical Studies*, 146(1):117–138.
- Pitt, D. (2011). Introspection, phenomenality, and the availability of intentional content. In Bayne, T. and Montague, M., editors, *Cognitive Phenomenology*, page 141. OUP.
- Putnam, H. (1975). The meaning of 'meaning'. *Minnesota Studies in the Philosophy of Science*, 7:131–193.
- Putnam, H. (1981). *Reason, Truth, and History*. Cambridge University Press.
- Seager, W. E. and Bourget, D. (2007). Representationalism about consciousness. In Velmans, M. and Schneider, S., editors, *The Blackwell Companion to Consciousness*, pages 261–276. Blackwell.
- Searle, J. R. (1989). Consciousness, unconsciousness, and intentionality. *Philosophical Topics*, 17(1):193–209.

- Searle, J. R. (1990). Consciousness, explanatory inversion and cognitive science. *Behavioral and Brain Sciences*, 13(1):585–642.
- Searle, J. R. (1992). *The Rediscovery of the Mind*. MIT Press.
- Siewert, C. (1998). *The Significance of Consciousness*. Princeton University Press.
- Siewert, C. (2011). Phenomenal thought. In Bayne, T. and Montague, M., editors, *Cognitive Phenomenology*, page 236. Oxford University Press.
- Smithies, D. (2011). What is the role of consciousness in demonstrative thought? *Journal of Philosophy*, 108(1):5–34.
- Smithies, D. (2013a). The nature of cognitive phenomenology. *Philosophy Compass*, 8(8):744–754.
- Smithies, D. (2013b). The significance of cognitive phenomenology. *Philosophy Compass*, 8(8):731–743.
- Smithies, D. (2014). The phenomenal basis of epistemic justification. In Kallestrup, J. and Sprevak, M., editors, *New Waves in Philosophy of Mind*, pages 98–124. Palgrave MacMillan.
- Smithies, D. (forthcoming). *The Epistemic Role of Consciousness*. Oxford University Press.
- Strawson, G. (1994). *Mental Reality*. MIT Press.
- Strawson, G. (2008). Real intentionality 3: Why intentionality entails consciousness. In Strawson, G., editor, *Synthesis Philosophica*, pages 279–297. Oxford University Press.
- Strawson, G. (2011). Cognitive phenomenology: Real life. In Bayne, T. and Montague, M., editors, *Cognitive phenomenology*, pages 285–325. Oxford University Press.
- Tye, M. (1995). A representational theory of pains and their phenomenal character. *Philosophical Perspectives*, 9:223–39.
- Tye, M. (2008). The experience of emotion: An intentionalist theory. *Revue Internationale de Philosophie*, 62:25–50.