Phenomenal Intentionality:
How intentionality arises from phenomenal consciousness

Angela Mendelovici

Draft July 12, 2016
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 of intentionality we are after. Chapter 1 proposes an ostensive way of fixing reference on intentionality, while Chapter 2 specifies what kind of theory we are after and overviews two sources of theory-independent knowledge of intentionality that we can use to test our theories: introspection and considerations of psychological role.
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, because if we are not clear on what is our target, then it won’t be clear just what our theories say.

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.

1.1 Aboutness and directedness

Intentionality is sometimes initially characterized as the “aboutness” or “directedness” of mental states (and perhaps other items) to things that may or may not exist. Some mental states seem to be about or directed at things. For example, a perceptual experience of a cup is in some way directed at a cup, a thought that it is raining is about the proposition that it is raining,
and a belief in Santa Claus is about Santa Claus or the proposition that Santa Claus exists.¹

This notion of intentionality in terms of “aboutness” or “directedness” has roots in Brentano’s oft-cited passage:

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)

While the characterization of intentionality in terms of aboutness and directedness arguably succeeds in gesturing towards the phenomenon of interest, we should not take it to be a definition, not even a 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), like 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”

¹For some representative examples of this way of characterizing intentionality, at least as a first pass, see Jacob (2003), Siewert (2006), Byrne (2006), Kim (1998, p. 21), Searle (2004, p. 112) and O’Madagain (2014). For example, 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.”
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 at 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. I want to suggest that this allows us to
ostensively define “intentionality” as this mental 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 consider 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 them as being “of” or “about” things or ways things are or
might be, or as “saying something.” We might say they are “about” some

\(^2\)Chisholm (1957a) criticizes Brentano’s definition as being too fuzzy and suggests
instead a linguistic criterion of intentionality. See also Speaks 2010b for criticisms of
definitions of “intentionality” in terms of “aboutness.”
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 is tempting to describe using representational terms. We might describe them as being “about” things or as “saying something” about how things are. We might say that they are “about” our experiences, that they “say” that it’s time for a cup of coffee, etc.\(^3\)

These observations should be obvious and uncontroversial. We notice in ourselves certain mental states such as those mentioned above. Further, we notice that they have a certain feature, and we are tempted to describe this this feature using representational terms like “aboutness” and “directedness.” What is controversial is just what this feature consists in, how it relates to any other features of mental states (e.g., any physical or functional features they might have), and what exactly any given mental state is about or says. But it should be obvious and uncontroversial that we can notice that we have mental states with features that we’re tempted to describe in representational terms. This allows us to fix reference on our target ostensively:

**Intentionality** The feature of mental states that we can at least sometimes notice introspectively in ourselves in mundane cases such as those described above and are tempted to describe using representational terms such as “about,” “of,” “represent,” “present,” or “saying something.”\(^4\)

This definition of “intentionality” fixes reference on its target ostensively by pointing to instances of what appear to be the same kind of thing. These instances, which are readily apparent to introspection, form our base of

---

\(^3\)I 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 §1.4.1.

\(^4\)Definitions of key terms can be found in the glossary on p. 296.
paradigm cases of intentionality, the cases that are available to form our initial sample for the purposes of our ostensive definition. The ostensive definition takes intentionality to be the feature of paradigm cases that we are tempted to describe using representational terms.\footnote{This kind of ostensive definition of intentionality has been developed and defended by Mendelovici 2010 and Kriegel 2011b.}

The ostensive definition arguably picks out precisely what the characterizations of intentionality in terms of aboutness and directedness are gesturing that, and so it does justice to those characterizations. However, 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 using them.

Note that, although we are using introspection to fix reference on intentionality, the ostensive definition does not require that all instances of intentionality are introspectively accessible. Presumably, intentionality is a fairly natural kind (though it need not be a unified kind), allowing us to get a grip on it by pointing to some of its instances, even if we cannot point to all of its instances, or if we cannot point to all of its instances using introspection. As far as my reference-fixing definition is concerned, the feature of introspectible states that we fix reference on might be had by states that are not introspectible, such as unconscious beliefs or unconscious states involved in linguistic processing. My definition also does not rule out the possibility of non-mental intentionality.

We can use this ostensive definition of “intentionality” to define some related notions: Intentional properties are ways things are with respect to their intentionality, or intentional ways things are. Intentional states are instantiations of intentional properties. Intentional states are not the same thing as intentional mental states, which are mental states that involve, 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.

Intentional properties and intentional states can be said to have intentional contents, which are what they “say” or are “directed at.” 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.\(^6\) When a state, property, or other item has a certain content, we can say that it (intentionally) represents that content.\(^7\) For example, the judgment that grass is green represents the content <grass is green>.

This starting point is fairly non-committal. Our introspective observations so far are compatible with various possible theories of intentionality, such as theories on which the phenomenon we observe is a matter of causal or other tracking relations, functional roles, or phenomenal consciousness. They are also compatible with views on which intentionality is a matter of being directly presented with worldly items, as on direct realist views of perception, and with adverbialist views of perceptual experience, on which perceptual experiences are modifications of subjects. These theories do not deny that there is intentionality in the ostensively defined sense; rather, they are attempts at accounting for (perhaps certain types of) intentionality.

Relatedly, our introspective observations do not reveal the nature of contents. Contents could turn out to be ordinary objects and properties, propositions, facts, sense data, ideas in a world of forms, ways of represent-

\(^6\)When 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.

\(^7\)I sometimes use “represent” more broadly to describe representation-like phenomena that are not instances of intentionality, but context should disambiguate. The alternative “to intend” is too awkward.
ing, properties of intentional states, or even intentional states or properties themselves, or it could turn out that they lack a unified nature. My approach is compatible with views of contents that take them to be items that exist independently of us that we entertain or otherwise represent, as well as with views that take them to be intrinsic features of our minds. What we’ve noticed through introspection does not automatically decide between any of these views.

Our introspective observations also do not prejudge any issues regarding the vehicles of intentionality, which are the bearers of intentional properties. The vehicles of intentionality could be, for example, subjects, symbols in a language of thought, brain states, internal states, or immaterial souls. For simplicity, however, I will take the vehicles of intentionality to be internal items that I will call (mental) representations. Since different intentional states will involve different vehicles of representation, this way of speaking allows us to talk about intentional states while remaining non-committal on their contents, which is useful when the content of a particular intentional state is under dispute.\footnote{Those who think subjects are the vehicles of intentionality can replace talk of mental representations with talk of internal states or capacities that allow for or correspond to subjects’ representation of various contents and adjust my claims accordingly. Presumably, when subjects have various intentional states, this at least corresponds to their being in certain internal states, even if these internal states are not properly considered vehicles of representation.}

\section{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 gesture towards. This section considers alternative ways of defining “intentionality” and explores how they can come apart from the ostensive way. §1.3.5 explains why I prefer the ostensive way.
1.3.1 Folk psychology

One approach to intentionality is to take it to be a posit in our 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. We might then take intentionality to be whatever plays a particular role in folk psychological theory.\(^9\)

Such a definition of “intentionality” in terms of folk psychology might not pick out the same thing as the ostensive definition. 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, or that the folk psychological notion of intentionality picks out a broader or narrower class of states than does the ostensive definition.

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.\(^10\) But the ostensively defined phenomenon might still exist. So, the folk psychological notion might fail to pick out the same thing as the ostensive definition.

---


\(^10\)Paul 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 to be merely a posit in folk psychology, then it will turn out that there is no intentionality. (Notice that the Churchlands do not think there are no intentional states of any sort (see Churchland 1989b).)
1.3.2 The mind/brain sciences

Another approach to intentionality takes it to be a posit in scientific studies of the mind and brain. For instance, some brands of cognitive science aim to explain mental processes and behavior in terms of operations over internal states that are described as carrying information, representing, or having content, and it is not uncommon for neuroscientific theories to speak of neural structures as representing or carrying information about 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.\(^{11}\)

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 role 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. Cummins (1994, pp. 278-9), 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.

It could turn out that these alternative definitions pick out the same thing as the ostensive definition. Maybe the intentionality that we can at least sometimes introspectively notice is just what cognitive scientists or other mind/brain scientists are talking about. But it also might turn out that the best elucidation of the notions implicit in the mind/brain sciences pick out a different feature of internal states from the one we ostensively picked out through introspective observation. One prima facie reason to think this might be the case is that it makes sense to ascribe at least some of the kinds of representational states operative in cognitive science to artifacts that we might

\(^{11}\)Fodor (1987) takes intentionality to be a posit in computational cognitive science, as well as a posit in folk psychology. He takes cognitive science and folk psychology to point to the same thing. See also Millikan 1984.
not really believe to have genuine intentional powers, such as calculators and computers.

It could also turn out that two ways of defining “intentionality” do not pick out the same thing because the way 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. 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 it defines them 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 an explanatory posit in a theory of the generation of successful behavior.\(^{12}\)

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 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.\(^{13}\) It could turn out, then, that intentionality itself is causally impotent and it’s these other

\(^{12}\)Versions of this approach might also be versions of the approaches based on folk psychological or the mind/brain sciences described above.

\(^{13}\)See also Dretske’s distinction between representational facts and mere facts about representations (1995, p. 3).
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 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 that they either successfully refer to whatever they’re “of” or “about” or they don’t. 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 conditions of truth or reference.\(^\text{14}\)

This approach is certainly attractive. It promises to provide a substantive characterization of intentionality, and it does justice to the idea that intentionality connects 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

\(^{14}\text{See, e.g., Siewert (1998), Chalmers (2004), Siegel (2010), and Byrne (2009) for understandings of intentionality based on truth and reference.}\)
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 perceptually experience can exist or fail to exist.

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. Although such a theory has few present-day adherents, it is instructive to consider a theory of the perceptual instances of intentionality in terms of sense data, which might be taken to be mind-dependent mental particulars. 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 connect us with items in the external world when those items cause them (or when they bear some other relation to them). The first claim offers a story of perceptual intentionality, while the second claim offers a story of how perceptual intentionality connects us to the world. 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 connect us to the external world in certain conditions. But it could also turn out that the first part of the story leaves open whether and how sense data connect us to the external world. On such a theory, perceptual intentionality alone might not automatically give rise to conditions of truth and 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 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 in which there is no such thing as conditions of truth and reference, there would be nothing answering to the definition of “intentionality” in terms of truth and reference, though there would be something answering to the ostensively defined notion.

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 those things. However, I want to suggest that if we are interested in the phenomenon gestured at by “aboutness” and “directedness” talk, there is reason to prefer my fairly minimal ostensive definition. As we saw above, alternative definitions of “intentionality” build in assumptions about their target that are not present in the ostensive definition. For example, the folk psychological definition assumes that intentionality is whatever plays a certain role in folk psychological theories. This leaves someone who adopts these alternative approaches vulnerable to eliminativism about intentionality: If there is nothing that plays the relevant role, then there is no intentionality. 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 optional assumptions 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 non-committal ostensive definition, which does not build in such extra assumptions.

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 the paradigm cases, had the features invoked by the alternative definitions, we would still be left with the question of how the paradigm cases 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.
1.3.6 Why we’re not just 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. One might worry 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 debates on intentionality, there often is a real disagreement between competing theories that employ different definitions of “intentionality.” For many theorists, the target of the alternative definitions is at least meant to include the target of the ostensive definition, even if it includes other targets as well. Regardless of what additional features various theorists attribute to their targets, it seems that part of what they want to account for is the observed aboutness that I am taking to be central to the notion. For example, although Dretske (1995) does not explicitly provide a way of fixing reference on his target, he states that his tracking notion of intentionality covers anything answering to the term “intentionality”:

Brentano (1874) conjectured that a mark of the mental was intentionality. 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 with his notion of representation are the power to misrepresent, aboutness, aspectual shape (our ability to represent things in different ways), and directedness (p. 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, he presumably aims to include the ostensively defined phenomenon as part of his target.

Likewise, although Fodor (1987) explicitly defines his target in terms of folk psychology and truth, (roughly) taking intentionality to be that in virtue of which the mental state posits of folk psychology have truth conditions, he aims to capture our everyday representational perspective on the world:

There’s quite a lot of Greycat’s behavior that I want to explain by adverting to the way Greycat takes the world to be; how he represents things. For example: It’s part of my story about why Greycat turns up in the kitchen in the morning that Greycat has a story about his bowl; and that, in Greycat’s story, the bowl figures as—it’s represented as being—a likely locus of food. (Fodor 1987, p. x, emphasis in original)

Presumably, Greycat’s representational states are of the same kind as the paradigm intentional states we notice in ourselves. So, although Fodor is interested in a phenomenon posited by folk psychology, his target seems to include intentionality in my sense. This is also evident in Fodor’s discussion of trains of thought, which he takes to involve the kinds of psychological states he is interested in:

Here, for example, is Sherlock Holmes doing his thing at the end of “The Speckled Band”:

I instantly reconsidered my position when . . . it became clear to me that whatever danger threatened an occu-
pant of the room couldn’t come either from the window or the door. My attention was speedily drawn, as I have already remarked to you, to this ventilator, and to the bell-rope which hung down to the bed. The discovery that this was a dummy, and that the bed was clamped to the floor, instantly gave rise to the suspicion that the rope was there as a bridge for something passing through the hole, and coming to the bed. The idea of a snake instantly occurred to me, and when I coupled it with my knowledge that the Doctor was furnished with a supply of the creatures from India I felt that I was probably on the right track.

The passage purports to be a bit of reconstructive psychology: a capsule history of the sequence of mental states which brought Holmes first to suspect, then to believe, that the doctor did it with his pet snake. (Fodor 1987, p. 13-4)

Sherlock Holmes’ monologue seems to assert an introspective awareness of his represented contents (for instance, he says: “The idea of a snake instantly occurred to me”), and Fodor takes whatever Holmes is talking about to be included within the scope of his target. This suggests that Fodor’s target includes my paradigm intentional states.

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 paradigm cases to illustrate their claims, such as judgments concerning farmyard 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. This appeal to

15This is especially clear in discussions of the disjunction problem, which partly rely on
introspectible cases suggests that whatever else theorists who fix reference on their target in ways other than the ostensive way are trying to do, they are also trying to explain intentionality in my sense.16

1.4 Worries with the ostensive way of fixing reference to intentionality

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

1.4.1 Standing beliefs, standing desires, and other standing states

One might agree with my suggestion of defining “intentionality” ostensively, 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 you are 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.17

My reason for not including standing states in my stock of paradigm cases is that we do not have the same kind of immediate introspective access to

---

16 See also §10.2 for discussion of the extent to which my conclusions rest on terminological assumptions.

17 Thanks to David Bourget and Tim Bayne for suggesting worries along these lines.
them as we do in the case of introspectively accessible occurrent states. While we can immediately observe the intentionality of my preferred paradigm cases, we cannot immediately observe our standing states or their features. We might infer that we have standing states on the basis of our noticing that we have corresponding occurrent states, or on the basis of a folk psychological theory of mind and behavior, but these ways of knowing about them are relatively indirect compared to our ways of accessing my preferred paradigm cases. Relatedly, given our indirect access to standing states, their existence is less certain for us than that of my introspectively accessible paradigm cases. Indeed, in Chapter 8, I will eventually argue that many standing states do not exist. 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, and this is why I include introspectively accessible cases, but not standing states, in my initial stock of examples.

Notice also that if, as the objector seems 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.

Not including 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 a case, 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, such a terminological difference does not make a substantive difference in what follows: Although my target is intentionality, in my sense, I also aim to account for standing states like standing beliefs and standing desires, and other (perhaps merely) alleged instances of intentionality, even if they do not turn out to be instances of intentionality, in my sense.
1.4.2 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 the introspectively observable cases we used to ostensively fix reference on intentionality are actually instances of two different kinds of phenomenon, which I am mistakenly lumping together.

Of course, when we use multiple samples in an ostensive definition, there is always a risk that the samples 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 regulation of bodily functions. Their similarity to one another and distance from other phenomena suggests that there is an interesting natural kind that they both belong to.\(^\text{18}\)

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

\(^{18}\)Of 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.
point will not steer us too far in the wrong direction. Assuming our target is whatever “aboutness” and “directedness” talk gesture towards, 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 towards, 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 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. 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 picked out intentionality requires that it be a unified phenomenon.

Of 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).

1.4.3 Perceptual experiences don’t have content

One might object that it is a mistake to think of perceptual states, or at least some types of perceptual states, as intentional at all. Rather, perceptual
experiences need to be interpreted before they give rise to any intentionality. For example, consider the visual experience one enjoys when one views a red ball. One might argue that this experience is compatible with multiple external-world possibilities. It is compatible with there being a red ball in normal lighting conditions, with there being a white ball lit by red light, etc. The experience does not by itself “say” which of these possibilities is the case. It is silent between them. So, one might argue, perhaps it doesn’t represent the ball as being any particular color at all. Perhaps, instead, it is a further state, such as a judgment or interpretation of the initial state, that represents the ball as having a particular color. On this view, there is a distinction between non-intentional mental features of some perceptual experiences, which we may call their “raw matter,” and further judgments or states about or based on raw matter, or “interpretations.” On this picture, intentionality only appears at the level of interpretation; raw matter is not intentional on its own. One might object that if this view is correct, then my approach to intentionality is too liberal: it includes both raw matter and interpretations, whereas we should only include interpretations.19

This objection is misguided. If the above view is correct, then it is not in fact the case that my approach is too liberal. 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 manage to pick out uninterpreted raw matter, if there were such a thing. Instead, if the above view is correct, my method would pick out interpreted raw matter. Ultimately, a complete theory of intentionality should isolate the components of interpreted raw matter and distinguish their contributions to the phenomenon of intentionality that we observe and want to explain. But nothing in my ostensive definition rules out such a possibility.

19Travis 2004 presents a view along these lines.
1.5 Conclusion

I’ve proposed a minimally committal ostensive way of fixing reference on intentionality. The central aim of this book is to develop 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, for example, 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 7, I argue that intentional thought content does not specify the conditions of truth and reference that we care about; instead, a different kind of phenomenon, derived mental representation, is connected to the conditions of truth and reference that we care about. Chapter 9 argues that it is not even clear that intentionality gives us conditions of truth and reference without the help of further ingredients. In Chapter 8, I argue that folk psychological notions of content most closely correspond to a combination of intentional content and content derived from intentional content. In the same chapter, I also argue that the kind of representation implicit in various mind sciences is distinct from intentionality, and that nonconscious occurrent states might satisfy the mind/brain sciences’ notions of representation, but lack genuine intentionality. On the resulting picture, then, intentionality is a matter of phenomenal consciousness, and many of the various other roles that are sometimes used to pick it out are in fact played by something else.
Chapter 2

Goals and methodology

The previous chapter fixed on our target, intentionality. This chapter considers what exactly we want to know about our target, and proposes some ways in which we might come to know it.

§2.1 considers what it would take to provide a theory of intentionality and the goals around which I will structure most of the subsequent discussion. §2.2 proposes two theory-independent ways in which we can know about our intentional states and that we can use to test competing theories of intentionality: introspection and considerations of psychological role.

2.1 What is a theory of intentionality?

A theory of intentionality is a theory that describes the deep nature of intentionality, where intentionality’s (deep) nature is what it really is, metaphysically speaking. For example, a theory of intentionality might tell us that intentionality is a tracking relation to instantiated properties, a relation of isomorphism between a functionally defined system of representations and abstract propositions, or a primitive relation to objects and their properties.

My aim is to provide a theory of intentionality that specifies the nature of all actual and possible intentional states. But I will structure much of
my discussion around a less ambitious goal, that of providing a theory that specifies what gives rise to actual instances of original intentionality. Let me explain these three qualifications:

A gives rise to B (or, equivalently, B arises from A) when B is identical to, fully grounded in, constituted by, or realized by A. A theory that tells us what gives rise to intentionality tells us what exactly intentionality comes from, though it might remain neutral on how exactly intentionality comes from it. For example, a theory of intentionality might claim that intentionality arises from tracking relations obtaining between internal states and items in the environment, while remaining neutral on whether intentionality is identical to this tracking relation, grounded in it, or arising from it in some other way.

Actual, as opposed to merely possible, intentional states are intentional states existing in the actual world. A theory of intentionality might account for all actual intentional states while allowing for the possibility of intentional states that it cannot account for. For example, a theory of intentionality in terms of tracking might allow that in other possible worlds intentionality is a primitive phenomenon.¹

Original intentionality is intentionality that does not derive from other instances of intentionality. Original intentionality can be contrasted with derived intentionality, which is intentionality that derives from other instances of intentionality. For example, one might think that linguistic expressions have derived intentionality, which is derived from the original intentionality of mental states, e.g., from our thoughts, beliefs, interpretations, or communicative intentions.

A few more definitions are in order: Originally intentional properties are ways things are with respect to their original intentionality, or originally

¹Compare: A physicalist theory of mental states can be neutral on the question of whether non-physical mental states are possible. In the same way, a tracking theory of intentionality can be neutral on whether it is possible for intentionality to arise from something other than tracking.
intentional ways things are, and an *originally intentional state* is an instantiation of an originally intentional property. We can say that something *originally represents* a content when it instantiates an originally intentional property representing that content. Correspondingly, *derivatively intentional properties* are ways things are with respect to their derived intentionality, or derivatively intentional ways things are. A *derivatively intentional state* is an instance of a derivatively intentional property, and we can say that something *derivatively (intentionally) represents* a content when it instantiates a derivatively intentional property representing that content.

Note that although it is sometimes thought that the line between original and derived intentionality is to be drawn between mental instances of intentionality and non-mental instances, many views of intentionality accept mental instances of derived intentionality. We will soon see that this is a key part of most versions of the phenomenal intentionality theory, the view that I will eventually defend, but it is also true of many alternative views that accept a language of thought-like picture on which internally unstructured representations, representations that do not contain other representations as proper parts, come together to form internally structured representations whose contents are determined by their constituent representations and the ways they are combined. One natural way of understanding this idea takes the intentionality of internally structured representations to be derived from their constituent internally unstructured representations and their mode of combination.\(^2\)

While I will eventually propose a theory of intentionality that specifies the nature of all actual and possible intentional states, I will structure much of my discussion around the more modest goal of providing a theory of intentionality that specifies what gives rise to actual instances of original intentionality.

\(^2\)See Bourget (2010a), who suggests that intentionality that is obtained through composition is a kind of derived intentionality.
My reason for this is that it allows us to characterize the various approaches to intentionality in terms of their lowest common denominators: in terms where they take actual-world instances of intentionality to come from, or, in other words, in terms of what they take to be the actual-world “source” (Kriegel 2011b, 2013b) of intentionality. Once we have settled on the source of actual-world instances of intentionality, we can turn to providing a more complete theory of intentionality, one that settles other questions about intentionality, including those of how exactly intentionality arises from its source, how it might arise in other possible worlds, and whether and how original intentionality can yield derived intentionality.

2.2 Theory-independent knowledge of intentionality

As we will see in Chapters 3–5, many theories of intentionality make predictions as to which intentional contents we represent, so theory-independent access to such facts can help us decide between them. This section describes two theory-independent ways of finding out what contents we represent: introspection and observed psychological roles.

2.2.1 Introspection

In §1.2, I argued that introspection provides access to our intentional states, though it may not provide us access to various facts about the nature of intentionality. Here, I want to suggest that introspection can provide theory-independent knowledge of intentionality by telling us, at least in some cases, which intentional contents we represent.

In order to clarify what introspection can and cannot tell us about intentionality, we can contrast the notion of the deep nature of intentionality, introduced in the previous section, with that of what I will call its “superficial
character.” Recall that intentionality’s deep nature is what it is, at bottom, metaphysically speaking. We can also speak of the deep nature of intentional properties, states, and contents, which is what they are, metaphysically speaking. For example, contents might be sets of possible worlds, structured propositions consisting of objects and properties, or Fregean senses, and intentional properties might be tracking relations, functional properties, or phenomenal properties.

It is doubtful that introspection reveals to us the deep nature of intentionality. For example, introspection does not tell us whether contents are sets of possible worlds or structured propositions, or whether intentionality is a tracking relation or a matter of functional roles. Since a theory of intentionality is a theory that specifies the deep nature of intentionality, this means that introspection does not reveal to us the correct theory of intentionality.

Although our introspective observations are neutral on the nature of intentionality, they do tell us that we have intentional states, which is all I needed in order to fix reference on our target in Chapter 1. They also tell us a bit more than that: They at least sometimes tell us which intentional contents we represent. For example, introspection might tell me that I represent the content <there is a cat in my lap>. Although I might not be able to tell from introspection what is the deep nature of my intentional state or its content, I can tell which content I am representing. This knowledge might allow me to distinguish the content I represent from other contents, e.g., from the content <there is an octopus before me>. Furthermore, I do not know that I am representing <there is a cat in my lap> because my intentional state has, say, a number, symbol, or other marker that I have come to associate with it. Rather, I know that I am representing that particular content because I have some sort of access to features of this content that characterize it as the content it is and that distinguish it from other contents. One reason to think this is that my introspective knowledge of the content <there is a cat in my lap> allows me to discern some of its entailment relations with other contents,
e.g., that it entails <there is something in my lap>. Knowledge of a content’s entailment relations presumably involves knowledge of something more than simply a number or other marker for the content, though it needn’t require knowledge of the content’s deep nature. What it involves is knowledge of what I will call the **superficial character** of a content (and, by extension, of the states and properties that represent it), the features of the content (and, by extension, the properties and states that represent it) that characterize it as the content that it is and that distinguish it from other contents.

My claim is that introspection at least partially reveals the superficial character of at least some of the contents we represent. Note that this is compatible with introspection failing to grant full access to the superficial character of all our represented contents. For example, the superficial characters of unconsciously represented contents might be entirely inaccessible to introspection, and introspection might not be able to fully access the superficial characters of broad contents, contents the representation of which depends on relations to the environment. However, it is natural to suppose that we have access to which contents we represent in paradigm cases of intentionality (see §1.2), since these are the cases we have some kind of introspective access to.³

³See Mendelovici (forthcoming a) for the superficial character/deep nature distinction and its application to the debate on propositionalism. See also Bayne and Spener (2010) for a congenial discussion of the limits of introspective access.

It can be helpful to consider an analogy: You can know which pieces of furniture a room contains without knowing the deep nature of furniture, i.e., without knowing whether items of furniture are physical, functional, or even mental in nature. And the way you typically come to have such information is not by coming to associate particular items of furniture with numbers or other arbitrary markers, but rather by knowing certain of their distinguishing features, e.g., their shapes and functions. In the same way, you can know some of the features of your intentional properties, states, and contents without knowing their deep natures.
Above, I suggested that at least some theories of intentionality make predictions about which intentional states particular subjects have. These predictions can be construed as being or at least including predictions about the superficial characters of the relevant states. By providing us with access to the superficial character of paradigm cases, introspection provides us with a theory-independent way of knowing what at least some intentional states represent, and hence of testing particular theories of intentionality (see especially Chapters 3 and 5).\(^4\)

## 2.2.2 Psychological involvement

The second theory-independent way of finding out about our intentional states is through their psychological roles. Although it is pretheoretically unclear exactly what intentional states do and just how they do it, it should be fairly uncontroversial that our represented contents play some role in the mental economy. This role might be a role in connection to prior or subsequent intentional states, a role in connection to behaviors, or simply a role in constituting our representational phenomenology or our “grasped” representational perspective on the world. More generally, represented contents behave as if they’re there; they play some psychological role. We can say that an intentional state has a psychologically involved content when its representing this content plays a psychological role appropriate to which content it is, i.e., to its superficial character.\(^5\)

What psychological involvement amounts to might differ between occurrent and standing intentional states. The reason for this is that, while both occurrent and standing intentional states might be said to represent their

\(^4\)In a similar spirit, Georgalis (2006) recommends the following principle of methodological chauvinism: “The foundation for the development of any theory of representation should use a restricted database consisting of elements established from the first-person perspective to be representations or representational states.” (p. 285)

\(^5\)As Millikan (1997) writes, “we cannot suppose that a representation could be a mental representation, a representation for mind, yet that its representational value was independent of its effect upon mind.” (p. 515)
contents, it is only in the case of occurrent intentional states that contents are
thoughts, experienced, or, more generally, entertained. Since the contents
of occurrent states are entertained, they are available to be put to various
uses, e.g., in reasoning, drawing inferences, and generating behaviors. They
might also or instead contribute to our overall phenomenology, including our
subjective experiences and any phenomenology we might have of “grasping”
certain contents. And they might also or instead be introspectively accessible
and form potential targets of higher-order mental states, mental states about
our mental states. Note that not all contents of occurrent states need play
all of the above-mentioned psychological roles in order to qualify as being
psychologically involved; any given represented content need only play one of
them.

Unlike the contents of occurrent states, the contents of standing states
are not entertained, and so they are not immediately available to be put to
various uses. However, presumably, such represented contents are in principle
available to be used. For example, my standing belief that the Acropolis is in
Athens might not always be influencing my further thoughts and behaviors,
but it should be available to influence my thoughts and behaviors when needed.
Like contents that are entertained, the contents of standing states should
exhibit psychological involvement, though what psychological involvement
amounts to for them might differ from what it amounts to for the contents of
occurrent states.

Though I want to insist that represented contents exhibit psychological
involvement, I think it is premature at this point to speculate on exactly what

---

6Note that entertaining need not be conscious, and we need not be aware of all the
contents that we are entertaining. For example, when we think that the cat is on the mat,
we entertain the content <the cat is on the mat>, but if this thought is not conscious, we
need not be consciously aware of its content.

7Note that taking introspective accessibility to be a way in which contents can be
psychologically involved does not make introspection redundant as a way of knowing about
intentional states. Introspection reveals intentional contents, providing us specific data
points that a theory of intentionality should get right, while considerations of psychological
involvement provide a weaker form of evidence.
psychological roles they play. For example, it is not clear whether represented contents are causally related to each other and behavior, or if all relevant causal work is being done by the vehicles that carry them in a way that reflects the semantic relations between the vehicles’ contents. It is also an open question which contents are introspectively accessible. All I want to say, and all I will need to say for the purposes of my arguments, is that contents play some such roles.

Considerations concerning psychological involvement provide us with a theory-independent way of knowing about intentionality. If we assume that intentional states play psychological roles appropriate to their contents, we can take their observed psychological roles to provide evidence regarding their contents. Perhaps less obviously, considerations concerning psychological involvement also provide us some insight into the deep nature of intentional states, since it might turn out that intentional states have to be a certain way in order for them to be psychologically involved in the first place. Several of the arguments in this book will make use of the assumption that represented contents are psychologically involved to support claims about the superficial characters and deep nature of represented contents (see especially Chapters 3, 5, and 9).

### 2.3 Conclusion

In summary, a theory of intentionality is a theory that describes the deep nature of intentionality. Though the theory of intentionality that I will propose is much more ambitious, I will structure much of the initial discussion around the question of what gives rise to all actual instances of original intentionality. As we will see, by providing us with a theory-independent access to the contents of at least some intentional states, introspection and considerations of psychological involvement can help us settle this question.

With these preliminaries under our belt, we are now ready to begin our
search for a true theory of intentionality. Part II considers approaches to intentionality that take it to arise from tracking or functional roles, and Part III turns to my favored approach, the phenomenal intentionality theory.