

The Phenomenal Basis of Intentionality

Angela Mendelovici

Penultimate Draft

Forthcoming with OUP in May 2018

Contents

Preface	v
Overview	viii
I Introduction	1
1 Fixing Reference on Intentionality	3
1.1 Aboutness and directedness	4
1.2 The ostensive way of fixing reference	5
1.3 Other ways of fixing reference	10
1.4 Worries with the ostensive definition	18
1.5 Conclusion	24
2 Goals and Methodology	26
2.1 What is a theory of intentionality?	26
2.2 Theory-independent access to intentionality	29
2.3 Conclusion	36
II Alternative Theories of Intentionality	37
3 The Mismatch Problem for Tracking Theories	39
3.1 Tracking theories	39

3.2	Overview of the mismatch problem for tracking theories	42
3.3	Background and assumptions	43
3.4	A mismatch case: Perceptual color representations	46
3.5	Other mismatch cases	53
3.6	Objections	56
3.7	Reliable misrepresentation and the significance of tracking . .	70
3.8	Conclusion	73
	Appendix A: Objections to the mismatch problem	74
4	Functional Role Theories and Tracking Theories Again	88
4.1	The functional role theory	89
4.2	Worries with short-arm functional role theories	91
4.3	Worries with long-arm functional role theories	96
4.4	The Real Problem with both tracking theories and functional role theories	100
4.5	Conclusion	102
III	The Phenomenal Intentionality Theory	103
5	The Phenomenal Intentionality Theory	105
5.1	The phenomenal intentionality theory	106
5.2	Arguments for PIT	110
5.3	Identity PIT	119
5.4	Challenging cases	123
5.5	Conclusion	128
	Appendix B: The extent of phenomenal intentionality	129
	Appendix C: The multiple arisability of intentional states	133
6	PIT's Status as a Theory of Intentionality	140
6.1	Is PIT a theory of intentionality in terms of phenomenal con- sciousness?	141

6.2	Is PIT trivial?	146
6.3	Is PIT interesting?	148
6.4	Is PIT naturalistic?	150
6.5	Conclusion	154
IV	Challenging Cases	155
7	Thought	157
7.1	The challenge for PIT from thought	157
7.2	Thoughts' phenomenal contents	162
7.3	Self-ascriptivism about thoughts' alleged contents	178
7.4	Is derived mental representation a type of intentionality? . . .	195
7.5	Conclusion: PIT about thought	197
	Appendix D: Derived mental representation in perception	198
	Appendix E: Attitudes	201
8	Nonconscious States	206
8.1	The problem with nonconscious states	207
8.2	Derivativist strategies	209
8.3	Standing states	218
8.4	Allegedly nonconscious occurrent states	238
8.5	Conclusion	247
V	The Aspect View	249
9	Is Intentionality a Relation to a Content?	251
9.1	The relation view and the aspect view	253
9.2	Two worries with the relation view	257
9.3	The alleged virtues of the relation view	266
9.4	Conclusion	295

Appendix F: The aspect view and adverbialism	297
Appendix G: Contents as first- or second-order states or properties	300
Appendix H: An internal theory of truth and reference	304
 VI Conclusion	 310
 10 Conclusion: Intentionality and Other Related Phenomena	 311
10.1 Return to other ways of fixing reference on intentionality . . .	313
10.2 Radical internalism	317
 Glossary	 319
 Bibliography	 333

Overview

The aim of this book is to defend a radically internalist theory of intentionality, the aboutness or directedness of mental states, on which intentionality is simply identical to phenomenal consciousness, which is an intrinsic, non-relational feature of mental life. This view has been described to me as obviously false, unfashionable, and flying in the face of everyday intuition and cognitive science. It has also been described to me as trivially true and uninteresting. I aim to defend a version of this view that is true but not trivial, interesting but not false, and surprisingly conciliatory with our intuitive and scientific understanding of the mind.

My target, intentionality, can be understood as the observed “aboutness” or “directedness” of mental states. We introspectively notice that many mental states in some way or other seem to “present,” “represent,” or be “about” things. For example, you might notice that your current visual experiences represent a page before you, some marks of various shapes and colors, and perhaps the words that these marks form. You might also notice that your current thoughts represent that there is a page with marks and words before you, something to do with your own mental states, or a need for a cup of coffee. Intentionality, roughly, is this phenomenon of aboutness or directedness that we notice in these and other everyday cases.

My aim is to offer a theory of intentionality, a theory that describes the deep nature of intentionality, or, in other words, that tells us what intentionality really *is*, metaphysically speaking. Examples of theories of intentionality

include tracking theories, on which the most basic kind of intentionality is a causal or other kind of tracking relation between internal representations and items in the world (see, e.g., Dretske 1986 and Fodor 1987), and functional role theories, on which the most basic kind of intentionality is a matter of internal states' functional dispositions with respect to other internal states and perhaps also with respect to items in the environment (see, e.g., Harman 1987 and Block 1986).

This book proposes a very different kind of theory of intentionality, the phenomenal intentionality theory (PIT), which takes the most basic kind of intentionality to arise from a conceptually distinct mental feature, phenomenal consciousness, the felt, subjective, or “what it’s like” (Nagel 1974) aspect of mental life. This and related views have recently been defended by various authors, including Horgan and Tienson (2002), Loar (2003), Farkas (2008b, 2008a), Strawson (2008a), Siewert (1998), Montague (2010), Bourget (2010a), Mendelovici (2010), Kriegel (2011), Pitt (2004, 2009), Pautz (2013a), and Mendelovici and Bourget (2014), and have historical roots in the works of Brentano (1874) and Husserl (1900). This book proposes a version of PIT that is not only motivated on in-principle grounds but also empirically adequate in that it can accommodate all cases of intentionality, including those that are commonly thought to pose problems for PIT.

I proceed as follows: Chapter 1 of Part I fixes reference on our target, intentionality. I argue that while the notions of aboutness and directedness gesture toward this target, they are too fuzzy to provide us with a firm grip on it. I propose to replace these notions with an ostensive reference-fixing definition, which can be contrasted with other candidate definitions that take intentionality to be whatever plays certain roles, such as roles in folk psychological or scientific theories of behavior, roles in securing truth and reference, or simply roles in an explanation of how we get around in the world. On my approach, intentionality is a phenomenon we observe and want to explain, rather than a posit in a theory primarily aimed at explaining

something else.

Chapter 2 of Part I specifies the kind of theory of intentionality we are after and describes two theory-independent ways of knowing about our intentional states: introspection and consideration of psychological role.

Part II considers and argues against what I take to be the two main competitors to my favored approach to intentionality, tracking and functional role theories. Chapter 3 of Part II argues that tracking theories face a mismatch problem: there are cases in which we represent a content that does not match anything we can plausibly be said to track. The tracking theory, then, is empirically inadequate, since it cannot accommodate all the required cases. Chapter 4 of Part II argues that the mismatch problem also afflicts the best versions of the functional role theory. Now, while the mismatch problem shows *that* the tracking theory and the best versions of the functional role theory are false, it does not pinpoint the precise reasons for their failure. Chapter 4 further argues that the fundamental problem with the theories is that tracking relations and functional roles simply do not have what it takes to give rise to intentionality.

Part III turns to my favored approach to intentionality, the phenomenal intentionality theory (PIT), on which the most basic kind of intentionality arises from phenomenal consciousness. Chapter 5 of Part III presents and motivates PIT. I argue that, unlike tracking theories and functional role theories, PIT provides the right kinds of ingredients to account for intentionality and is not clearly empirically inadequate. I distinguish between different versions of PIT, focusing especially on my favored version, strong identity PIT, which, roughly, takes every intentional property to be identical to some phenomenal property. Chapter 6 of Part III considers and responds to some theoretical worries with PIT, such as that it is not naturalistic.

Part IV further supports PIT by considering certain challenging cases for the view. In doing so, it fleshes out my favored version of strong identity PIT and shows that it is both interesting and tenable. Chapter 7 of Part IV

considers the challenge raised by the case of thoughts, which appear to be rich in intentional content but poor in phenomenal character. I argue that thoughts have a kind of content that does indeed arise from their fairly impoverished phenomenal characters, though this content is correspondingly impoverished. I further argue that, although thoughts do not phenomenally represent many of their alleged contents, they do the next best thing: they *derivatively* represent them. I propose self-ascriptivism, a view on which we derivatively represent various contents by ascribing them to ourselves, which is a matter of being disposed to have thoughts accepting ourselves or our phenomenal contents as representing these further contents. Although, as I argue, the resulting kind of derived mental representation does not qualify as a kind of *intentionality*, it qualifies as a kind of representation on a broad sense of the term.

Another important challenge for PIT is that of accounting for mental states that we take to be intentional but that appear to have no phenomenal character. Such states include standing states, like beliefs and desires that we are not currently entertaining, as well as occurrent states that we are simply not aware of, such as nonconscious states involved in language processing, blindsight, and early visual processing. Chapter 8 of Part IV addresses these challenges. I argue that standing states are not genuinely intentional states. However, I also suggest that self-ascriptivism can be extended to accommodate standing state contents and perhaps even standing states in their entirety.

Chapter 8 also argues that many nonconscious occurrent states, such as states involved in early visual processing, are neither intentional nor derivatively representational. While this position might seem fairly extreme, even “flying in the face of cognitive science,” it is arguably very much in line with the standard view on the matter. It agrees with the standard view that such occurrent states track or carry information about various items in the environment and play various functional roles, and it also agrees that they represent various items, if all we mean by “representation” is something

that boils down to tracking, carrying information, or having a functional role. The key disagreement with the standard view does not concern nonconscious occurrent states, but rather *conscious* occurrent states.

Part V, which consists in only one chapter, Chapter 9, turns to the question of whether intentionality is a relation to distinctly existing entities that play the role of contents or whether, instead, intentionality is simply an aspect of intentional states or subjects. I argue in favor of the latter **aspect view** of intentionality. While it might be thought that the alternative relation view has various virtues that the aspect view lacks, such as according with common sense, allowing for public contents, helping us make sense of structured intentional states, and accounting for conditions of truth and reference, I argue that the aspect view fares no worse than the relation view when it comes to these virtues.

The main goal of this book is to offer, flesh out, and defend a theory of intentionality, but it also has a secondary aim. As I mentioned above, Chapter 1 will argue that it is possible to get a good grip on the phenomenon of intentionality without defining it in terms of truth and reference, our abilities to get around in the world, folk psychology, or the scientific study of the mind. Throughout this book, I return to these alleged roles of intentionality and argue that it turns out that most of them are not in fact played by intentionality itself but by various closely related phenomena: The relevant ability to get around in the world is explained by a combination of intentionality, tracking, and functional roles; the notions of representation implicit in folk psychology don't correspond to intentionality but to some combination of intentionality and derived representation; conditions of truth and reference might end up requiring something more than merely having intentional states, like a primitive correspondence relation or our specifications of how we'd like to be interpreted; and the notions of representation invoked in the mind-brain sciences are often a matter of tracking relations and functional roles. The concluding chapter, Chapter 10, returns to the alleged roles of

intentionality and summarizes these findings. The end result is a picture on which intentionality, as picked out ostensively, is a matter of phenomenal consciousness, and the various other roles intentionality is sometimes thought to play are in fact often played by distinct, although sometimes closely related, phenomena.

Part I

Introduction

My aim is to provide a theory of intentionality. Before comparing competing theories of intentionality, it is important to fix firmly on our target and to get clear on what kind of theory we are after. Chapter 1 proposes an ostensive way of fixing reference on intentionality, while Chapter 2 specifies what kind of theory we want and overviews two sources of theory-independent knowledge of intentionality that we can use to test our theories: introspection and considerations of psychological role.

Chapter 1

Fixing Reference on Intentionality

The aim of this book is to provide a theory of intentionality. The aim of this chapter is to clarify just what a theory of intentionality is a theory *of*. It is important to get clear on this before we start. A theory of intentionality is a theory that tells us that intentionality has a particular nature, but if it is unclear just what “intentionality” refers to, then it is unclear what it is that such a theory says has that nature.

I propose to get clear on our target by defining it ostensively using introspectively accessible paradigm cases. My ostensive definition can be contrasted with alternative definitions that may or may not end up picking out the same thing. I will suggest that the ostensive definition does a better job of capturing the core notion we are interested in. But first, I will say something about why common characterizations of intentionality in terms of “aboutness” and “directedness,” though they succeed in gesturing toward our target, do not provide a satisfactory way of fixing firmly upon it.

1.1 Aboutness and directedness

Intentionality is sometimes characterized, at least as a first pass, as the “aboutness” or “directedness” of mental states (and perhaps other items) to things that may or may not exist. We might say that a perceptual experience of a cup is “directed” at a cup, that a thought that it is raining is “about” the putative fact that it is raining, and that a belief in Santa Claus is “about” Santa Claus or the putative fact that Santa Claus exists.

This characterization of intentionality has roots in an oft-cited passage from Brentano, who is often credited with introducing the notion of intentionality to contemporary discussions:

Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence 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. Every mental phenomenon includes something as object within itself, although they do not do so in the same way. In presentation, something is presented, in judgment something is affirmed or denied, in love loved, in hate hated, in desire desired and so on. (Brentano 1874, p. 88)

There are many examples of contemporary characterizations of intentionality in terms of “aboutness” and “directedness.” For instance, Siewert (2006) writes: “Intentionality has to do with the directedness or aboutness of mental states—the fact that, for example, one’s thinking is *of* or *about* something” (p. 1). Similarly, Speaks (2010b) writes: “The closest thing to a synonym for intentionality is ‘aboutness’; something exhibits intentionality if and only if it is about something” (p. 398).¹

¹For additional representative examples of this way of characterizing intentionality, see Jacob 2003, Byrne 2006, Kim 1998, p. 21, Searle 2004, p. 112, and O’Madagain 2014.

In light of the widespread acceptance of such characterizations of intentionality in terms of aboutness and related notions, I will take it as given that such characterizations at least *gesture* toward the phenomenon of interest. However, despite this, the characterization of intentionality in terms of aboutness or directedness would not make a good definition, not even a good reference-fixing definition, of “intentionality.” As a definition of “intentionality,” it is too fuzzy and metaphorical to give us a firm grip on our target. It is simply not clear what is being said when we say that a mental state is “directed at” or “about” something, especially if this thing need not exist. An experience of a cup is not literally pointed in the direction of a cup (which may not even exist), in the way that a finger or an arrow might point to a cup, and a thought is not literally pointed in the direction of a proposition, which might be an abstract entity having no spatial location at all. If we take “aboutness” and “directedness” talk to supply a definition of “intentionality,” it is simply not clear what this definition says.²

1.2 The ostensive way of fixing reference

Although “aboutness” and “directedness” talk do not provide us with a satisfactory definition of “intentionality,” they do gesture toward the phenomenon of interest. I want to suggest that what is doing the work when we use “aboutness” and “directedness” talk to fix on intentionality is a prior grasp we have on the phenomenon. My suggestion for defining “intentionality,” then, is to look past our descriptions of this phenomenon in terms of aboutness and related notions and focus instead on the phenomenon thus described. This is possible because we have a special access to this mental feature independent of any fuzzy or metaphorical descriptions: We can directly notice it through introspection, at least in some cases. This allows us to ostensively define

²Chisholm (1957a) criticizes Brentano’s definition as being too fuzzy and suggests instead a linguistic criterion of intentionality. Speaks (2010b) also argues that we should not *define* “intentionality” in terms of aboutness; see n. 5 for discussion.

“intentionality” as this feature, whatever it is, that we at least sometimes notice in ourselves and are tempted to describe using representational terms like “aboutness” and “directedness.”

In order to flesh out this suggestion, let us begin by considering some cases of mental states that we are tempted to describe using representational terms like “aboutness” and “directedness.” Take your present perceptual experiences: You might be visually experiencing some marks on a page, pens on your desk, or parts of your body. Likewise, you might be enjoying auditory experiences of voices, music, or various noises. These experiences have a certain feature, a feature we are tempted to describe using representational terms like “aboutness,” “directedness,” “ofness,” or “saying something.” We might describe these experiences as being “of” or “about” things or ways things are or might be, or as “saying” that things are a particular way. We might say they are “about” some marks on a page, that they “say” that these marks are in front of you, and so on.

Now consider the thoughts you are currently having. You might be thinking about your experiences, desiring another cup of coffee, or judging that I am pointing out the obvious. Like perceptual experiences, these thoughts have a feature that it is tempting to describe using representational terms. We might describe these thoughts as being “about” things or as “saying” that things are a certain way. We might say that they are “about” our experiences, that they “say” that I am pointing out the obvious, and so on.³

The above examples show that we have mental states that have a certain feature that we at least sometimes introspectively notice and are tempted to describe using representational terms, such as “about,” “of,” “represent,” “present,” and “saying something.” *That* feature, whatever it is, is intentionality.

We can put things more precisely as follows: Call the mundane, everyday

³I take the category of thoughts to include occurrent beliefs, occurrent desires, and other occurrent “cognitive” states but not standing beliefs, standing desires, or other standing states. See §1.4.2.

cases such as those described above our **paradigm cases** of intentionality. These are the cases that will form our initial sample of cases of intentionality for the purposes of our ostensive definition. Then, we can define “intentionality” as follows:

Intentionality The feature that in paradigm cases we sometimes both (i) notice introspectively in ourselves and (ii) are tempted to describe using representational terms, such as “about,” “of,” “represent,” “present,” or “saying something.”⁴

It is important to emphasize that the feature picked out by my definition is the feature of paradigm cases that we at least sometimes *both* introspectively notice *and* are tempted to describe representationally. This allows that there are features of paradigm cases that we *either* introspectively notice *or* are tempted to describe representationally, but not both, and that do not qualify as intentional. For example, the definition does not by itself rule out the view that paradigm cases have introspectively accessible phenomenal features that are distinct from intentionality.

It is also important to emphasize that, although we are using introspection to fix reference on intentionality, the ostensive definition does not rule out the possibility of instances of intentionality that are not introspectively accessible, or even instances of intentionality that are not mental. Such cases would not be paradigm cases of intentionality, but they would nonetheless be cases of intentionality so long as they had the relevant feature exemplified by paradigm cases. For example, as far as my definition is concerned, it could turn out that nonconscious beliefs and the nonconscious states posited by cognitive science, which, presumably, are not introspectively accessible, are instances of intentionality. For the same reasons, the ostensive definition does not rule out the possibility of instances of intentionality that we are not tempted to describe representationally. For example, it does not rule out the possibility

⁴Definitions of key terms can be found in the glossary on p. 319.

of moods and afterimages being instances of intentionality, even though we (arguably) are not tempted to describe them representationally.

The ostensive definition arguably both does justice to the intuition behind the characterization of intentionality in terms of “aboutness” and “directedness” and is an improvement over a definition of “intentionality” in terms of this characterization. If I am right, “aboutness” talk aims to characterize a phenomenon that we have an antecedent grasp on. My ostensive definition picks out precisely that phenomenon, so it does justice to the intuition behind characterizations of intentionality in terms of “aboutness” and “directedness.” It offers an improvement over a definition of “intentionality” in terms of such characterizations, since it fixes firmly on our target. Unlike a definition of “intentionality” simply as aboutness or directedness, it avoids being fuzzy or metaphorical, since it merely *mentions* our fuzzy and metaphorical representational terms rather than *use* them. (Of course, it uses the term “representational term,” but this is a term picking out a class of terms rather than a representational term itself.)⁵

We can use this ostensive definition of “intentionality” to define some related notions: **Intentional properties** are ways things are or might be with respect to their intentionality, or intentional ways things are or might be, and **intentional states** are instantiations of intentional properties. As I am using the terms, intentional states are not the same thing as **intentional mental**

⁵One might object that “aboutness” talk gestures at reference rather than at the ostensively defined phenomenon (but see Crane 2013, pp. 8–9 for a convincing argument against this). After all, one might argue, we sometimes say that mental states that fail to refer, like a thought that Santa Claus exists, are not really about anything at all.

Now, we might agree that “aboutness” talk is sometimes used to pick out reference but disagree that this means that “aboutness” talk, *when used to characterize intentionality*, gestures at reference, since whatever “aboutness” talk is supposed to gesture at is normally taken to include mental states in which there is a failure of reference, such as the thought that Santa Claus exists. In any case, if “aboutness” talk is normally used to pick out reference, this only further supports my claim that characterizations of intentionality in terms of “aboutness” would not provide an adequate definition of intentionality. Speaks (2010b) makes a similar point, arguing that for these reasons the “characterization of intentionality as aboutness is only true to a first approximation” (p. 398).

states, which are mental states that include, but may not be exhausted by, the instantiation of intentional properties. For example, a judgment that grass is green might involve the instantiation of the intentional property of representing that grass is green together with a particular non-intentional “judgment” component. So, it is an intentional mental state but not an intentional state.⁶

What intentional properties and intentional states “say” or are “directed at” are their intentional contents. More precisely, we can think of intentional content as follows: When we introspectively notice intentional states, we notice the general phenomenon that we are tempted to describe as “directedness” or “saying something.” But we also notice something we are tempted to describe as *what* our mental states are “directed at” or what they “say”; this is their **(intentional) content**.⁷ When a state, property, or other item has a certain intentional content, we can say that it **(intentionally) represents** that content.⁸ For example, the judgment that grass is green represents the content <grass is green>.⁹

It is worth emphasizing that my starting point is fairly noncommittal in that my definition of “intentionality” and the introspective observations it is based on do not prejudge questions concerning the nature of intentionality.

⁶The term “intentional state” is often used to mean what I mean by “intentional mental state.” I deviate from this usage because my discussion focuses on instantiations of intentional properties, so it is useful for me to reserve the term “intentional state” for them.

⁷When we introspectively notice intentionality, we do so at least in part by introspectively noticing our contents. Indeed, it might be that there is nothing more to notice when we notice intentionality than these intentional contents.

⁸I sometimes use “represent” more broadly to describe representation-like phenomena that are not instances of intentionality, but context should disambiguate. The alternative would be to use a special term, like “intend,” for having an intentional content, but this would be too awkward.

⁹Contents might include propositional contents, like <grass is green>, but might also include proprietal or objectual contents, like <green> and <George> (see Montague 2007, Grzankowski 2013, and Mendelovici forthcoming c, MS). (The notions of objectual and proprietal contents are something like Crane’s [2013] notions of contents and objects, respectively, though not equivalent.)

As far as they are concerned, intentionality might end up being a causal or other tracking relation, a matter of the functional roles of internal states, or a matter of phenomenal consciousness. My starting point is neutral on these and other possible views of intentionality. Likewise, my starting point does not prejudge questions concerning the nature of contents. Contents might turn out to be ordinary objects and properties, propositions, facts, sense data, ideas in a world of forms, ways of representing, properties of intentional states, or even intentional states or properties themselves.

Relatedly, my starting point does not prejudge any issues regarding the **vehicles of intentionality**, which are the bearers of intentional properties. The vehicles of intentionality could turn out to be, for example, subjects, symbols in a language of thought, brain states, internal states, or immaterial souls.

For simplicity, however, I will assume that the vehicles of intentionality are internal items that I will call **(mental) representations**. Since different intentional states involve different vehicles of representation, this way of speaking allows us to talk about intentional states while remaining noncommittal on their contents, which is useful when the content of a particular intentional state is under dispute.¹⁰

1.3 Other ways of fixing reference

I have recommended an ostensive way of fixing firmly upon the phenomenon that the fuzzy and metaphorical notions of aboutness and directedness merely

¹⁰In what follows, I will also sometimes assume something like a language of thought view (Fodor 1975), on which there are subpropositional vehicles of representation, like RED and CAT, which come together to constitute complex vehicles of representation representing complex contents. Apart from, I think, being largely correct, this assumption provides a useful way of talking about our particular representational capacities via their alleged vehicles. However, this assumption, along with the assumption that vehicles are internal items, can be discharged by replacing talk of mental representations with the more awkward talk of intentional capacities and amending my discussion appropriately.

See Ryder 2009 for a useful overview of different views of the structure of representations.

gesture toward. This section considers some alternative ways of defining “intentionality” and shows that they might not pick out the same thing as the ostensive definition (§§1.3.1–1.3.4). It then argues that if what we are interested in is the phenomenon that “aboutness” talk gestures at, the ostensive definition is preferable to these alternatives (§1.3.5).

1.3.1 Folk psychology

One approach to intentionality defines it in terms of its role in a third-personal folk psychological, or common sense, theory of mind and behavior. We attribute beliefs, desires, and other mental states to each other, and we take these states to be related to one another in various ways and to have various other features. A definition of “intentionality” in terms of folk psychology takes intentionality to be whatever plays a particular role in such a folk psychological theory.¹¹

Such a definition might not pick out the same thing as the ostensive definition. For instance, it could turn out that what the ostensive definition picks out lacks some of the extra features attributed to it by folk psychology. If these features are considered crucial by folk psychology, then the folk psychological notion will not pick out the ostensively defined phenomenon. For example, it could turn out that the ostensively defined phenomenon does not play certain causal roles considered crucial by folk psychology.

The folk psychological definition and the ostensive definition might also pick out different things if the folk psychological definition fails to pick out anything at all. Suppose that folk psychology is hopelessly false. Then its theoretical terms, including those putatively referring to intentional states, will fail to refer, and it will turn out that what it calls “intentionality” does not exist.¹² But the ostensively defined phenomenon might still exist. So,

¹¹See Sellars 1956, Lewis 1974, Braddon-Mitchell and Jackson 1996, and Fodor 1987 for this way of fixing reference on intentionality.

¹²Paul and Patricia Churchland argue that folk psychology is false and hence that its posits fail to refer (see, e.g., Churchland 1981). If they are right, and if we take intentionality

the folk psychological notion might fail to pick out the same thing as the ostensive definition.

1.3.2 The mind-brain sciences

Another approach to intentionality takes it to be a posit in scientific approaches to the mind and brain. For instance, some approaches in cognitive science aim to explain mental processes and behavior in terms of operations over internal states that are described as carrying information or “representing” various contents, and it is not uncommon for neuroscientific theories to speak of neural structures as carrying information about or “representing” their causes. A suggestion for an alternative way of picking out our target, then, takes intentionality to be a posit in the mind-brain sciences.¹³

There are interesting questions in the philosophy of science surrounding the notions of representation operative in various disciplines and research programs. What are these notions of representation? What roles do they play? Do different research programs use the same notion of representation? Some philosophers explicitly claim to be trying to answer these types of questions and not the types of questions I’m concerned with.¹⁴

It could turn out that this approach picks out the same thing as the phenomenon we noticed introspectively in ourselves. But it also might turn out that the best elucidations of the notions implicit in the mind-brain sciences pick out different features of internal states than the one we ostensively picked out through introspective observation. One *prima facie* reason to think this

to be merely a posit in folk psychology, then it will turn out that there is no intentionality. (Note that the Churchlands do not think there are no intentional states of any sort; see Churchland 1989b.)

¹³See, e.g., Fodor 1987, Millikan 1984, and Cummins 1994. Note that Fodor takes intentionality to be a posit in computational cognitive science, as well as a posit in folk psychology; on his view, cognitive science and folk psychology point to the same thing.

¹⁴Cummins (1994, pp. 278–279), for instance, specifically claims to be describing a notion of representation that is useful for computational theories of cognition, but not necessarily for the kinds of representation implicit in folk psychology.

might be the case is that it makes sense to ascribe at least some of the kinds of representational states operative in the mind-brain sciences to artifacts that we might not really believe to have genuine intentional powers, such as calculators and computers. This suggests that, at best, the ostensibly defined phenomenon is a species of whatever representational phenomenon is picked out by the mind-brain sciences. At worst, it is something else entirely.

It could also turn out that the two ways of defining “intentionality” do not pick out the same thing because the definition based on the mind-brain sciences does not pick out anything at all. Perhaps the best understanding of talk of representation in the mind-brain sciences takes representational notions to be merely a dispensable fiction (see Egan 2010). Then the mind-brain sciences do not really posit representational states after all. Another possibility is that they do posit representational states, but nothing plays the roles they are defined in terms of, so the notions of intentionality based on the mind-brain sciences fail to refer. Again, this shows that this way of defining “intentionality” might not pick out the same thing as the ostensive way.

1.3.3 Getting around in the world

For the most part, we manage to acquire the things we need, avoid the things that are harmful to us, perform sophisticated actions involving multiple steps, and, more generally, get around in the world fairly successfully. It is quite plausible that we do this by means of internal representations of the world. Inspired by this way of thinking, we might take intentionality to be whatever explains successful behavior in the relevant way.¹⁵

Behavioral phenomena such as those listed above call out for explanation, and it may very well be that the phenomenon we fixed on with our ostensive definition is a crucial part of this explanation. What is less clear is exactly

¹⁵Versions of this approach might also be versions of the approaches based on folk psychology or the mind-brain sciences described earlier.

what role the ostensively defined phenomenon plays. Vehicles of intentionality might have properties apart from intentional properties, such as syntactic, neural, or other broadly physical or functional properties.¹⁶ It could turn out, then, that intentionality itself is causally impotent and it's these other properties of vehicles of intentionality—say, their syntactic properties—that are responsible for their usefulness in helping us get around in the world. If this (unhappy) situation were the case, an explanation of successful behavior might not involve the ostensively defined phenomenon, and so the approach to defining “intentionality” based on getting around in the world might fix reference on something other than what the ostensive definition picks out.

The two definitions might also fail to pick out the same thing in skeptical scenarios in which we do not in fact manage to acquire the things we need, avoid the things that are harmful to us, or generally manage to get around in the world successfully, perhaps because we are brains in vats or dreaming, disembodied souls. In such scenarios, the definition based on getting around in the world would fail to fix reference on anything, since nothing in fact helps us get by in the world in the way required, but the ostensive definition would not fail to refer.

1.3.4 Truth and reference

If mental states “say something,” then it seems to follow that what they say can be either true or false, and if mental states are “of” or “about” something, then it seems to follow that they can either refer or fail to refer to whatever they’re “of” or “about.” So, perhaps we can use the notions of truth or reference to fix on our target. One such approach takes intentionality to be the having of conditions of truth or reference, while another takes intentionality to be that which gives rise to the having of conditions of truth

¹⁶See also Dretske’s distinction between representational facts and mere facts about representations (1995, p. 3).

or reference.¹⁷

This approach is certainly attractive. It promises to provide a substantive characterization of intentionality, defining it in terms of one of its apparently key features, and it does justice to the intuitive idea that intentionality serves to connect us to the external world, the world outside the mind. It also does justice to the idea that, at least when they are successful, there is some existing thing that intentional states are in some sense directed at or about. The approach also seems fairly unobjectionable. It certainly seems that intentional states have conditions of truth and reference, that what we think can be true or false, or that an object we represent can exist or fail to exist. The claim that intentional states have conditions of truth and reference is quite uncontroversial.

But, like the other alternative definitions of “intentionality” considered above, definitions in terms of truth and reference might fail to pick out the phenomenon picked out by the ostensive definition. They would fail to pick out the ostensively defined phenomenon if intentionality did not automatically connect us to the world without the help of additional ingredients. Consider, for the sake of illustration, a sense data theory of perceptual intentionality (a view, of course, which has few contemporary adherents¹⁸). On one way of characterizing such a theory, it can be divided into two main claims: First, perceptual intentionality is a relation of awareness to sense data. Second, sense data refer to whatever external items cause them (or bear some other special relation to them). The first claim offers a story of perceptual intentionality in terms of sense data, while the second claim offers a story of truth and reference in terms of causation (or other relations). It could turn out that the second part of the story is secured by the first part, that the nature of sense data and our relation to them makes it the case that intentional states

¹⁷See, e.g., Siewert 1998, Chalmers 2004, Siegel 2010, and Byrne 2009 for understandings of intentionality based on truth, reference, or the closely related notion of accuracy.

¹⁸In the 2009 PhilPapers Surveys, 3.1% of respondents at leading English-speaking universities reported holding a sense-datum theory (Bourget and Chalmers 2014).

refer to certain items in certain conditions. But it could also turn out that the first part of the story leaves open whether and how sense data, and hence intentionality, is connected to the world through reference. On such a theory, perceptual intentionality *alone* might not automatically give rise to conditions of reference.

Something similar might be true of other pictures of intentionality. Consider a Frege-inspired picture on which intentionality is a matter of being appropriately related to abstract senses, and truth and reference are a matter of how senses connect with the world. Depending on how we characterize senses, their connecting with the world might not occur automatically but might instead require an extra ingredient, a “satisfaction” relation or some such. So, if truth and reference require the ostensively defined phenomenon in combination with something else, a definition of “intentionality” in terms of conditions of truth and reference or what gives rise to them will pick out this combined phenomenon rather than the ostensively defined phenomenon alone.

Relatedly, the definition in terms of truth and reference would fail to pick out the ostensively defined phenomenon if it fails to refer because having conditions of truth and reference requires something in addition to the ostensively defined phenomenon of intentionality and our mental states exhibit the ostensively defined phenomenon but lack the additional ingredients. In such a case, there would be no such things as conditions of truth and reference, and so there would be nothing answering to the definition of “intentionality” in terms of truth and reference. But there would still be something answering to the ostensive definition.

1.3.5 Why we should prefer the ostensive definition

I have outlined various alternatives to the ostensive definition of “intentionality” and argued that these alternative definitions might pick out something other than the ostensively defined phenomenon. Now, of course, there is

no arguing over definitions. Different definitions of “intentionality” might pick out different things, and we are free to theorize about any of them. However, I want to suggest that if we are interested in the phenomenon gestured at by “aboutness” and “directedness” talk, we should prefer my fairly minimal ostensive definition. As we’ve seen, alternative definitions of “intentionality” build in assumptions about their target that are not present in the ostensive definition. This leaves someone who adopts these alternative definitions vulnerable to eliminativist threats: If there is nothing that satisfies the additional assumptions, then there is no intentionality, in their sense. But, I want to suggest, the phenomenon gestured at by “aboutness” and “directedness” talk is not vulnerable to eliminativism in the same way, which suggests that the extra assumptions that are built into the alternative definitions are substantive claims about, rather than defining features of, the phenomenon we gesture at with “aboutness” and “directedness” talk. If this is right, then there is reason to prefer my fairly noncommittal ostensive definition.

Here is a thought experiment that supports this point: Suppose that folk psychology is horribly mistaken, the mind-brain sciences have no need for a notion of content at all, and, relatedly, our best account of how we successfully get by in the world doesn’t either. Suppose further that there is no determinate fact about how mental states are supposed to correspond to the world, and so there are no such things as truth and reference. On this scenario, none of the alternative ways of defining “intentionality” manage to pick out anything at all. Still, on this scenario, we might introspect and notice paradigm cases of intentionality. We might notice perceptual experiences and thoughts that seem to be “about” or “directed” at something or that seem to “say something.” And we might want to know how this “aboutness” arises. This curiosity would not be misdirected, a mere result of our ignorance that the alternative definitions fail to refer. Even if we knew that nothing, not even paradigm cases, had the features invoked by the alternative definitions, we would still be left with the question of how thoughts and experiences

get to have the features we are tempted to describe using representational vocabulary like “aboutness” and “directedness.” What this shows is that observation of paradigm cases by itself gives rise to curiosity about “aboutness” and “directedness,” which suggests that my ostensive definition best captures the notion of intentionality that such talk gestures at.

1.4 Worries with the ostensive definition

I now turn to some potential worries one might have with the ostensive definition of intentionality.

1.4.1 Are we talking past each other?

When different theorists pick out their topic of interest in different ways, there is a danger that they end up talking past one another. Suppose that the folk psychologically defined phenomenon is distinct from the ostensively defined phenomenon. It seems that there is no real disagreement to be had between someone who takes the nature of the folk psychological phenomenon to be N and someone who takes the nature of the ostensively defined phenomenon to be M. The two theorists take different phenomena to have different natures.

However, it seems to me that in the case of many debates on intentionality, there often is a real disagreement between competing theories that employ different definitions of “intentionality.” Many theorists who define “intentionality” in alternative ways still intend their theories to account for the phenomenon that we introspectively observe and are tempted to describe representationally, the phenomenon gestured to by “aboutness” talk. For example, although Dretske (1995) does not employ an ostensive definition of his target, he states that his tracking theory covers anything answering to the term “intentionality”:

Brentano (1874) conjectured that a mark of the mental was inten-

tionality. Whatever, exactly, Brentano meant by intentionality, and whether or not he was right about its being a feature of all, and only, mental events, most philosophers take intentional characteristics (variously understood) to be distinctive of a great many mental phenomena. What follows is a brief catalog of those aspects of intentionality that have figured most prominently in the recent literature. In each case we find that a representational account of the mind provides a satisfying explanation of intentionality. (Dretske 1995, p. 28)

The aspects of intentionality that Dretske claims to accommodate are the power to misrepresent, aboutness, aspectual shape (roughly, our ability to represent things in different ways), and directedness (pp. 28–34). In effect, Dretske claims that all there is to any kind of intentionality-like phenomenon we have any reason to believe in is captured by his account. So, even though he does not define “intentionality” in my ostensive way, at least part of what he aims to account for is the ostensively defined phenomenon, the phenomenon gestured at by “aboutness” talk.¹⁹

Further reason to think that many theorists who define “intentionality” in one of the ways I reject aim to be targeting a phenomenon that at least includes the ostensively defined phenomenon is that they often use what appear to be introspectively accessible paradigm cases to illustrate their claims, such as judgments concerning barnyard animals and hallucinations of pink rats and daggers. Although there could be non-introspectively-accessible intentional states with such contents, the examples are usually supposed to be of the kinds of states that are or at least could be introspectively accessible.²⁰

¹⁹Similarly, Fodor (1990) explicitly states that his theory solves “Brentano’s problem” (pp. 127–128).

²⁰This is especially clear in discussions of the disjunction problem, which partly rely on intuitions about what is represented in possible cases (e.g., Baker 1989 and Fodor 1990). If these intuitions aren’t supposed to be intuitions about the kinds of contents that a subject might notice in herself, then it’s not clear where they are supposed to come from and why we should put any weight on them.

This appeal to introspectible cases suggests that whatever else theorists who fix reference on their target in ways other than my own are trying to do, they are also often trying to explain intentionality in my sense.

1.4.2 Standing states

One might agree with my suggestion of defining “intentionality” ostensibly but find my choice of paradigms overly restrictive. All my paradigm cases are **occurrent states**, mental states that are used, entertained, or otherwise active at the time at which they are had, such as judgments and perceptual states. But we might also want to include in our stock of paradigms some **standing states**, mental states that need not be used, entertained, or otherwise active at the time at which they are had, such as beliefs and desires that one is not currently entertaining. For example, the belief you had five minutes ago that the Acropolis is in Athens is a standing state, and one might suggest that it is a prime example of intentionality.²¹

My reason for not including standing states in my stock of paradigm cases is that we do not have the same kind of introspective access to them as we do to introspectively accessible occurrent states. While we can simply observe the intentionality of my preferred paradigm cases, we cannot observe our standing states or their features. Instead, we *infer* that we have standing states, perhaps on the basis of our noticing that we sometimes have corresponding occurrent states or on the basis of a folk psychological theory of mind and behavior. These ways of knowing about them are relatively indirect compared to our ways of accessing my preferred paradigm cases, and, relatedly, their existence is less certain for us than that of my introspectively accessible paradigm cases. In short, then, we have an especially secure epistemic access to introspectively accessible cases of intentionality, one that we do not have to standing states, which is why I do not include standing states in my paradigm cases.

²¹Thanks to David Bourget and Tim Bayne for pressing me on this worry.

Notice also that if, as the objector is likely to hold, the proposed additional paradigms have the feature that we notice in my paradigm cases, then my choice to not include them has no effect: My ostensive definition will cover them as well, and including them in our initial stock is unnecessary. We end up picking out the same thing either way.

Excluding standing states only has an effect on what we end up picking out if the proposed additional paradigms either do not exist or do not have the feature picked out by my ostensive definition. But, in such cases, it is arguably a virtue of my ostensive definition that it keeps the clearly intentional introspectively accessible feature of mental states separate from whatever it is that the proposed additional paradigms have. In any case, the terminological difference between me and someone who adopts an ostensive definition including standing states as paradigms does not make a substantive difference in what follows: Although my target is intentionality, in my sense, I also offer an account of standing states, as well as other alleged instances of intentionality (see Part IV). The overall story offered would be the same whether or not we include standing states in our paradigm cases, though it would be stated differently.

1.4.3 Perception and thought

The ostensive definition fixes on our target by pointing to several of its instances, including instances in perception and instances in thought. But one might worry that these instances do not belong to a unified natural kind. Perhaps we end up picking out a disjunctive kind, consisting of two distinct natural kinds. This might be the case if perceptual states are importantly different from “cognitive” states such as thoughts, and so what we might call “perceptual intentionality” is not the same kind of thing as what we might call “cognitive intentionality.” If this is the case, then my paradigm cases are actually instances of two different kinds of phenomenon, which I am mistakenly lumping together.

Of course, when we use multiple examples in an ostensive definition, there is always a risk that they are very different in their natures and we end up picking out a disjunctive kind. One response to this sort of worry is that this scenario is unlikely. While it remains a possibility that I've picked out a disjunctive kind, it at least initially seems that the observations concerning thought and those concerning perception are similar in important ways. Both thought and perception are readily described as “saying something” or being “directed at” something. And both intentionality in perception and intentionality in thought seem fairly distant from other kinds of phenomena, such as reflexive behaviors and the automatic control of vital functions. Their similarity to one another and distance from other phenomena suggest that there is an interesting natural kind that they both belong to.²²

A second response to this worry is that even if it turns out that perceptual intentionality and cognitive intentionality are very different phenomena that do not form a unified natural kind, this is not a problem, since my starting point will not steer us too far in the wrong direction. Assuming our target is whatever “aboutness” and “directedness” talk gesture toward, the problem with approaches to fixing reference on our target that I want to reject is that they risk missing our target entirely. For example, defining our target as an explanatory posit in a theory of behavior risks missing our target if intentionality does not play the requisite role in generating behavior. If nothing plays the requisite role (say, because the relevant parts of the theory are false), then it will turn out that there is no intentionality. If something plays this role, but it is not whatever “aboutness” and “directedness” talk gesture toward, then it will turn out that there is intentionality, but that it's not the same thing as our targeted phenomenon. In contrast, picking out a disjunctive kind does not carry with it the risk of missing our target. Perhaps

²²Of course, even if intentionality is a unified natural kind, perceptual representation and representation in thought might end up forming two more specific distinct natural kinds. This does not affect my claims here, since the issue is over whether intentionality is a unified natural kind, not whether it has various distinct subkinds.

perceptual intentionality and cognitive intentionality are two entirely different kinds of things. Then we would need two distinct, and perhaps unrelated, theories to explain them, and if we start off thinking of perceptual content and thought content as relevantly similar, then it might take longer to reach such a conclusion. However, such a conclusion has not been ruled out from the start, because nothing in the way we fixed on intentionality requires that it be a unified phenomenon.²³

1.4.4 Perceptual states don't have contents

One might object that perceptual states don't have contents, at least not prior to an act of "interpretation." For example, one might argue that the visual experience one enjoys when one views a red ball is neutral between multiple external-world possibilities, such as that there is a red ball in normal lighting conditions, that there is a white ball lit by red light, etc. The experience does not by itself "say" which of these possibilities is the case, so it does not represent the ball as being any particular color at all. Instead, a further state, such as a judgment, interprets the perceptual state and commits us to one or another possibility. On this view, there is a distinction between non-intentional mental features of some perceptual experiences, which we might call their "raw matter," and further states that interpret them, or "interpretations." Interpretations are intentional, but mere raw matter is not. One might object that if this view is correct, then my ostensive definition is too permissive: it includes both raw matter and interpretations, whereas we should only include interpretations.²⁴

However, if the above view is correct, then it is not in fact the case that my

²³Of course, whatever their apparent similarities, perceptual intentionality and cognitive intentionality also seem quite different in certain respects. For instance, perceptual intentionality is more vivid, detailed, and closely related to phenomenology than is cognitive intentionality. Eventually, I will offer a view of intentionality that begins to explain both the similarities and differences between intentionality in perception and in thought (see especially Chapter 7).

²⁴Travis (2004) presents a view along these lines.

approach is too permissive. My observations pick out *intentional states*, states that “say” something, not non-intentional components or contributors to those states. And so, my way of picking out the phenomenon of intentionality isn’t meant to and wouldn’t in fact pick out uninterpreted raw matter, if there were such a thing. Instead, my method would pick out interpreted raw matter, or interpretations. It might then turn out that some of our allegedly perceptual paradigm states are not mere perceptual states but instead thoughts or combinations of perceptual states and thoughts. If the above view is in fact correct, then, ultimately, a complete theory of intentionality should isolate the components of interpreted raw matter and distinguish their contributions to intentionality. Note that nothing in my ostensive definition rules out such a view from the get-go.

1.5 Conclusion

The central aim of this book is to offer a theory of intentionality, the phenomenon we at least sometimes notice introspectively in ourselves and are tempted to describe using representational terms, and which, I’ve suggested, is the phenomenon that “aboutness” talk gestures at. This book develops a theory of this ostensively defined phenomenon in terms of a conceptually distinct mental feature, phenomenal consciousness, the “what it’s like” of mental states.

This book also has a secondary aim. This chapter considered and rejected alternative ways of fixing reference on intentionality via some of its alleged additional roles. Thus far, I have argued that intentionality *might not* play these roles. A secondary line of argument in this book argues that intentionality alone in fact *does not* play many of these roles. Many of them are played by something else. In Chapter 3, I argue that a crucial part of a story of how representations contribute to successful behavior must invoke non-intentional features of representations, namely, their tracking relations

to external items. In Chapter 8, I argue that folk psychological notions of content most closely correspond to a combination of intentional content and derived mental representational content. In Chapter 8, I also argue that the kinds of representation implicit in the mind-brain sciences are distinct from intentionality and that nonconscious occurrent states might satisfy these notions of representation but lack intentionality. Chapter 9 argues that it is not even clear that intentionality gives us conditions of truth and reference without the help of further ingredients.

On the resulting picture, then, intentionality is a matter of phenomenal consciousness, and many of the other roles that are sometimes used to pick it out are in fact played by something else.