

## 1

# The Experiential Origins of Intentionality

## Introduction/Abstract

The purpose of this book is to develop an account of intentionality that assigns a central theoretical role to a kind of intentionality proper to conscious experience. The purpose of the present chapter is to argue that we **should** have such an account of intentionality; that the intentionality proper to our conscious experience—“experiential intentionality”—**deserves** a central theoretical role in our account of intentionality. I will argue that our conception of intentionality is grounded in our grasp of experiential intentionality, and that to that extent, experiential intentionality is the origin of all intentionality. I call this the **thesis of the experiential origins of intentionality**, or the **experiential origin thesis** for short.

The plan of the chapter is as follows. I start, in §1.1, with a general model of how we form our conceptions of certain phenomena, including the phenomenon of intentionality. In §1.2, I present an argument to the effect that the application of the general model to the case of intentionality generates the result that our conception of intentionality is grounded in our grasp of experiential intentionality. I close, in §1.3, with some elucidations of the central notion of experiential intentionality.

Before starting, a note to the reader: in the course of the discussion, I will often present very precise formulations of central theses and definitions, in the form of indented biconditionals; by and large, these need not be read in order to follow the discussion, and the reader is herewith entreated to skip them if s/he finds no joy in going over them.

## 10 THE SOURCES OF INTENTIONALITY

I.1. The Concept of Intentionality  
and Anchoring Instances

On the assumption that intentionality is a natural kind notion, understanding how we form our conception of intentionality requires prior understanding of how we form natural kind concepts in general. This is of course an immense area of research, and we will have to treat it with considerable brevity here.<sup>1</sup> In §I.1.1, I propose a general model of the formation of natural kind concepts—or more accurately, of a certain kind of formation of a certain kind of natural kind concept—that I call the **anchoring-instance model**. In §I.1.2, I suggest an application of this general model to the concept of intentionality that leads to the experiential origin thesis. (The argument for that application will be presented in §I.2.)

## . . . The Anchoring-Instance Model of Concept Formation

In this subsection, I discuss three toy models of concept formation, defending the third. All three models are decidedly underdeveloped, as concept formation is a highly rich topic, joined to a vast literature, and there is no hope to discuss it adequately as a mere preliminary to the main subject matter of this book. However, many of the details of how each of the three models could play out do not affect the implications I will be seeking for our conception of intentionality. What I will end up endorsing may thus be better thought of as a family of models of concept formation rather than one specific model; my concern will be with the application of any model from this family to the concept of intentionality.

There is no special reason to expect that a single type of model of concept formation would cover the variety of manners in which we can form concepts. Different types of concept, in particular, may well be formed in different ways. I will focus here on the **non-parasitic** formation of **observational natural** kind concepts. Let me explain what I mean by these specifications.

For present purposes, let a natural kind concept be any concept whose instances all seem to share a “natural similarity,” that is, a similarity that is in the nature of things themselves rather than due to our faculties and/or activities.<sup>2</sup> The model I will propose for concept formation might be modified to accommodate other kind concepts, but my concern here is with the formation of **natural** kind concepts.

More specifically, my concern is with **observational** natural kind concepts, where a natural kind concept is observational (relative to a subject S) just in case it is nomologically possible (for S) to observe instances of the concepts. Thus, the concept of a horse is an observational natural kind concept (relative to the normal human subject), because we (normal human subjects) can observe horses; the concept of a lepton is a non-observational natural kind concept, because we cannot observe leptons.<sup>3</sup> The model of concept formation I will propose may or may not be correct for non-observational natural kind concepts, but in any case I am proposing it only for observational ones.

More specifically yet, my concern here is with what might be called “non-parasitic” formation of observational natural kind concepts. Consider our concept of a mammal.<sup>4</sup> As it happens, most of us formed our concept of a mammal by reading (or hearing) about mammals. But this manner of concept formation is parasitic on a more basic manner, namely, that by which those who wrote what we read (or said what we heard) about mammals formed their concept of a mammal.<sup>5</sup> Forming a concept of a mammal on the basis of something like testimony is parasitic on forming it on the basis of perceptual encounter with mammals. It is this sort of non-parasitic concept formation, involving encounter with instances that fall under the concept, that concerns me here.<sup>6</sup>

One traditional model of such concept formation goes something like this. A subject goes out into the world and encounters many particulars, including animals. Among some subsets of those particulars she notices certain systematic and apparently objective similarities, and she therefore groups them together. For example, she notices a similarity among dogs, cats, and horses: they are all furry, four-legged land animals with faces featuring two eyes, a nose, and a mouth similarly spatially arranged. On this basis, she specifies a criterion for membership in a group in terms of observable features of prospective members: being a furry, four-legged land animal with the right kind of face. This criterion, which typically consists in a number of jointly sufficient and severally necessary conditions for membership in the relevant kind, is effectively her concept of a mammal: animals the subject has not encountered qualify as belonging to the group just in case they satisfy the criterion. More generally, and to a first approximation, the model is this:

## 12 THE SOURCES OF INTENTIONALITY

- (C) For any natural kind *K* and particular *X*, there is a criterion *C*, such that *X* qualifies as a *K* iff *X* satisfies *C*.

We may call this the **crierial model**. Importantly, the “criterion” it adverts to must be specifiable in terms of observable features.<sup>7</sup>

One immediate difficulty with the criterial model, however, is that it cannot explain how we can come to **discover** that whales and bats, for example, are also mammals. On the criterial model, a particular is **discovered** to be a member of a kind by being **discovered** to satisfy the criterion. Thus, although interaction with certain nocturnal rodents (e.g., the chipmunk) is rare, and the concept of a mammal is not typically formed—not even in part—through interaction with them, one may discover later that they qualify as mammals as well, namely, by discovering that they too are furry four-legged land animals with good faces. However, the whale is none of those things, so it is unclear in what way it could be **discovered** to be a mammal. Perhaps a story could be told about how we might **decide** to count the whale as a mammal, by deciding to **change** our criterion for membership in the mammal kind. But the whale’s qualifying as a mammal was never a **decision**—it was a **discovery**. This means that our concept of a mammal was ready to subsume the whale all along, and did not need to be **changed** in order to do so.

The lesson from the criterial model’s shortcoming is that the concept of a mammal appeals to certain hidden features of mammals that are taken to go more deeply into what makes them belong together. The manifest similarity between them is always taken to be merely a symptom of a deeper underlying commonality, and our concept of a mammal adverts to this underlying commonality in classifying a given particular as a mammal.<sup>8</sup> In the first instance, the relevant hidden commonality may be still **observable**, even if not immediately, easily, or typically **observed**: being warm-blooded, giving live birth, nursing the young through a mammary gland.<sup>9</sup> But ultimately, there is a deeper unobservable commonality among mammals at the genetic level. We know, for example, that the whale evolved from the hippopotamus, and we know this because whale DNA is extremely similar to hippopotamus DNA. There is, more generally, some underlying genetic commonality among all mammals that makes them mammals (as well as a genetic peculiarity that makes other animals not mammals). This genetic commonality (and peculiarity) is the “underlying nature” of the property of being a mammal.

These observations may inspire a model of natural kind concept formation that makes the concepts appeal solely to such underlying hidden properties. On such a model, acquiring the concept of a mammal fully is a matter of coming to know the underlying nature of mammals. Let *G* be the underlying genetic commonality among mammals. According to the present model, our concept of a mammal is such that a particular qualifies as a mammal just in case it has *G*. More generally, and to a first approximation:

(UN) For any natural kind *K* and particular *X*, there is an underlying nature *N*, such that a particular *X* qualifies as a *K* iff *X* has *N*.<sup>10</sup>

Call this the **underlying-nature model**. Unlike the criterial model, it can readily account for the **discovery** that the whale is a mammal: it is simply the discovery that the whale has the right underlying nature—it has *G*.

The problem with the underlying-nature model is the opposite of the problem with the criterial model. In the criterial model, too **little** was discoverable. In the underlying-nature model, too **much** is discoverable. Thus, further research could turn up evidence that, initial appearances (and studies) notwithstanding, dogs, cats, and horses do not have *G*. By the lights of the underlying-nature model, this would effectively be evidence that dogs, cats, and horses do not qualify as mammals. For that matter, it could in principle be discovered that there are no mammals other than the whale and the hippopotamus: all other animals lack *G*. In fact, it is not clear that the model, as stated above, can rule out the possibility that it be discovered that there are no mammals at all, namely, in case it turns out that no animal has *G*.<sup>11</sup>

These are absurd consequences. The lesson here seems to be that there are **some** a priori restrictions on what can in principle be discovered regarding the extension of the concept of a mammal, hence some a priori conditions that anything must satisfy in order to qualify as a mammal. In particular, the following seems to be a priori: one's concept of a mammal is such that at least **some** (and probably **most**) of the particulars one encountered during the formation of one's concept of a mammal, and noticed to have the manifest, observable features associated with mammals, must in fact qualify as mammals. That is to say, it is a priori that some (most?) furry, four-legged land animals with good faces encountered during concept formation are mammals. It can in principle be discovered that, surprisingly, dogs do not qualify as mammals. But it **cannot** be discovered that dogs, cats, horses, cows, goats, and every other manifestly similar animal a

## 14 THE SOURCES OF INTENTIONALITY

subject encountered during the formation of her concept of a mammal do not qualify. For the formation of a concept of a mammal requires an early stage in which certain animals are grouped together and used collectively as an anchor for the construction of a kind category. The status of these anchoring instances as mammals may be negotiable individually, but not collectively: each one may turn out not to be a mammal after all, but they cannot **all** turn out not to be mammals.

These observations suggest to me a compromise model that combines the virtues of the criterial and underlying-nature models, a model where **both** observable features and underlying natures are appealed to by natural kind concepts in classifying given particulars as mammals. On a natural model, a subject goes out into the world and encounters many particulars, including many, such as dogs, cats, and horses, that are furry, four-legged land animals with good faces. On this basis alone, she forms the concept of a kind members of which have the same underlying nature as (most of) the furry, four-legged land animals with good faces that she has encountered; this is effectively her concept of a mammal. She may later discover that other animals, such as the chipmunk, are also furry, four-legged, etc. And she may later yet discover that yet other animals, such as the whale, also have the underlying nature that the original encountered animals have, even though they are not furry, four-legged, etc. But her concept of a mammal was always ready to subsume these new instances.

Let us say that the dog, the chipmunk, and the whale are each an **instance** of the concept of a mammal; that the dog and the chipmunk are, but the whale is not, a **manifest instance** of the concept of a mammal; and that the dog is, but the chipmunk and the whale are not, an **anchoring instance** of the concept of a mammal.<sup>12</sup> To a first approximation, an anchoring instance is a manifest instance that the subject encountered in the process of forming the relevant concept; a manifest instance is one that exhibits the relevant observable features.<sup>13</sup> According to the model under consideration, the concept of mammal is such that a particular qualifies just in case it has the same underlying nature as most anchoring mammals. More generally, and to a first approximation:

- (AI) For any natural kind *K* and particular *X*, there is a nature *N*, such that *X* qualifies as a *K* iff (i) most anchoring instances of *K* have *N*, and (ii) *X* has *N*.<sup>14</sup>

Call this the **anchoring-instance model**. Unlike the criterial model, it can account for the discovery that the whale is a mammal: it is the discovery that the whale has the same underlying nature as dogs, cats, horses, etc. And unlike the underlying-nature model, it can account for it being undiscoverable that no “manifest mammal” is really a mammal: for that would involve discovering that none of the anchoring instances has the same underlying nature as most of the anchoring instances—a logical contradiction.<sup>15</sup> The anchoring-instance model thus avoids the main pitfalls of its competition.<sup>16</sup>

I close this subsection with ten clarifications of (AI). The first six elucidate central notions in (AI), the last four warn about what is **not** implied by (AI).<sup>17</sup>

First, for something to “qualify” as a K is for it to meet the condition specified in (hence fall in the extension of) the concept of a K. For example, qualifying as a mammal is meeting the condition specified by (thus falling in the extension of) the concept of a mammal. By “the” concept of a mammal, I mean the concept that the normal subject would form under normal conditions. The normality of the subject can be construed either statistically or teleologically: a statistically normal subject is a subject that behaves relevantly similarly to most other subjects; a teleologically normal subject is one whose concept-forming mechanisms function as they are supposed to.<sup>18</sup> The teleological notion of normality is probably the more relevant one here, but presumably the anchoring-instance model itself is neutral on such matters. It may be that an informed-ness condition should be added here: “the” concept of a mammal is the concept that a normal **well-informed** subject would form under normal conditions (here an elucidation of “well-informed” would have to be appended). Also, it is possible that in some areas the normality of the subject would have to be relativized in certain ways (e.g., to a culture or a time); the general shape of the account would not be overly affected.

Secondly, there are far-from-trivial questions surrounding what a “concept” is. Two main approaches treat the concepts as abstract entities or alternatively as mental entities. Here I have in mind concepts as mental entities, but it may well be that the anchoring-instance model can operate with concepts as abstract entities as well. In any case, what kind of mental or abstract entity a concept is is something the model is neutral on. One common and congenial approach individuates concepts (construed as mental entities) in terms of their possession conditions (see Peacocke 1992). Of course, what

## 16 THE SOURCES OF INTENTIONALITY

it takes for the subject to possess a concept is no less debatable. Presumably, she needs to have certain discriminatory, recognitional, and classificatory abilities, and probably further facts must hold of her as well. I would insist that, to possess a concept, the subject need not know any associated public-language noun, or be aware of herself as possessing the concept (or for that matter understand the notion of a concept). What exactly is required is an issue that needs to be resolved independently of the anchoring-instance model, and its resolution could be plugged into the model to generate a more explicit account.

Thirdly, the “formation” of a concept is presumably the process that takes the subject from a state of not possessing the concept to a state of possessing it. One natural view is that this process takes as many months or years as the infant or toddler requires before acquiring the relevant abilities (discrimination etc.). A very different view is that there is no clean-cut division between an initial stage of concept formation and a later period of mere concept “maintenance,” but rather a single “dynamic,” cascading process of ever-evolving, always-live concept formation and maintenance, such that the process of concept formation can never be truly said to have ended. An intermediate view would let the relevant process remain “live” for many years, but still weigh early stages more heavily than later ones (and perhaps let it die out at some point). The anchoring-instance model is neutral on which is the best option here. It should simply recruit whichever view generates the most accurate predictions and retrodictions for the model as a whole.

Fourthly, the notion of “encounter” is fairly straightforward in its application to **tokens**, but has an interesting application to **types** that is less straightforward. One might say that one encountered dragonfruit in one’s travels in the East, with the thought that one is now familiar with the **type** dragonfruit thanks to one’s seeing and touching **token** dragonfruits. The term “encounter” can be used both for the seeing/touching of the tokens and for the ensuing familiarity with the type. Consequently, when we speak of encounter with dogs, cats, and horses, we may have in mind either token-encounter or type-encounter. Since anchoring instances are manifest instances encountered during concept formation, there are consequently two ways of understanding what counts as an anchoring instance. The stricter view would require token-encounter, counting as anchoring instances of S’s concept of a mammal only those individual manifest mammals that S “met”

(this particular cat, that particular horse). The more laid-back view would include also individual mammals manifestly belonging to the relevant sub-kind (all cats, all horses), thus requiring only type-encounter of anchoring instances.<sup>19</sup> This second approach strikes me as the most likely to return the right results with respect to who possesses which concepts, but again the anchoring-instance model need not be definitionally tied to this particular take on encounter rather than the other.

A fifth issue concerns what counts as “most” of the anchoring instances. One could require a simple majority, or alternatively require some substantial majority, say 80 percent. Here too, the theoretical choice between the many options should be made by considering which option is such that, when it is plugged into (AI), (AI) returns the right results on concept possession.

Sixthly, (AI) is formulated in terms of “particulars,” a notion with straightforward application for natural kind concepts canonically expressed by count nouns (such as “mammal”). The application is less clear when it comes to natural kind concepts canonically expressed by mass nouns, notably substance terms such as “water” and “gold.” However, for such natural kind concepts, the thesis can be reformulated in terms of samples, or quantities, or some such type of entity. More conveniently, we may simply allow the term “particular” to range over those types of entity as well. In this (admittedly technical) usage, samples and quantities are particulars.

Seventhly, although the model is tailored to cases in which the underlying nature of anchoring instances is uniform and clear-cut, it can be extended to apply to cases where the situation is murkier. Thus, suppose that some natural kind concept is such that, as it turns out, no majority of its anchoring instances has any underlying commonality. Different minority subsets of the anchoring instances may have local commonalities, perhaps with family-resemblance overlaps, but no majority subset has a commonality. How does the anchoring-instance model treat this case? There are several ways to go here, it seems to me, but the most natural would be to say that there is no natural kind in the area, and that therefore the relevant concept is an empty natural kind concept or a non-natural (“manifest”) kind concept.<sup>20</sup> In another scenario, there might be a natural kind concept such that, as it turns out, all of its anchoring instances have two independent (causally and explanatorily insulated) underlying natures  $N_1$  and  $N_2$ , but some non-anchoring particulars have only  $N_1$  and others only  $N_2$ . How

## 18 THE SOURCES OF INTENTIONALITY

does the model handle this scenario? Again, there are several ways to go, but to my mind the most natural are to say either (a) that there are two separate natural kinds here, and thus the concept is ambiguous, or (b) that the concept unambiguously picks out a conjunctive natural kind.<sup>21</sup> There may be other challenging scenarios, but I see no reason to think that any would render the anchoring–instance model predictively impotent.

Eighthly, observe that the model does **not** imply that all anchoring instances are genuine instances—only that most are. For it is true of any anchoring instance that it can turn out not to be an instance at all. This way of putting things may have an initial air of paradox about it, but in fact is perfectly coherent. Consider Twardowski's (1894) distinction between determining and modifying adjectives: whereas a brown car is always a car, and so “brown” functions here as a **determining** adjective, a fake barn is never a barn, and so “fake” functions as a **modifying** adjective. We can introduce a third kind of adjective, the “neutral adjective,” which guarantees neither determination nor modification of the predicate it accompanies—sometimes it determines, sometimes it modifies. The adjective “rubber” is a neutral adjective: a rubber tire is a kind of tire, but a rubber duck is not a kind of duck. In the present context, the adjective “anchoring” functions as a neutral adjective in the same sense: some anchoring instances are instances, but some are not. Whether a given anchoring instance is an instance depends on whether its underlying nature is the same as the underlying nature of most of the anchoring instances it is one of. Hence the anchoring–instance model does not prejudge whether a given anchoring instance is indeed an instance. At the same time, it does prejudge that **most** of the anchoring instances are instances (or else that the concept is empty). Observe that this makes the anchoring–instance model non-circular and reductive: it does not account (circularly or non-reductively) for being an F in terms of being a special sort of F (an “anchoring” F), but rather accounts for being an F in terms of being a G (an “anchoring-F”).<sup>22</sup>

Ninthly, anchoring instances should not be confused with prototypes—the theoretical role of the former in the present model is quite different from that of the latter in prototype theories of the structure of concepts.<sup>23</sup> There are similarities between the two, but also important differences. The main similarity is that both anchoring and prototypical instances are proper subsets of all instances, set apart from, and in some sense assigned some priority over, other instances.<sup>24</sup> The differences go much deeper, however.

Most notably, some anchors are not prototypes: a bed is an anchoring instance in most subjects' concept of furniture, but appears not to be a prototypical instance.<sup>25</sup> For it appears that beds do not score high on typicality tests (Rosch 1975), even though they are manifest instances of furniture one normally encounters when forming one's concept of furniture. More generally, there are many more anchors than prototypes: whereas chairs and tables are plausibly the only prototypes of furniture, anchors include beds, cabinets, bookcases, and more. The deeper issue here is that anchoring instances enjoy a special **epistemological** status, but no special **metaphysical** status, whereas prototypical instances enjoy a special metaphysical status as well. Thus an anchoring instance of some kind K is in no sense **more of an instance** of K than a non-anchoring instance, whereas a prototypical instance is in some sense more of an instance.<sup>26</sup>

Tenthly, though, it is important to appreciate that the anchoring-instance model of the **formation** of natural kind concepts is consistent with a prototype model of the **structure** of such concepts.<sup>27</sup> In fact, it is consistent with many different theses on the metaphysics of natural kinds and the semantics of natural kind terms and concepts. Perhaps most importantly, it is compatible with the main doctrines currently widely accepted among philosophers about natural kinds and natural kind terms/concepts: that natural kinds are "real" and "objective" (as per realism about natural kinds) that have scientifically discoverable essences (as per scientific essentialism), and that natural kind terms and concepts are rigid designators that allow for a posteriori necessary truths (as per Kripkean semantics).<sup>28</sup>

Hopefully these remarks make the anchoring-instance model sufficiently clear to proceed. The remarks should be taken as **preliminary**, and the model as a **framework** in need of much further elaboration that might enrich it considerably and/or modify aspects of it importantly. As noted at the opening, it is in fact best thought of as a **family** of views. As long as further elaborations and modifications do not undermine certain essential elements of the model, the limited exposition offered here could be used in an application to the concept of intentionality. This is the task of the next subsection.

### . . . Application to the Concept of Intentionality

According to the anchoring-instance model, a particular qualifies as an instance of some observational natural kind concept just in case it has the same underlying nature as most anchoring instances of that concept, i.e., has

## 20 THE SOURCES OF INTENTIONALITY

the same underlying nature as most of the manifest instances the subject encountered during the formation of her concept of that kind. This model applies to intentionality on the assumption—highly plausible—that the concept of intentionality is an observational natural kind concept.<sup>29</sup> The application yields that an item qualifies as intentional just in case it has the same underlying nature as most anchoring instances of intentionality. Now: if we add to this that all anchoring instances of the concept of intentionality are experiential-intentional states, we obtain that an item qualifies as intentional just in case it has the same underlying nature as most experiential-intentional states, or more accurately, just in case it has the same underlying nature as most of the experiential-intentional states that serve as anchoring instances of intentionality.

In the next section, I will argue that we should add this extra thesis: all the anchoring instances of (the concept of) intentionality are experiential-intentional states.<sup>30</sup> With this thesis in hand, we will be in a position to use the anchoring-instance model and the claim that the concept of intentionality is an observational natural kind concept to reason as follows:

1. When K is an observational natural kind concept, an item X qualifies as a K iff X has the same underlying nature as most anchoring instances of K;
2. The concept of intentionality is an observational natural kind concept; therefore,
3. An item X qualifies as intentional iff X has the same underlying nature as most anchoring instances of intentionality;
4. There is a subset of experiential-intentional states, such that an item X is an anchoring instance of intentionality iff X is a member of that subset; therefore,
5. There is a subset of experiential-intentional states, such that an item X qualifies as intentional iff X has the same underlying nature as most members of that subset.<sup>31</sup>

Note that this presentation of the reasoning does not commit to whether the subset of experiential-intentional states that serve as anchoring instances of intentionality is a proper or improper subset. For it may well be that although all anchoring instances of intentionality are experiential-intentional states, not all experiential-intentional states are anchoring instances of intentionality.<sup>32</sup>

Recall that an anchoring instance is a manifest instance encountered during the process of the concept's formation. So there are two ways in which some experiential-intentional states might not be anchoring instances of intentionality. One is by failing to exhibit the manifest features of intentional states. This is admittedly not very plausible. The other way is much more plausible, however: it might be that some experiential-intentional states are not **encountered** instances, instances that have been encountered during the (normal subject's non-parasitic) formation of the concept of intentionality.<sup>33</sup> For example, during the formation of my conception of intentionality, I had not yet undergone, hence not yet encountered, gustatory experiences as of snake meat or as of sea urchin sushi.<sup>34</sup> These intentional experiences came relatively late in life, and so are not encountered instances in any relevant sense. Quite plausibly, then, not all experiential-intentional states are anchoring instances of intentionality, but only a **proper** subset thereof, namely, the set of **encountered** experiential-intentional states.<sup>35</sup> If so, we can formulate the above reasoning more precisely, and more pointedly, as follows:

1. For any item  $X$  and observational natural kind  $K$ ,  $X$  qualifies as a  $K$  iff there is a nature  $N$ , such that (i) most of the anchoring instances of  $K$  have  $N$  and (ii)  $X$  has  $N$ ;
2. The concept of intentionality is an observational natural kind concept; therefore,
3. For any item  $X$ ,  $X$  qualifies as intentional iff there is a nature  $N$ , such that (i) most of the anchoring instances of intentionality have  $N$  and (ii)  $X$  has  $N$ ;
4. For any item  $X$ ,  $X$  is an anchoring instance of intentionality iff  $X$  is an encountered experiential-intentional state; therefore,
5. For any item  $X$ ,  $X$  qualifies as intentional iff there is a nature  $N$ , such that (i) most encountered experiential-intentional states have  $N$  and (ii)  $X$  has  $N$ .

As noted, the key premise in this reasoning is 4, and it will be defended in the next section.

The conclusion of this reasoning can be thought of as one possible formulation of the experiential origin (EO) thesis, since it grounds our conception of intentionality in our grasp of experiential intentionality. The formulation is this:

## 22 THE SOURCES OF INTENTIONALITY

- (EO) For any item  $X$ ,  $X$  qualifies as intentional iff there is a nature  $N$ , such that (i) most encountered experiential-intentional states have  $N$  and (ii)  $X$  has  $N$ .

There is a straightforward sense in which (EO) captures the **spirit** of the experiential origin thesis, and can thus legitimately serve as its **letter**.

Since (EO) is an application of the anchoring-instance model to the concept of intentionality, the clarifications offered above of the general model apply here as well. First, for an item to qualify as intentional is for it to fall in the extension of the concept of intentionality, where this should be taken as short for “the normally circumstanced (teleologically) normal subject’s concept of intentionality.” Secondly, different versions of (EO) will have different views on the ontology of the concept of intentionality, and thirdly, different versions will have different views on what is involved in the concept’s formation. Fourthly, the set of “encountered” experiential-intentional states can be construed either as the set of encountered tokens or as the set of tokens that belong to the encountered types (and my suspicion is that the latter construal would return the better results).<sup>36</sup> Fifthly, different versions of (EO) will differ on how many encountered experiential-intentional states count as “most.” Sixthly, intentional items—items with intentional properties—are all clearly particulars.<sup>37</sup> Seventhly, (EO) does not rule out the possibility that the concept of intentionality is an empty concept or a conjunctive concept. Eighthly, (EO) does not imply that all anchoring instances of intentionality are really instances, only that most are. Ninthly, to say that only experiential-intentional states are anchoring instances of the concept of intentionality is not to say that only experiential-intentional states are **prototypical** intentional states—though these two theses would be somewhat similarly inspired, and the latter thesis may well be independently plausible.<sup>38</sup> Tenthly, (EO) is compatible with any number of rather popular claims about the metaphysics of intentional properties and the semantics of intentional terms, including that the former are “real” and “objective” and that the latter are rigid designators.

In addition, one more central notion must be elucidated in (EO) that was not part of (AI), namely, the notion of an “experiential-intentional” state. I will dwell on this more patiently in §1.3.1, but the general idea is that a mental state is experiential-intentional when it is intentional in virtue of being experiential. It is not self-evident that there are such states

(see discussion in §I.3.2), but perceptual experiences are natural candidates: a visual experience as of a pear or a tulip, an auditory experience as of trumpets, and an olfactory experience as of ground coffee are all intentional experiences. Emotional and somatic states, such as joys and itches, are clearly often experiential, though their status as intentional is subject to debate. Conversely, cognitive and conative states, such as thoughts and intentions, are typically intentional, but their status as experiential is uncertain (see discussion in §I.3.3). Depending on how the relevant debates are settled, these may or may not turn out to be potential anchoring instances of intentionality if (EO) is true. The case for (EO) that I will make in §2 is supposed to be insensitive to the scope of experiential intentionality, though obviously the project of grounding all intentionality in experiential intentionality is liable to go more smoothly the more inclusive the prospective grounds.

It will be noted that the experiential origin thesis, as construed here, is in the first instance an **epistemological** thesis: a thesis about the grounding of the **concept** of intentionality, not the **property** of intentionality. It does have a straightforward implication for the property, namely, well, that for an item to instantiate the property picked out by the concept of intentionality, it must have the same underlying nature as most experiential-intentional states. But at its core, the experiential origin thesis is a thesis about our grip on, hence epistemic relation to, the phenomenon of intentionality. A more immediate implication of its is for theory construction: if our grip on intentionality is indeed based upon our grasp of experiential intentionality, then the theory of intentionality ought to be organized, in some sense, around our understanding of the phenomenon of experiential intentionality.

It should also be noted that the experiential origin thesis is not by itself in any tension with the notion that the origin of intentionality is a certain kind of tracking relation. For if the experiential-intentional states that function as anchoring instances of intentionality turn out to have their experiential-intentional content in virtue of tracking the environment, it would be true of the origin of intentionality both that it is experience-based and that it is tracking-based. Indeed, the prospects for such a marriage of experience-first and tracking-first approaches will be taken up in the next chapter.<sup>39</sup>

In the reasoning leading to (EO), Premise 4 is clearly the key: only experiential-intentional states are anchoring instances of intentionality. This will be defended in the next section. Premise 1 is the anchoring-instance

## 24 THE SOURCES OF INTENTIONALITY

model, motivated in the previous subsection, and Premise 2—that intentionality is an observational natural kind concept—is a relatively uncontroversial claim. It could be resisted by denying either the existence of the concept, or its being a natural kind concept, or its being an observational natural kind concept. But none of these options is particularly plausible.

It is certainly hard to imagine that we have no concept of intentionality. Recall that talk of “the” concept of X should be understood as short for “the normally circumstanced normal subject’s concept of X.” It is implausible that the normal subject has no concept of the intentional. She may not have the **term** “intentionality,” but she has a battery of other terms—most notably, “about” and “directed”—with which she can express her concept(ion) of intentionality. Without such a conception, she would be unable, for example, to understand simple traffic signs, which in fact she does understand.<sup>40</sup>

It is not much more plausible to deny that the concept of intentionality is not a natural kind concept or an observational kind concept. Recall that an observational natural kind concept, as understood here, is a kind concept whose instances resemble independently of us but can be observed by us (thus are observable but enjoy an observer-independent similarity). Some intentional items are **perceptually** observed, some are only **introspectively** observed, and some are unobserved (perhaps even unobservable). But for the concept of intentionality to qualify as an observational kind concept, it suffices that it be nomologically possible for us to observe **some** instances of it, in one way or another.<sup>41</sup> Meanwhile, on the face of it, it is clear that the concept of intentionality is also a natural kind concept, since the commonality among its instances—that peculiar directedness one finds in all intentional items—does not appear to be due to our activities or institutions, or to be in any sense “projected” from our thinking on intentionality or our practices of intentional ascription.<sup>42</sup> (Furthermore, if the concept of intentionality is not a natural kind concept, then the role of the anchoring instances is liable to become even greater: in all likelihood, something would qualify as intentional just in case it was manifestly, or superficially, similar to them, instead of just in case it had the same underlying nature as them.<sup>43</sup> Thus there is hardly a danger of **over**playing the role of experiential-intentional states in proceeding as though the concept of intentionality is a natural kind concept.)

To summarize, the purpose of this subsection has been to propose an application of the anchoring-instance model to the concept of intentionality

that would pave the way for an anchoring of all intentionality in experiential intentionality. The case for this depends on two additional assumptions. One is that the concept of intentionality is the kind of concept the anchoring-instance model is supposed to cover, which I have just argued. The other is that all anchoring instances of the concept of intentionality are instances of experiential intentionality, which is argued next. If both assumptions are accepted, it follows that an item is intentional just in case it has the same underlying nature as most experiential-intentional states the normally circumstanced normal subject encounters during her formation of her conception of aboutness—and to that extent, that an item is intentional just in case it is appropriately related to experiential-intentional states.

### 1.2. Experiential Intentionality the Anchor

The anchoring instances of the concept of intentionality are instances encountered during the concept's formation. Given that the concept is observational, presumably encounter here means something like **observational contact** with the encountered. My contention is that the only instances of intentionality with which we have observational contact are experiential-intentional states. We may therefore argue as follows:

1. All the anchoring instances of intentionality are such that we have observational contact with them;
2. The only instances of intentionality with which we have observational contact are experiential-intentional states; therefore,
3. All the anchoring instances of intentionality are experiential-intentional states.

This is effectively a sub-argument for Premise 4 in the argument for (EO) in §1.1.2.<sup>44</sup>

The first premise of this argument falls out of the anchoring-instance model. An objector might protest that, in light of where this seems to lead us, perhaps it was a mistake to put so much weight on encounter with items that fall under a concept. The objection is essentially that there is an undue empiricism creviced in the anchoring-instance model of concept formation. Thus what has to be abandoned is the anchoring-instance model, or perhaps just the part of the model that requires anchoring instances to be encountered.<sup>45</sup>

## 26 THE SOURCES OF INTENTIONALITY

In response, however, we must keep in mind that the anchoring-instance model is supposed to apply only to observational concepts—concepts whose instances **can** be observed. I would insist that while we may form certain natural kind concepts not on the basis of encounter with instances, this surely characterizes exclusively concepts instances of which we **cannot** observe. It is extremely odd to think that we may (non-parasitically) form the concept of a kind instances of which we can and routinely do encounter but form it independently of our encounter with those instances.

The controversial premise in the sub-argument is the second: why think that experiential-intentional states are the only ones we can have observational contact with? On the one hand, it seems to me entirely natural to suppose that this is so. When an image as of a smiling octopus suddenly and involuntarily pops into my mind, I am aware of my octopus experience, and seem to be aware of it in a direct, non-inferential, and in some admittedly elusive sense observational manner. However, while this seems to me a supposition so natural as to defy doubt, I recognize that others have found it very much doubtful, indeed have found the whole notion of observational contact with conscious experience overly mysterious. The purpose of this section is to provide **an argument** for the claim that we have direct observational contact with some of our experiential-intentional item but with no other intentional states.

The argument is from inference to the best explanation. I will first argue that there is a certain asymmetry in our ascription of experiential-intentional states to ourselves, on the one hand, and our ascription of any other kinds of intentional state, on the other (§1.2.1). I will then argue that the best explanation of this asymmetry is that we have observational contact with our own experiential-intentional states but with no other intentional items (§1.2.2).<sup>46</sup> I will close with discussion of some possible objections (§1.2.3).

### . . . An Asymmetry of Ascription

To a first approximation, the asymmetry I have in mind is that while all other intentional ascription requires the deployment of normative principles of charity, ascription of experiential-intentional states to oneself does not.<sup>47</sup> That is, while the ascription of non-experiential-intentional states is **always** governed, and thus mediated, by a normative principle of charity, the ascription of experiential-intentional states is sometimes not.<sup>48</sup> This

asymmetry is grounded in a more specific one, which will come out in the discussion to follow.

As our starting point, let us take Davidson's claim that ascription of intentional states to persons is supposed to maximize the intelligibility of these persons' overt behavior, and therefore competent ascription is governed by a cluster of normative principles sometimes loosely referred to as "the principle of charity."<sup>49</sup> These include the principles that (by the interpreter's own lights) persons' beliefs are mostly true and coherent, that their desires are mostly good and mutually satisfiable (and/or suitably prioritized), and that their beliefs and desires mostly conspire to constitute good reasons for action.<sup>50</sup>

The reason for claiming that competent intentional ascription is governed by such principles is not always clear in Davidson's writings, but does come through quite nicely in the following passage (Davidson 1974: 18):

If you see a ketch sailing by and your companion says, "Look at that handsome yawl," you may be faced with a problem of interpretation. One natural possibility is that your friend has mistaken a ketch for a yawl, and has formed a false belief. But if his vision is good and his line of sight favorable it is even more plausible that he does not use the word "yawl" quite as you do, and has made no mistake at all about the position of the jigger on the passing yacht.

From which Davidson concludes (*Ibid.*; italics mine):

[I]f we merely know that someone holds a certain sentence to be true, we know neither what he means by the sentence nor what belief his holding it true represents. His holding the sentence true is thus **the vector of two forces**.

Confronted with a situation in which someone exclaims "Look at this tiger!" while pointing at (what we know is) a pigeon, two coherent interpretations are open to us: (a) that this person believes a tiger is present, and takes the word "tiger" to express the concept of a tiger; (b) that the person believes a pigeon is present, and takes the word "tiger" to express the concept of a pigeon.<sup>51</sup> Two points should be appreciated in this example. First, any competent interpreter would choose the second interpretation. Indeed, we may reasonably consider it a constraint on the competency of an interpreter that s/he choose (b) in a situation such as this. Secondly, there is nothing in the data at the interpreter's disposal (i.e., the exclamation) to favor either interpretation. What favors (b) is only the principle of charity. Thus the reason to

## 28 THE SOURCES OF INTENTIONALITY

hold that intentional ascription is governed by principles of charity—that is, that principles of charity must be operative in competent interpretation—is that without them interpretation is wildly underdetermined.

This observation applies not only to verbal behavior, but to all behavior. As Dennett (1971) notes, if we see a person opening her umbrella when it starts raining, we instantly ascribe to her the desire to stay dry and the belief that opening the umbrella will help. But our data are fully consistent with ascribing to her a desire to get wet and a belief that opening the umbrella will further that cause. What makes us ascribe to her the first belief-desire pair rather than the second has nothing to do with the behavioral data, which are also the vectors of two forces (as Davidson puts it in the passage above). It has to do rather with the fact that (by our lights) the first belief is true and the first desire is good (in the relevant sense), while the second belief is false and the second desire bad.<sup>52</sup> That is to say, what makes us ascribe to her the first belief-desire pair is charity.<sup>53</sup>

In fact, it is natural to treat the case of verbal behavior as just a special case. In the general case, the data for competent intentional ascription—that on the basis of which we ascribe intentional states—are the vectors of two intentional forces, a cognitive force (in the form of a belief or some other doxastic attitude) and a conative force (in the form of a desire or some other pro attitude).<sup>54</sup> The verbal case is one in which the desire is to perform a certain linguistic act (e.g., express or communicate a belief that *p*) and the belief is that uttering certain words is likely to achieve that (e.g., manage to express the belief that *p*).<sup>55</sup> Thus, under the interpretation we would naturally adopt, insofar as we are competent interpreters, the person who says “Look at that tiger!” while pointing at a pigeon wants to give voice to her belief that a pigeon is present and believes that uttering “Look at this tiger!” will achieve that.

The reason to take intentional ascription to be inherently normative, then, is what we may call the **vector-of-two-forces observation**. The observation is that the data upon which intentional ascriptions are based cannot typically decide between a number of competing ascriptions. In an ordinary situation calling for the ascription of an intentional state, there is typically one correct (i.e., competent) ascription to make, but a great many alternative ascriptions fully and equally consistent with the data at the interpreter’s disposal. That is, many very different intentional states can be ascribed consistently with the data, but typically only one can be so ascribed **competently**.<sup>56</sup> Davidson

infers—very reasonably, it seems to me—that principles of charity must be operative in such a situation, and serve to rule out inappropriate ascriptions that are nonetheless consistent with the data.<sup>57</sup>

What I want to argue, however, is that while the Davidsonian insight is cogent, it applies only to the ascription of non-experiential intentionality, as well as the ascription of experiential intentionality to others, but not to the ascription of experiential intentionality to oneself. More precisely, my claim is that although most intentional ascription is indeed based on data that are the vectors of two forces, and thus requires the deployment of principles of charity, ascription of experiential-intentional states to oneself relies on data that are *not* the vectors of two forces, and thus does not involve principles of charity. I further want to argue that the best explanation of this is that ascription of experiential-intentional states to oneself is based on observational contact with the states ascribed, whereas other ascription is not.

To streamline the discussion, let us introduce the terms “exp-intentional” for experiential-intentional and “nexp-intentional” for non-experiential-intentional and distinguish explicitly four kinds of intentional ascription:<sup>58</sup>

- a. ascription of non-experiential-intentional states to another (“third-person nexp-intentional ascription”)
- b. ascription of non-experiential-intentional states to oneself (“first-person nexp-intentional ascription”)
- c. ascription of experiential-intentional states to another (“third-person exp-intentional ascription”)
- d. ascription of experiential-intentional states to oneself (“first-person exp-intentional ascription”)

My claim can now be put as follows: while (a)–(c) are based on data that are the vectors of two forces, (d) is not. This asymmetry calls for explanation, and I now turn to argue that the best explanation is that (d), and only (d), is based on observational contact with the intentional state ascribed.<sup>59</sup>

### . . . Explaining the Asymmetry

In this subsection, I want to examine more closely each of the four types of intentional ascription distinguished above, and claim that the crucial asymmetry I am claiming is best explained in terms of the existence of observational contact with exp-intentional states only. For each type of ascription, I will consider simple examples that bring out particularly crisply the central

## 30 THE SOURCES OF INTENTIONALITY

model for how the relevant type of ascription works.<sup>60</sup> We will see that the mechanics of first-person exp-intentional ascription are crucially different from those of other intentional ascriptions.

Let us start with a prototypical case of third-person ascription. Suppose a person says “It’s a nice day” and on that basis you ascribe to her the desire to communicate the belief that it is a nice day, and the belief that just that concatenation of sounds will manage to communicate that belief.<sup>61</sup> In what sense are your data for this ascription the vectors of two forces? The only relevant datum here is the person’s utterance, her verbal behavior. The utterance of “It’s a nice day” is the vector of two forces in that neither the belief by itself nor the desire by itself could causally explain it. Combined with a different desire, the same belief would cause a different utterance, and combined with a different belief, the same desire would cause a different utterance. For example, combined with the desire to mislead, the same belief might result in uttering “It’s an ugly day”; combined with the belief that “day” means night and “night” means day, the same desire would result in uttering “It’s a nice night.” It is only the conspiracy of a belief and a desire that can cause the utterance, never a belief in isolation from any desire or a desire in isolation from any belief. Accordingly, the datum at your disposal offers no support for ascription of the belief in isolation from the desire or the desire in isolation from the belief.<sup>62</sup> More generally, it is impossible to use the datum at your disposal to ascribe to the person any **single** state; only pairs of complementary states can be ascribed by way of explaining her verbal behavior. It is in this sense that your datum is the vector of two forces.<sup>63,64</sup>

It is clear that a case such as this is one of third-person ascription, but there is an interesting question as to whether it is one of exp-or nexp-intentional ascription. On the one hand, the concepts of belief and desire are the main currency of traditional discussions of intentionality that have tended to disregard the difference between experiential and non-experiential intentionality and treat all intentionality as cut from the same cloth. On the other hand, upon examination of a specific case such as this, it seems to me that when a person exclaims “It’s a nice day,” she is giving voice to a conscious occurrent thought, and as will be discussed in §1.3.3, there are good reasons to hold that such thoughts have a proprietary experiential character and are thus, in one important sense, **cognitive experiences**. The same is true of this person’s conscious desire to share her thought, which conscious desire has

an experiential character (a felt quality of pull to action that comes with a distinct and varying dimension of experiential intensity) and may be referred to as a **conative experience**. My own tendency, then, would be to construe the case just discussed as involving third-person **exp**-intentional ascription.<sup>65</sup> In any event, we can consider other cases of third-person intentional ascription that belong more clearly to one category or the other.

Consider first a clearer case of third-person **nexp**-intentional ascription. Suppose that during a departmental meeting one of your colleagues rails with furious fervor ostensibly against a rather meaningless new college procedure. The colleague and you are quite close, and you know that she is in general a fragile, sadly bitter, but fundamentally well-meaning and good-hearted person who is occasionally frequented by free-floating anger energies, a predicament she seems to cope with by seeking vaguely proper objects of indignation and moral outrage to latch her preexisting anger energies onto. This coping mechanism allows her to evince her anger without perturbing her self-conception as a good person, and do so in a socially acceptable manner. All of this, let us suppose, is entirely opaque to the colleague herself, despite years of therapy. Witnessing her latest tirade, you ascribe to her the desire to evince some new anger energies while preserving her favorable self-conception, together with a belief that launching this tirade would accomplish that. In this case, it is clear that you ascribe to your colleague an unconscious belief-desire pair that exhibits only non-experiential intentionality. It is also clear that the datum at your disposal—her tirade—is the vector of two forces: the desire to evince anger energies without sacrificing a positive self-conception would not lead to the tirade if it were combined, for example, with the belief that the tirade would actually compromise that self-conception; nor would the belief that such a tirade would successfully evince anger energies without compromising the self-conception lead to the tirade if it were combined, say, with the desire to avoid evincing any anger energies.<sup>66</sup>

Consider next a clear case of third-person **exp**-intentional ascription.<sup>67</sup> Suppose you ascribe to **another** a visual experience as of a table. One way you might come to ascribe such an experience is by positing it as cause of table beliefs which, in conspiracy with conversation desires, cause the person's table conversation. In such a case, the ascription behaves essentially as third-person **nexp**-intentional ascription does, and the data are clearly vectors of two forces.<sup>68</sup> Another way you might come to ascribe to her

## 32 THE SOURCES OF INTENTIONALITY

the table experience is by observing the table itself and the position of the person's visual organs, and concluding, perhaps through simulation, that the person must be seeing the table. This kind of ascription may or may not behave completely differently, but in any case it seems to be even more immediately imbued with charity: you assume that the person's perception is **veridical**—and do so without evidence. (You could certainly **produce** evidence, e.g. by probing the person for verbal report, but such evidence would consist in just more behavioral data, which would of course be vectors of two forces.) It appears, then, that third-person exp-intentional ascription is either based on data that are the vectors of two forces or even more immediately governed by principles of charity.

I conclude that all third-person intentional ascription—both exp- and nexp-intentional—is based on data that are the vectors of two forces, or calls even more immediately for charitable interpretation. On a natural model of how such ascription works, this is due to the fact that the beliefs and desires we ascribe to others, whether exp- or nexp-intentional, are posited as behind-the-scene causes of behavior, including verbal behavior. The data at the ascriber's disposal are bits of behavior, which are taken to be the effects of hidden mental (indeed intentional) causes. Thus ascription of intentional states to others is based on **causal inference**: we make inferences about others' intentional states on the basis of observing behavior in the same way we generally make inferences about hidden causes on the basis of observed effects. Just as on the basis of observing smoke you infer that there is unobserved fire, so on the basis of observing the person's utterance you infer that she has certain unobserved intentional states.

Consider, by contrast, a prototypical case of first-person exp-intentional ascription. Suppose you ascribe to yourself a visual experience as of a table. The basis on which you make this ascription, at least in central cases, does not consist in behavioral data of any sort. It is not as though you notice a certain behavior on your part and infer that you must be undergoing a visual experience as of a table. (Perhaps there are cases of first-person exp-intentional ascription that do work like this, but those are surely the pathological or abnormal cases, and in any case not **all** the cases.) It is not immediately clear what should be considered the datum or data of this kind of ascription, but it **is** rather clear that you **can** on the basis of the data at your disposal—whatever they turn out to be—ascribe to yourself an experience as of a table **without at the same time ascribing to yourself any other state**.<sup>69</sup> In

particular, there is no desire-like state (a pro attitude) you must ascribe to yourself in tandem with your experience.<sup>70</sup> You can on the basis of your data ascribe to yourself a **single** intentional state; no conspiracy with other intentional states is needed to explain anything.<sup>71</sup>

If this is true, then the mechanics of ascription must be different in this case from the case of third-person (exp- and nexp-)intentional ascription. That is, the ascription cannot be based here on causal inference from observed (behavioral) effects to hidden (intentional) causes. The natural model is rather as follows. Insofar as it is appropriate to speak of data for ascription here, the only relevant datum seems to be a certain deliverance of introspection, namely, that it introspectively seems to you that you are having a table experience. Explaining this introspective deliverance does not seem to call for the positing of any intentional state other than the table experience. In fact, no substantive inference at all is needed to ascribe to yourself the state which it introspectively seems to you that you are in. Rather, all that is required is what is sometimes referred to in the case of perception as **endorsement**. Just as you form the perceptual judgment that there is a chair before you because (a) it **perceptually** seems to you that there is and (b) you endorse this perceptual seeming, so you form the introspective judgment that you are having an experience as of a table because (a) it **introspectively** seems to you that you are and (b) you endorse this introspective seeming.<sup>72</sup> In both cases, something seems to be the case and you endorse the seeming. In this kind of intentional ascription, then, your data are certain deliverances of introspection, and once the deliverances are endorsed, you ascribe an exp-intentional state to yourself; or perhaps the endorsing of such deliverances **constitutes** the ascribing.

There is thus a contrast between the mechanics of first-person exp-intentional ascription and third-person (exp- and nexp-)intentional ascription.<sup>73</sup> The former is based on endorsement of introspective seemings, the latter on causal inference from behavior. This is hardly deniable: as noted, when you ascribe to yourself a perceptual experience as of a table, you do not observe putative causal effects of your experience and infer on their basis the existence of a hidden experiential cause. Rather, you seem to make the ascription on the basis of observing, in some (not unproblematic) sense, the experience itself—observing, that is, the very state which you ascribe.<sup>74</sup> And this is **why** your experience of the table **can** yield the relevant deliverance all by itself and in isolation from any other intentional state.<sup>75</sup> More generally,

## 34 THE SOURCES OF INTENTIONALITY

this is why the relevant introspective deliverances are **not** the vectors of two forces. In this way, the “hypothesis” that we have introspective observational contact with our own exp-intentional states **explains** the fact that the data for these states’ ascription are not vectors of two forces.<sup>76</sup>

It is important to stress that, in saying this, one need not be indulging in any mystery-mongering about first-person access. Introspection has been claimed sometimes to be infallible, such that if it delivers that **p**, then **p**; sometimes to be immune to error of a certain type (e.g., through misidentification), such that if it delivers that **p**, and further conditions are met, then **p**; sometimes to be incorrigible, such that if it delivers that **p**, then the deliverance is incontrovertibly justified; sometimes to be self-intimating, such that if **p**, then it delivers that **p**; and so on.<sup>77</sup> None of these claims follows, however, from the assertion that we have introspective observational contact with our intentional experiences. Perceptual observational contact with perceivables is not commonly thought to secure any of these extraordinary epistemic privileges, and introspective observational contact with introspectibles need not be thought to do so either.<sup>78</sup>

In any event, to repeat, all this is supposed to apply to the **central**, not necessarily **universal**, manner in which first-person exp-intentional ascription proceeds. There may well be cases of such ascription that do not fit the observation-cum-endorsement model. But for our present purposes, it is sufficient that **some** exp-intentional ascriptions do—provided that **no** nexp-intentional ascriptions do. For this would suggest that only exp-intentional states can be observed, and therefore that only they can serve as anchoring instances of the concept of intentionality. To complete the case for this, then, we must show that first-person nexp-intentional ascription is based on data that are the vectors of two forces and thus requires the employment of a principle of charity.

First-person nexp-intentional ascription involves ascribing to oneself nexp-intentional states, paradigmatically unconscious beliefs and desires.<sup>79</sup> More generally, however, our concern here is with the self-ascription of non-experiential-intentional states. Such states seem to fall into three categories: (i) dispositional or tacit beliefs and desires; (ii) Freudian suppressed or repressed beliefs, desires, and emotions; (iii) occurrent sub-personal states typically posited in the context of cognitive-scientific explanations of behavior. Let us consider the central models for ascription of states in each category.

Consider first the first-person ascription of a dispositional belief, say the belief that Michael Jordan is not a three-headed kangaroo. The central model for how we ascribe such states to ourselves seems to me to be the one presented by Evans. Evans (1982: 225–6) points out that when someone asks you whether you believe that there will be a third world war, you do not start monitoring your internal states in search of an item that fits the description “belief that there will be a third world war”; rather, you start considering whether there will in fact be a third world war. You attend not to psychological facts, but to geopolitical ones. That is, you seek evidence not for your believing the proposition, but for the proposition itself.<sup>80</sup> If you find such evidence, you ascribe to yourself the belief that there will be a third world war, but do not do so on the basis of observing the belief.<sup>81</sup> Thus you charitably ascribe to yourself (non-conscious) beliefs in what you (consciously) take to be true.<sup>82</sup> In a way, this is the most direct application of charity possible: we consciously launch an investigation into the question of a third world war, and when the investigation is completed, that which we consider to be the correct answer is ascribed to ourselves as a content of a belief—even though no belief item was ever sought in the course of the investigation.

In addition, it is not entirely clear that there really **are** dispositional nexp-intentional states. Plausibly, it is not the case that a person **dispositionally believes** that *p*, but rather that she is **disposed to believe** that *p* (and likewise for desire). On this view, there are no dispositional intentional states, only dispositions to enter (occurrent) intentional states. The motivation for this view is that any explanatory burden that dispositional intentional states are called forth to meet can be equally met, and more parsimoniously, by dispositions to enter intentional states (Manfredi 1993, Audi 1994).<sup>83</sup> There may be an **intuitive** cost in denying the existence of dispositional states, but there is no **explanatory** cost.<sup>84</sup> This is important, because when we ascribe non-experiential intentional states, the only evidence we can have for their existence is based on the behavioral effects their existence would explain. So lack of explanatory power vis-à-vis those behavioral effects would remove any rationale for positing them in the first place. Thus because no explanatory gain can be found in positing dispositional states generally, there is no reason to suppose that some nexp-intentional states are dispositional.<sup>85</sup> Needless to say, if there are no dispositional intentional states, then **a fortiori** we do not have observational contact with them, and

## 36 THE SOURCES OF INTENTIONALITY

so they cannot be cited to undermine the claim that only exp-intentional states are observable.

Consider now the case of Freudian states. One way we might ascribe to ourselves such states is exactly as we would ascribe them to another. Suppose one day in therapy session your angry colleague reaches an important insight about her relationship to anger. In consequence, immediately after her departmental meeting outburst, she ascribes to herself the same pair of unconscious belief and desire that you ascribe to her: the desire to evince anger energies while preserving her self-conception as a good person along with the belief that lashing out at a meaningless college procedure will serve that purpose. Both belief and desire are unconscious here, but not quite dispositional or “standing”; certainly the belief is occurrent. It is first-person ascription of Freudian suppressed belief and desire. This kind of first-person nexp-intentional ascription obviously works precisely as the corresponding third-person nexp-intentional ascription you were engaged in. Accordingly, the datum for ascription at your colleague’s disposal—her outburst behavior—is the vector of two forces: she cannot ascribe to herself the aforementioned belief in isolation from the aforementioned desire, nor the desire in isolation from the belief.

There are other cases of first-person Freudian nexp-intentional ascription that rely on a wider array of indirect behavioral and introspective data. For example, you might infer that you must be sub-consciously expecting bad news in a looming phone conversation from the fact that you have been postponing picking up the phone for a few hours already (a behavioral datum) and have been finding it difficult to concentrate during that time (an introspective datum). Inasmuch as you rely on the behavioral datum (not picking up the phone), the mechanics of ascription are the same as in **third**-person nexp-intentional ascription, so the datum is a vector of two forces. Thus, expecting bad news would explain not picking up the phone only in conspiracy with a desire for **good** news (or at least for there not to be bad news). More interestingly, the same applies to the introspective datum: expecting bad news would explain the difficulty in concentrating only in conspiracy with a desire for good news. If you desired bad news, the expectation of bad news would not explain your concentration problems. Thus even introspective data for first-person Freudian nexp-intentional ascription are vectors of two forces.

It is instructive to consider why it is that the introspective data behave so differently in exp-intentional versus nexp-intentional ascription (i.e., are

the vectors of two forces only for the latter). The reason is quite clearly that in the former case, but not in the latter, the data pertain directly to that which is ascribed, in the sense that the relevant deliverances of introspection are that one seems to be in the very intentional state one eventually ascribes to oneself. Because of this, first-person exp-intentional ascription involves nothing but endorsement of seemings, whereas first-person nexp-intentional ascription involves a more substantial inference—essentially, a causal inference from broadly experiential effects to behind-the-scene nexp-intentional causes.<sup>86</sup>

Finally, let us consider first-person ascription of non-experiential subpersonal intentional states. There are many examples of such states posited in cognitive science: Marr's (1982) 2.5D sketches, Milner and Goodale's (1995) dorsal-stream visual representations, blindsight, and subliminal perception would be examples drawn just from vision science. It is worth noting that ascription of such states does not normally take place in the folk, since the folk are typically unaware that such states exist (in contradistinction with Freudian states, which by now are much more folkloresque). In any case, since subjects have no introspective access to such states, they can ascribe such states to themselves only on the basis of behavioral data.<sup>87</sup> If so, the model for first-person ascription of such states would be the same as the model for their third-person ascription: it would be based on inference from observed behavioral effects to unobserved intentional causes. In both cases, the ascription is based on data that are the vectors of two forces.<sup>88</sup>

I conclude that first-person ascription of all three types of nexp-intentional state is based on data that are the vectors of two forces, or else involves a principle of charity even more directly (as in Evans' model for first-person ascription of standing beliefs).<sup>89</sup> Given that the same holds of third-person nexp-intentional ascription, I further conclude that **all** nexp-intentional ascription is based on such data, and that therefore there is an important epistemological asymmetry between non-experiential intentionality and experiential intentionality. The asymmetry is that while **all** nexp-intentional ascription is governed by charity, **some** exp-intentional ascription is not.<sup>90</sup>

I have endeavored to show that the best—certainly most natural—explanation for this asymmetry is that we have observational contact with at least some of our exp-intentional states, and none with any nexp-intentional states. Against the background of the anchoring-instance model—in particular, the notion that all anchoring instances of the concept of intentionality

## 38 THE SOURCES OF INTENTIONALITY

are instances with which we have observational contact—it follows that only a subject's own exp-intentional states can serve as anchoring instances of her concept of intentionality. Since this is the key Premise 4 in the argument (presented in §I.1.2) for the experiential origin thesis, we now have a sustained argument for that thesis—the thesis that an item qualifies as intentional just in case it has the same underlying nature as one's (observed, encountered) exp-intentional states. I close this section with consideration of some objections.

## . . . Objections and Replies

The argument of the section, recall, is this:

1. All the anchoring instances of intentionality are such that we have observational contact with them;
2. The only instances of intentionality with which we have observational contact are experiential-intentional states; therefore,
3. All the anchoring instances of intentionality are experiential-intentional states.

Objections to this argument can be divided into three groups: objections to the first premise, objections to the second premise, and objections directly targeting the conclusion. Below, I consider examples of each.

Objections to the first premise cannot quite suggest that some anchoring instances of intentionality require no observational contact, since that requirement falls out of the anchoring-instance model of the formation of observational natural kind concepts. However, they **can** simply reject the anchoring-instance model, or its application to the concept of intentionality. I have already defended the model's application to the concept (at the end of §I.1.2), but an objector might insist that the model itself is misguided. For one thing, concept formation is a psychological process, and any account of it is beholden to the vicissitudes of empirical research. Yet my presentation of the model was suspiciously free of empirical encumbrances.

In response, I would point out that although the matter of concept formation is certainly empirical, it may be that while the empirical research needed to establish one model over another at a high level of specificity is quite involved, at a low level of specificity certain relatively mundane observations might suffice. As I noted at the beginning of §I.1.1, my concern is better thought of as with **families** of (underdeveloped) models rather than

with specific (well-developed) models. That is, the level of specificity at which the anchoring-instance model is formulated does not call for particularly involved empirical research. Perhaps in putting it forward a philosopher would be wearing her scientist cap, just as scientists oftentimes advance important hypotheses wearing their philosopher caps, but if I am right, the considerations supporting the anchoring-instance model are available to the **unprofessional** scientist.<sup>91</sup>

Most objections to the argument of this section would presumably target its second premise. They would deny either (a) that we have observational contact with exp-intentional states or (b) that we do **not** have observational contact with nexp-intentional items.

The first option is to deny that we have observational contact with exp-intentional states, perhaps on the grounds that introspection is not observational or quasi-perceptual in the way suggested. What is sometimes known as the “perceptual model” of introspection, which construes introspection as crucially analogous to perception, may simply be misguided. Typically, the crucial analogy is understood to be that in both perception and introspection, there is a kind of direct “epistemic contact” with the perceived or introspected. Some philosophers have objected to such a perceptual model from the left, arguing that the relationship between introspection and the introspected is less intimate than between perception and the perceived. Others have objected to it from the right, arguing that the relationship is **more** intimate. My response is different in each case.

One inspiration for the left-wing objection might be the view that introspection is fundamentally directed at the world, rather than at experience, due to the so-called transparency of experience (Harman 1990).<sup>92</sup> Thus, according to Dretske’s (1995) “displaced perception” model of introspection, introspective judgments are not generated by endorsing introspective seemings, but rather by **inference** from perceptual experiences. Just as one hears that the postman has arrived by hearing the dog bark, so one introspects that one has a table experience by seeing the table. There is no direct epistemic contact with the postman, and none with the table experience.

In response, it is important to note that regardless of whether ascription is mediated by displaced perception, it is still a fact that one can ascribe the table experience to oneself without at the same time ascribing any other mental state, something one cannot do with one’s nexp-intentional states. That is, the asymmetry of ascription that underlies our argument from

## 40 THE SOURCES OF INTENTIONALITY

inference to the best explanation still holds. The proponent of displaced perception is of course allowed to reject our explanation of the asymmetry, but she owes us her own superior explanation. Yet it is not at all clear how the displaced perception model is supposed to explain the asymmetry.<sup>93</sup> Indeed, it is not clear what **resources** the model has for explaining the asymmetry: the asymmetry would be a complete mystery if introspection were nothing but displaced perception.<sup>94</sup>

A different left-wing objector is the philosopher who holds that there is nothing at all like perceptual immediacy in introspection, because introspection is always doxastic or intellectual: introspecting my table experience is simply **thinking** that I am having that experience (see, e.g., Rosenthal 1993).<sup>95</sup> Again, however, this does nothing to either undermine the ascriptive asymmetry pointed out above or offer a better explanation of it. More generally, the left-wing objector cannot simply reject our view of introspection without addressing the argument **for** it.<sup>96</sup>

Consider now the right-wing objector. This is the philosopher who holds that, unlike in perception, the relata of introspection are not “independent existences,” i.e., not entities which may persist in the other’s absence (see, e.g., Shoemaker 1994).<sup>97</sup> The term “acquaintance” is sometimes used to denote such an epistemic relation. The objection before us is that introspection involves acquaintance not observation. This objection strikes me as an **ignoratio elenchi**: what matters for the ability of exp-intentional states to serve as anchoring instances is that normal subjects have some kind of direct epistemic contact with them. Whether that epistemic contact is “observation,” because the contacted is numerically distinct from one’s introspective judgment, or “acquaintance,” because it is not, is beside the point.<sup>98</sup>

So much for the first option for rejecting the second premise of the argument—arguing that we do not have observational contact with exp-intentional states. The second option is to concede that we do but insist that we also have observational contact with some **non-exp**-intentional items. This again may be divided into two kinds of objection: one suggesting that we have observational contact with non-experiential mental intentional states, and one suggesting that we have such contact with altogether non-mental intentional items.

The first kind of objection may be mounted by appeal to cases in which we seem to directly perceive the intentional states of others. Thus, it has sometimes been claimed that just by attending to facial expressions, one can

**see** that a friend is nervous or distressed, say about an upcoming meeting with a romantic prospect. It is natural to interpret such ascription as involving nothing but endorsement of a **perceptual** seeming: it seems to one perceptually that one's friend is distressed or nervous about **p**, and one endorses that seeming. This in turn might render the relevant data of ascription "single-forced."

My response is that although this kind of intentional ascription is immediate in a way that may tempt us to assimilate it to first-person exp-intentional ascription, the temptation trades on an equivocation in the notion of immediacy. When we see that a friend is in distress, our ascription of distress to her is immediate in the sense of not being mediated by conscious inference, instead being rather "automatic." However, it is clearly not immediate in the sense that we observe **the distress itself**. For we do not observe the distress itself, only its facial manifestation. Accordingly, the facial manifestation is the vector of two forces: it indicates distress only against the background of attribution to the friend of a strong desire that the meeting with the romantic prospect be successful. If one knew for a fact that the friend was utterly uninterested in the person she is about to meet, one would not be in a position to (competently) ascribe distress or nervousness to the friend on the basis of the relevant facial expression.<sup>99</sup>

The second kind of objection under consideration is that we have some perceptual observational contact with non-mental intentional items, such as paintings and road signs. Assuming that "observational contact" can come in both introspective and perceptual forms, one might suggest that we form our concept of intentionality through both introspective encounter with exp-intentional states and perceptual encounter with paintings, road signs, and so on. According to Siegel (2006), for example, we not only **understand**, but can also **see**, that some words on a page mean that the basketball game has been canceled. This would involve perceptual contact with a nexp-intentional item.

Note, however, that the intentionality of language and road signs, certainly, and arguably also that of paintings and photographs, is partly conventional and depends on interpretation. This means that a subject needs to understand the role and nature of the interpretive conventions before seeing such items **as intentional**. In effect, it means that the subject needs to already have a conception of intentionality before seeing such items as intentional. Thus inked or mouthed words do not present themselves as contentful to

## 42 THE SOURCES OF INTENTIONALITY

a subject who has no conception of intentionality whatsoever.<sup>100</sup> Given this, such items cannot serve as anchoring instances of our conception of intentionality. By contrast, exp-intentional states do present themselves to introspection as directed, as possessed of aboutness or ofness, before we have any conception of intentionality.<sup>101</sup>

The last kind of objection to the argument of this section simply rejects its conclusion, leaving it unspecified what exactly might have gone wrong in the argument. One reason for which one might reject the argument's conclusion is the suspicion that the resulting account of intentionality will fail to capture all intentionality. The objector suggests that we have a pre-theoretic conception of which things are intentional and which are not, and any account of intentionality must capture this pre-theoretic extension of intentionality more or less correctly. However, if we accept the conclusion of this section's argument, she claims, we will fail to do so, because much of what seems pre-theoretically intentional does not share an underlying nature with exp-intentional states. Indeed, the cases of language and pictures are in point: it is unlikely that linguistic occurrences of "table" have the same underlying nature as table experiences.

In response, one could insist that linguistic expressions may turn out to share a deep feature with exp-intentional states in virtue of which both are intentional. More plausibly, however, I am tempted to suggest that the intentionality of linguistic expressions, even if it fails to share an underlying nature with exp-intentional states, may nonetheless be somehow **accountable for in terms of** the intentionality of items that do share an underlying nature with exp-intentional states. If so, there is no need to posit further intentionality to account for the domain of phenomena in need of accounting for. All that is required is the identification of the general relation *R* that obtains between intentional items that do not exhibit the same underlying nature as encountered exp-intentional states and intentional items that do when, and only when, we account for the intentionality of the former in terms of that of the latter. Once this is done, we can offer the following augmented formulation of the experiential origin thesis:

- (EO★) There is a relation *R* and a nature *N*, such that (i) most encountered exp-intentional states have *N* and (ii) for any item *X*, *X* qualifies as intentional iff either (a) *X* has *N*, or (b) there is an item *Y*, such that (b-i) *Y* has *N* and (b-ii) *X* bears *R* to *Y*.

The strategy is to argue that (EO<sup>★</sup>) recovers the instances of intentionality that (EO) may fail to. All this depends, of course, on the feasibility of somehow accounting for all intentionality in terms of intentional items that have the same underlying nature as encountered exp-intentional states. I undertake to outline a way of doing this in Chapter 4.

There are probably other objections that could be raised to the argument of this section, and which have not occurred to me. But it strikes me that the argument ultimately rests on a relatively straightforward observation, namely, that there is an asymmetry between first-person exp-intentional ascription and other kinds of intentional ascription, inasmuch as the data for intentional ascription are crucially different, and differently culled. Fundamentally, the difference is that whereas ascription of all nexp-intentional states is based on observation of those states' **causal traces**, ascription of at least some exp-intentional states is based on observation of those states **themselves**, not their traces. If we accept this basic claim about intentional ascription, it becomes plausible that only exp-intentional states can serve as anchoring instances of intentionality. This in turn makes it plausible that any item qualifies as intentional only if it shares an underlying nature with exp-intentional states.

### 1.3. "Experiential Intentionality"

I have discussed the experiential origin thesis, and the central role it gives to experiential intentionality in our conception of intentionality, without clarifying exactly what is meant by the notion of experiential intentionality. The purpose of this final section is to fix ideas with respect to this notion of experiential intentionality. I take up the following three questions (§1.3.1–1.3.3, respectively): What would it take for there to be experiential intentionality? **Is there**, in fact, any experiential intentionality? And if so, **how much** experiential intentionality is there? I do not discuss these matters with the goal of settling them once and for all. Rather, the goal is to give the notion of experiential intentionality sufficient texture to bring through the meaning and significance of the experiential origin thesis.

#### . . . Definition

By "experiential intentionality," I mean the intentionality that conscious experiences have in virtue of being conscious experiences.<sup>102,103</sup> Conscious experiences have many properties, only some of which they have in virtue

## 44 THE SOURCES OF INTENTIONALITY

of being conscious experiences. Thus, my visual experience of the blue sky has the property of occurring on a Wednesday, but it does not have this property in virtue of being a conscious experience. It would have this property even if it were an unconscious, subliminal perception (or a blindsighted one).<sup>104</sup> Its being conscious is irrelevant to its occurring on a Wednesday. But my experience does have its bluish qualitative character in virtue of being a conscious experience. It would not have this qualitative character if it were not a conscious experience. Thus its being conscious/experiential is crucial to its having a qualitative character.

Let us say that every conscious experience has an **experiential character**, and that a conscious experience is the conscious experience it is, and is a conscious experience at all, because it has the experiential character it does, and has one at all. Then to say that some conscious experience X has an intentional property P in virtue of being a conscious experience is the same as to say that X has P in virtue of X's experiential character. We may thus define being an exp-intentional property through the following biconditional:

A priori, for any intentional property P and state X, P is an exp-intentional property of X iff there is an experiential character E, such that (i) X has E, (ii) X is P, and (iii) X is P in virtue of X having E.<sup>105</sup>

With this definition of an exp-intentional property in hand, we can define **experiential intentionality** as the determinable, or genus, of all (and only) exp-intentional properties.<sup>106</sup> And we can define an **exp-intentional state** as either (a) the exemplification of an exp-intentional property at a time or (b) a bare particular of the right kind that instantiates an exp-intentional property.<sup>107</sup>

The above definition of an exp-intentional property can be read in two ways, however, depending on how we understand "in virtue of." On the first reading, X's being P is somehow **grounded** in its having E, in a way that gives E metaphysical primacy over P.<sup>108</sup> The second is neutral on matters of grounding and primacy, and implies merely that X has its intentional property **in its capacity as**, or **qua**, a conscious experience, in the sense that a certain counterfactual is true of it: if it did not have E, it would not be P. Call the former reading **doctrinal** and the latter **neutral**.<sup>109</sup> The doctrinal definition of experiential intentionality is this:

A priori, for any intentional property P and state X, P is an exp-intentional property of X iff there is an experiential character E, such that (i) X has E, (ii) X is P, and (iii) X's being P is grounded in X's having E.

The neutral definition is this:

A priori, for any intentional property P and state X, P is an exp-intentional property of X iff there is an experiential character E, such that (i) X has E, (ii) X is P, and (iii) if X did not have E, X would not be P.

In order to avoid prejudging certain issues, I will work here with the neutral definition of “exp-intentional,” and hence of “experiential intentionality.”<sup>110</sup>

This is also the main reason I have chosen to conduct the discussion in this book using the term “experiential intentionality” rather than the more entrenched “phenomenal intentionality.” The latter seems to me to have implicitly acquired the doctrinal definition, perhaps due to the substantive commitments of the first philosophers to use it (Loar 2003, Horgan and Tienson 2002).<sup>111</sup> Thus, many today already hear the term “phenomenal intentionality” as implying the primacy of the phenomenal to the intentional.<sup>112</sup> However, my concern in this book is not to argue for the primacy of the experiential over the intentional, merely for the primacy of the experiential-intentional over the non-experiential-intentional. Whether the experiential is in some sense prior to the intentional within the realm of the experiential-intentional, or conversely the intentional prior to the experiential, or neither—say, because they are one and the same—is something I wish to remain silent on.<sup>113</sup>

Another reason to prefer “experiential intentionality” is that to many philosophers the term “phenomenal” connotes the presence of sensuous qualities: a phenomenal state, to them, must be a sensory (or somatic) state.<sup>114</sup> Since I am keen to insist that there are non-sensuous, purely intellectual (cognitive) experiences, such as the experience of suddenly grasping a mathematical proof, or of realizing the solution to a philosophical problem, but am not at all keen to quarrel about whether the term “phenomenal” should apply to such experiences, I choose to work with the term “experiential.”<sup>115,116</sup>

One last issue left open is how to define “experiential character.” One approach to this question is to understand “experience” intuitively, as whatever it means in its everyday sense.<sup>117</sup> Another is to pin our hopes on the phrase “what it is like for the subject” capturing uniquely experiential character.<sup>118</sup> Here I am going to take a more involved approach, which is to understand experiential character as that for which there is, or appears to be, an **explanatory gap** (Levine 1983).<sup>119</sup> More precisely, to a first approximation we may use

## 46 THE SOURCES OF INTENTIONALITY

the following rigidified definite description to fix the reference of “experiential”: a property *P* is experiential iff in the actual world, there is (or appears to be) an explanatory gap between *P* and physical properties.<sup>120</sup> We can then define the experiential character of a mental state *X* at a time *t* as the sum (or conjunction) of all experiential properties instantiated by *X* at *t*.<sup>121</sup>

## . . . Existence

With the definition of experiential intentionality in place, two substantive questions present themselves: is there in fact any experiential intentionality, and if so how much of it is there? That is, are there any mental states endowed with experiential intentionality, and if so how many kinds of mental state are so endowed? This subsection and the next offer brief discussions of these matters, mostly by way of reportage on others’ work.

Against the background of our definition of experiential intentionality, the claim that there **exists** such a thing amounts to the thesis that there is at least one instance of an exp-intentional property. Call this the **experiential intentionality thesis**: For some mental state *X* and intentional property *P*, *P* is an exp-intentional property of *X*.<sup>122</sup> Note that the thesis does not require **every** intentional property to be experiential, nor every experiential property to be intentional. It is merely an existential thesis.

There are several sources, and grades, of potential resistance to the experiential intentionality thesis. One relatively radical view is that experiences are never intentional (Travis 2004). A more common view is that experiences are often intentional, but never intentional **qua** experiences: they have intentional properties, and they have experiential properties, but they could have the very same intentional properties even if they had very different experiential properties (see, e.g., Block 1990, 1996).<sup>123</sup> In the literature, there appear to be two main reasons for accepting the experiential intentionality thesis in the face of such resistance.

One reason—unlikely to win converts but central in motivating sympathizers—is the claim that experiential intentionality is simply **introspectively manifest**: attending to one’s stream of consciousness in the right way brings out that some conscious episodes are intentional, and intentional because experiential.<sup>124</sup> Horgan and Tienson’s (2002) descriptions of certain conscious episodes are naturally understood as attempts to bring out something like that.

An argument with a more neutral starting point is due to Siewert (1998 Ch.6). He observes that some conscious experiences are assessable for

accuracy purely in virtue of what it is like for the subject to be in them, and without need of interpretation. They thus have accuracy conditions—an intentional property—in virtue of their experiential character. Suppose you undergo an experience with a reddish experiential character. If nothing around you is red, then your experience is assessable as inaccurate. If something is red, then your experience may be assessable as accurate (under the right conditions). Thus experiential character brings in its train accuracy conditions. Since having accuracy conditions is an intentional property, and experiential character is (trivially) an experiential property, these accuracy conditions constitute an exp-intentional property.

In this book, I am going to take for granted the existence of experiential intentionality. I do so not because—at least not primarily because—I think this is self-evident or otherwise not in need of justification; rather because I have nothing substantially new to say in its defense, and I want to focus on other issues.<sup>125</sup> The next question I want to consider concerns the **scope** of experiential intentionality.

### . . . Scope

Work on the scope of experiential intentionality has tended to fall into two categories, concerned with experiential intentionality **within** and **without** the sphere of perceptual experience.

Within that sphere, there are elements in perceptual experience that outrun sensory phenomenology narrowly construed but are claimed to nonetheless involve experiential intentionality. Perhaps the most systematic contribution to this kind of research is probably due to Siegel (2005, 2006), who argues that high-level properties are represented in perception. These include causation, meaning, and kind properties, among others.<sup>126</sup>

Outside the sphere of perceptual experience, three potential types of experiential intentionality are debated, instantiated by somatic, emotional, conative, and cognitive states. Somatic states, such as pains and pleasures, have historically often been treated as experiential but not intentional. More recently, it has been argued that such somatic states do have directedness: a toothache is an experience intentionally directed at a tooth and an armpit tickle is an experience intentionally directed at an armpit (see Armstrong 1962, Tye 1997). Ditto for emotions and moods: a disappointment with the New England Patriots' season is intentionally directed at their season, while a general, "undirected" depression is in reality a quiet, prolonged experience

## 48 THE SOURCES OF INTENTIONALITY

intentionally directed at the world as a whole and presenting it as dull and pointless (see Seager 1999 Ch.8).<sup>127</sup> With conative and cognitive states, historically it has been more common to conceive of them as intentional but not experiential, but again their experiential character has been stressed in more recent work. It is, in truth, hard to understand how desires, intentions, and other conative states could have been denied an experiential dimension, given that there is clearly a varying degree of experiential intensity or vivacity involved in them: one can want an ice cream very much or only mildly, where the difference is not only functional but also experiential.<sup>128</sup>

The liveliest area of debate here is the question of whether there are purely intellectual or cognitive experiences. Among views that accept such experiences, a further distinction can be made between those that allow similar cognitive experiences to have different intentional contents and those that insist that cognitive-experiential character determines intentional content.<sup>129</sup> In the literature, three main argumentative strategies have been wielded by way of defending the existence of (content-determining) cognitive experience: arguments from experiential overwhelm, arguments from experiential contrast, and arguments from first-person knowability.<sup>130</sup>

Arguments from experiential overwhelm attempt to identify specific, rather unusual cognitive states whose experiential character is supposed to be overwhelming in the way that approximates the experiential overwhelm involved in, say, visual and somatic experiences. Perhaps the best known instance of this argument is based on the sudden and acute onset of an experience of grasping something. Consider the following passage (Mangan 2001, Chudnoff forthcoming-a):<sup>131</sup>

A newspaper is better than a magazine. A seashore is a better place than the street. At first it is better to run than to walk. You may have to try several times. It takes some skill but it is easy to learn. Even young children can enjoy it. Once successful, complications are minimal. Birds seldom get too close. Rain, however, soaks in very fast. Too many people doing the same thing can also cause problems. One needs lots of room. If there are no complications it can be very peaceful. A rock will serve as an anchor. If things break loose from it, however, you will not get a second chance.

This passage elicits mostly puzzlement, until one is told that it is about **kites**. Once informed, one typically undergoes the acute onset of the grasp experience of which I speak.<sup>132</sup> Arguments of this sort are particularly effective

dialectically, but suffer from the limitation that they do not obviously establish the stronger kind of claim one might make about cognitive-experiential character, namely, that they determine intentional content. Yet it is this stronger claim that is of interest if we are interested in the existence of cognitive experiential intentionality.

In arguments from experiential contrast, two overall conscious episodes of a subject are presented, such that (i) it is intuitively clear that there is an experiential difference between the two, and (ii) the best explanation of this overall experiential difference is that one instantiates a cognitive experiential character but the other does not. Thus, Strawson (1994 Ch.1) argues for the existence of “understanding experience” by contrasting the overall experiences of a French speaker and a non-French-speaker listening to the news in French (see also Moore 1953). Strawson argues that there is an experiential difference between their overall conscious episodes, and that the difference is best accounted for in terms of an element of understanding-experience present only in the French speaker’s stream of consciousness.<sup>133</sup> As it stands, the argument does not clearly establish that different cognitive-experiential characters go along with different intentional contents, but it can be tweaked to do so. Consider that  $\hat{p} \hat{a} \hat{i} \hat{n}$  refers to an unpleasant sensation in English and to bread in French. Imagine a world where two languages are so similar graphically and phonetically that the very same passage can express a news report about a faraway war in one of them and a children’s bedtime story in the other. We can envisage two subjects listening to a reading of the passage and each understanding it in a different language. Here there would be an overall experiential difference that is best explained by supposing that one subject’s cognitive experience had one intentional content while the other’s had another intentional content.

In arguments from first-person knowability, it is pointed out that the knowledge we have of our occurrent thoughts is the kind of knowledge we have only of our conscious experiences, from which it is inferred that these thoughts must be conscious experiences (*qua* thoughts, that is, not only in virtue of their accompanying imagery).<sup>134</sup> Thus, Pitt (2004) argues that we have an immediate, non-inferential knowledge of some cognitive states, and that only conscious experiences are knowable in this way, which suggests that the cognitive states have a proprietary experiential character.<sup>135</sup> Pitt further argues that whenever cognitive states differ in intentional content, we can know the difference between them in this non-inferential

## 50 THE SOURCES OF INTENTIONALITY

way peculiar to experience, wherefore he concludes that cognitive experiential characters determine intentional content.

For my part, I am persuaded of the existence of cognitive experience by all three lines of reasoning, but perhaps most vividly by something like everyday experiential overwhelm: it simply seems that my inner life is much more interesting to me than it would be if my conscious experience consisted merely in perceptual experiences. Thus, when I contemplate the problem of determinism and free will, it is not the accompanying imagery that fascinates.<sup>136</sup> In addition, however, I develop a more theoretical argument for it in Kriegel forthcoming. Working with the conception of experiential properties as suitably explanatory-gap-able properties, I argue that we can conceive of a world in which subjects have no perceptual, somatic, or emotional experience, but where the philosophical anxiety captured by the explanatory gap nonetheless appropriately arises. This suggests that the suitably explanatory-gap-able property is instantiated in that world, even though perceptual, somatic, and emotional experiential characters are not—which in turn suggests that the explanatory-gap-able properties of the subjects there are cognitive properties.<sup>137</sup>

Much more can be (and has been) said about the scope of experiential intentionality. As noted, in this book I will assume rather dogmatically that there is such a thing as experiential intentionality. In the same vein, I will also assume that experiential intentionality extends beyond purely sensory experience to high-level perceptual experience and cognitive experience. Although I recognize that these assumptions are not beyond rational doubt, it so happens that I personally do not doubt them. My concern here, in any event, is to consider what sort of overall understanding of intentionality we may obtain if we assume the existence of experiential (including cognitive-experiential) intentionality and accept the experiential origin thesis argued for earlier in the chapter. The purpose of the present section has been merely to elucidate (and give texture to) the notion of experiential intentionality.

### Conclusion: The Experiential Origin Thesis and the Case for It

The purpose of the chapter has been to argue for the experiential origin thesis. In spirit, this is the thesis that our conception of intentionality is grounded in our grasp of experiential intentionality. I proposed a precisification

of the thesis that would provide its letter: an item qualifies as intentional by the lights of one's concept of intentionality if, and only if, it has the same underlying nature as exp-intentional state encountered during the formation of one's concept of intentionality, where a state is exp-intentional just in case it has an intentional content and an experiential character such that it would not have the content if it did not have the character.

The argument I offered proceeded essentially as follows. First, for certain kinds of concept, an item falls under the concept just in case it has the same underlying nature as most manifest instances encountered during the concept's formation. Secondly, the concept of intentionality is of the relevant kind. Thirdly, the manifest instances of intentionality we encounter during the formation of our concept of intentionality are exclusively exp-intentional states (because these are the only instances of intentionality we have observational contact with). It follows that an item falls under our concept of intentionality just in case it has the same underlying nature as most encountered exp-intentional states.

If we do accept the experiential origin thesis, the theory of intentionality assumes a certain structure, dividing into two main parts: an account of experiential intentionality and an account of non-experiential intentionality. Chapters 2 and 3 are devoted to experiential intentionality, each proposing one account that I find reasonably plausible (and eventually comparing those two accounts). Chapter 4 is devoted to non-experiential intentionality, considering several possible accounts thereof and arguing for the superiority of one among them. Chapter 5 revisits the general theory of intentionality in light of the proposals made in the chapters leading up to it.<sup>138</sup>

## Notes

1. There are familiar semantic, epistemological, and metaphysical questions surrounding natural kinds. These concern the determination of reference and extension of natural kind terms and concepts, the formation and acquisition of natural kind concepts, and the identity, existence, and membership conditions of the kinds themselves. Typically, accounts of natural kinds and natural kind terms/concepts come in comprehensive package deals driven by insights either from the semantic or from the metaphysical angle. Here I will focus, atypically, on the epistemological angle. The question I want to concentrate on concerns how we form natural kind concepts.

2. There are three contrast classes here. First, there are concepts of kinds whose similarity is due to our activities. Thus, artifactual kind concepts are not natural kind concepts, as the similarity among

## 52 THE SOURCES OF INTENTIONALITY

their instances is due to human activity: tables are similar because they were built by us with a purpose in mind. Likewise, social kinds are due to our legal or social institutions: the similarity among bachelors is only relative to the institution of marriage. Secondly, there might be concepts whose instances resemble only mind-dependently. This would be the case, according to subjectivists about value, or color, with our value or color concepts. Thirdly, there could also be concepts whose instances are such that there is **no** similarity of any kind among their instances (this may be a far-fetched possibility if we are liberal about what counts as similarity, but need not be ruled out a priori).

3. It is not immediately obvious how to draw the observation/not observation distinction. One option is to pattern the *prima facie* distinction between observational and non-observational (natural kind) concepts on the traditional distinction between observational and theoretical terms (e.g., an observational concept would be thought of as a concept expressed by an observational term, and a theoretical concept as a concept expressed by a theoretical term), but I am not keen to do so, given the doctrinal baggage that has tended to come with the distinction in the hands of logical positivists.

4. Arguably, mammals form a natural kind, and our conception thereof is as of a natural kind. That is, our concept of a mammal is a natural kind concept.

5. This is of course an oversimplification: those who wrote what we read probably read about mammals in other sources. But at the end of the chain is someone who formed a concept of a mammal not by reading about mammals but by interacting with them.

6. Certain non-observational natural kind concepts are such that **nobody** forms them on the basis of encounter with instances. Thus, nobody has ever encountered a lepton, yet our concept of a lepton is a natural kind concept. The models discussed in the text of this subsection therefore do not apply to such concepts.

7. The criterial model is very natural for certain kind concepts, such as the concept of a bachelor. To qualify as a bachelor is clearly to satisfy the conditions of being a man and being unmarried. In a way, the criterial model takes the formation of the concept of a bachelor to be paradigmatic and suggests that other concept formations should be understood along the same lines.

8. This is supported by empirical work in developmental psychology that demonstrates young children's appeal to hidden natures in classifying objects (Gelman and Markman 1986).

9. Linnaeus classified whales as mammals, in the mid-eighteenth century, on the basis of such in-principle-manifest, observable-but-typically-unobserved features as having hollow ears and a warm heart. I trust it is clear that although such features are in principle observable, they are not observed during the (non-parasitic) formation of normal persons' conception of a whale.

10. Note that in this formulation the underlying nature is said to be possessed by particulars. If this sounds clunky to our ears, and we think that underlying natures are of properties not particulars, then the formulation would have to be changed to refer to underlying common features rather than underlying natures, or perhaps better yet, to the underlying nature of commonalities. Here I will continue to refer to natures as had by particulars.

11. One could mean something else by the "underlying-nature model of natural kind concept formation," and on this alternative understanding these consequences may not follow. But here I understand by the "underlying-nature model of natural kind concept formation" exactly what (UN) states—no more and no less.

12. Notice that I am using the term "instance" to denote any item that falls under the concept. The term "instance" is sometimes used to denote only items that instantiate the corresponding property, but this is not how I use the term.

13. This will be refined and modified when the model assumes more shape below.

14. This could be made more explicit by introducing the anchoring instances with an operator and plugging in the explication of "anchoring instance" as a relevantly encountered manifest instance.

The result would be this: For any natural kind  $K$  and particular  $X$ , there are particulars  $k_1, \dots, k_n$  and a nature  $N$ , such that  $X$  qualifies as a  $K$  iff (i) all of  $k_1, \dots, k_n$  are encountered manifest instances of  $K$ , (ii) most of  $k_1, \dots, k_n$  have  $N$ , and (iii)  $X$  has  $N$ .

15. There is the empirically quite unlikely possibility that every single anchoring instance turns out to have a different underlying nature. In this case, the concept is either (a) an empty natural kind concept or (b) a non-natural kind concept (this I find more plausible). If (a), then it will be discoverable that there is no such thing as mammal-ness, which is not quite the same discovery as that there are no mammals. If (b), then in the scenario envisaged all the anchoring instances would qualify as mammals.

16. With the exact formulation of the anchoring-instance model in place, we can formulate a precise explication of the notion of an anchoring instance, based on its theoretical role in the model: A priori, for any particular  $X$  and natural kind  $K$ , there are items  $k_1, \dots, k_n$  and a nature  $N$ , such that  $X$  is an anchoring instance of  $K$  iff (i) for any  $Y$ ,  $Y$  qualifies as a  $K$  iff (i-a)  $Y$  has  $N$  and (i-b) most of  $k_1, \dots, k_n$  have  $N$ , and (ii)  $X \in k_1, \dots, k_n$ .

17. The six notions are: “qualify as,” “concept,” “formation,” “encounter,” “most,” and “particulars.”

18. Ditto for the normality of conditions: there are statistically normal conditions and teleologically normal conditions, the latter being the conditions that allow the subject’s concept-forming mechanisms to perform their function(s) as they are supposed to.

19. Note that it is still required here that the unperceived individuals be manifest, or observably, similar to the perceived one. On this view, if there are very unusual cats or horses, which do not look sufficiently like the individual cats and horses the subject token-encountered during the formation of her concept of a mammal, these unusual cats and horses—non-manifest cats and horses—do not qualify as anchoring instances of her concept of a mammal. But other cats and horses do.

20. I am enamored of the thought that in such circumstances the concept defaults into manifest kind concept status. But clearly this is not the only option for the model. Other options are to construe the concept as an empty natural kind concept, or perhaps as a concept that refers to a disjunctive kind or a concept that has divided reference.

21. Two further comments are in order. First, the conjunctive line would make the non-anchoring instances that have only one of the two natures not instances at all (since they would not share the conjunctive underlying nature exhibited by the anchoring instances). Secondly, this line requires there to be conjunctive natural kinds. Presumably, some of Armstrong’s (1978 Ch.15) arguments for the existence of conjunctive universals could be adapted to make the case that there are conjunctive natural kinds. In particular, the argument that universals are posited to explain objective similarity relations among particulars, and that conjunctive universals, such as the red square, are thus needed to explain the objective similarity among certain particulars, such as all red squares, would transfer straightforwardly to natural kinds.

22. As is common in philosophical practice, here and elsewhere I use hyphens to indicate that the parts of the expression are morphological parts but not syntactic parts (in the way “apple” is a morphological but not syntactic part of “pineapple”).

23. For a classic on prototype theory, see Rosch 1973.

24. More accurately, given the eighth clarification, we should say that the anchoring instances that are genuine instances are a proper subset of the genuine instances. Also, perhaps a qualification such as “typically” should be introduced for anchoring instances, since there is no reason to rule out a priori the possibility of a concept all of whose instances turn out to be anchoring instances.

25. There might in principle also be prototypical instances that are not anchoring instances, in case highly prototypical instances of some concept are rarely encountered and thus do not play an anchoring role in the formation of that concept. I cannot think of a concrete example of this, but see no reason to exclude the possibility.

## 54 THE SOURCES OF INTENTIONALITY

26. How to make sense of the notion of being “more of an instance” is of course not a trivial matter. Presumably different approaches to vagueness could be applied here in different ways, or fuzzy logic could be used to regiment talk of degrees of “instance-hood.”

27. It is not entirely clear to me what theories of the “structure” of concepts target—what notion of structure they have in mind and what phenomenon they hope the notion denotes. I do note, however, first, that typically prototype theorists present their theories as about the structure of concepts, and second, that in their mouths the notion seems to be metaphysical rather than epistemological.

28. It is less clear whether the model is compatible with a direct reference account of natural kind concepts. More likely, the model requires concepts to refer via a description such as “has the same underlying nature as most anchoring instances.” Correspondingly, the model is probably also incompatible with a pure causal theory of natural kind terms/concepts’ reference (Kripke 1972). This doctrine is much more controversial, however, and in my view faces systemic problems that have been pointed out more or less immediately after its inception. I will not discuss these issues here, as they would take us too far afield, but see Zemach 1976 and Kim 1977 for incisive criticism.

29. I will consider objections to this later on. These will take two forms: denying that the concept of intentionality is a natural kind concept, and accepting that it is a natural kind concept but denying that it is an observational one.

30. In what follows, I will often drop the parenthetical reference to concepts for smoothness of exposition. This is not intended to lure the reader into forgetting that my thesis concerns concepts rather than properties—I am quite explicit that the thesis is primarily epistemological.

31. A more precise formulation is this: 1) For any item  $X$  and natural kind  $K$ ,  $X$  qualifies as a  $K$  iff there is a nature  $N$ , such that (i) most of the anchoring instances of  $K$  have  $N$  and (ii)  $X$  has  $N$ ; 2) The concept of intentionality is a natural kind concept; therefore, 3) For any item  $X$ ,  $X$  qualifies as intentional iff there is a nature  $N$ , such that (i) most of the anchoring instances of intentionality have  $N$  and (ii)  $X$  has  $N$ ; 4) For any item  $X$ ,  $X$  is an anchoring instance of intentionality iff  $X$  is an encountered experiential-intentional state; therefore, 5) For any item  $X$ ,  $X$  qualifies as intentional iff there is a nature  $N$ , such that (i) most encountered experiential-intentional states have  $N$  and (ii)  $X$  has  $N$ .

32. Keep in mind, in any case, that to ground our conception of intentionality in our grasp of experiential intentionality, all that is required is that **only** experiential-intentional states be anchoring instances of our conception of intentionality; it is not required that **all** be.

33. It may be objected that some—many—normal subjects do not undergo non-parasitic formation of most of their concepts. This is true, but a **non-sequitur**. The claims I am making are only about concept formation in normal subjects when they form the concepts they do non-parasitically. I take it that parasitically formed concepts inherit their structure from non-parasitically formed ones.

34. As I will mention more explicitly soon, the kind of encounter at play here is more plausibly construed as (what I called in §1.1.1) type-encounter, that is, the encounter a subject has with a type of entity on the basis of directly perceiving (i.e., token-encountering) individual tokens.

35. The qualification “probably” is needed since it could still turn out that some encountered experiential-intentional states do not exhibit the relevant manifest features. As noted, this is not very probable, but is a coherent possibility.

36. This means that the set of encountered experiential-intentional states probably includes all the experiential-intentional states that have the same manifest properties as the token experiential-intentional states that have actually been “met.” So, for example, if a subject has “met” (had a “direct” awareness of, in some duly explained sense of “direct”) a visual experience as of a table, then all visual experiences as of tables (or at least all experientially type-identical visual experiences as of tables) count as “encountered.”

37. I am assuming here that **states** in general are particulars. This does not necessarily mean that they are **bare** particulars, as opposed to, say, property instantiations or tropes. For the latter are widely regarded to be abstract particulars (see Williams 1953), thus **a fortiori** particulars. Certainly, in any case, intentional states are not **quantities** or **samples** in the way units of water and gold are. Furthermore, propositions are likewise particulars: there is one of each (at most!). The fact that they are abstract entities does not affect their status as particulars.

38. It strikes me as quite plausible that the prototypes of intentionality are exclusively experiential-intentional states. However, this is an empirical thesis and the empirical confirmation for it would be difficult to generate (though I suspect an empirical disconfirmation would be even harder). Elsewhere, Terry Horgan and I have suggested a somewhat similar thesis, according to which all the prototypical **mental** states are phenomenally intentional states (Horgan and Kriegel 2008).

39. It is clear, nonetheless, that there is a temperamental difference—a difference in animating sensibilities—between the two approaches (see Montague 2010, Kriegel and Horgan forthcoming). More on this later.

40. Traffic signs are about something: a sign on Interstate 95 that says “New York, next exit” is **about** the sorts of actions that have to be taken in order to reach New York from where the reader is at.

41. This thin existential claim is highly plausible **prima facie**, but will also be argued for, in effect, in §1.2.

42. On some views, such as Dennett’s (1971, 1987) “intentional stance” theory, intentionality **is** observer-dependent in this way. We will, in fact, discuss some relevant ideas in §1.2 and §4.4. However, while the view has insightful things to say about the practice of intentional ascription, the inference from those things to the anti-realist, observer-dependent view of intentionality is somewhat mysterious, and to my mind has been soundly refuted by Fodor and LePore (1993), who comment (1993: 76): “It’s mysterious . . . either how you could make facts out of stances, or how stances could make facts disappear.”

43. I suspect that every concept we have we treat initially as potentially picking out a natural kind. We approach the kind open to the possibility that it has a deep underlying nature. If and when we realize that it does not, the manifest features of the anchoring instances become the kind’s “essential” feature, that is, the individuating features.

44. This is a rather informal sub-argument. Its conclusion is not exactly the same as Premise 4 above. The more precise formulation would be as follows: 1) For any item  $X$ ,  $X$  is an anchoring instance of intentionality iff  $X$  is an encountered intentional item; [anchoring-instance model] 2) For any item  $X$ ,  $X$  is an encountered intentional item iff the (normal) subject has observational contact with  $X$ ; [plausible assumption] 3) For any item  $X$ , the (normal) subject has observational contact with  $X$  only if  $X$  is an experiential-intentional state; [to be argued for] therefore, 4) For any item  $X$ ,  $X$  is an anchoring instance of intentionality only if  $X$  is an experiential-intentional state; [from 1,2,3] 5) For any item  $X$ ,  $X$  is an anchoring instance of intentionality if  $X$  is an encountered intentional item; [from 1] 6) For any item  $X$ ,  $X$  is an encountered intentional item if  $X$  is an encountered experiential-intentional state; [logical truism] therefore, 7) For any item  $X$ ,  $X$  is an anchoring instance of intentionality if  $X$  is an encountered experiential-intentional state; [from 5,6] therefore, 8) For any item  $X$ ,  $X$  is an anchoring instance of intentionality iff  $X$  is an encountered experiential-intentional state. [from 4,7]

45. The objector might point out that many natural kind concepts involve no encounter with instances; the concept of a lepton is a case in point, as are many other concepts of “theoretical entities.”

46. This provides a non-demonstrative argument for the second premise of the argument with which I opened this section, which in turn is a demonstrative argument for Premise 4 of the previous section’s argument for the experiential origin thesis.

47. Note, however, that there are many different senses in which intentionality is sometimes said to be normative. Sometimes it is said to be normative in aiming at truth, sometimes in connection with

## 56 THE SOURCES OF INTENTIONALITY

rule following, sometimes in other senses. The only sense that will concern us here, however, is the Davidsonian claim that principles of charity are needed for competent ascription.

48. The expression “non-experiential-intentional” is ambiguous, and I mean by it anything that has non-experiential intentionality, not anything that does not have experiential intentionality. (Obviously, experiential intentionality is not prior in any sense to everything else!)

49. I use the term “overt behavior” to refer to publicly observable, third-personally accessible behavior, i.e., bodily behavior. There is a sense of “behavior” in which many mental acts, e.g. calculating, count as behavior, namely, the sense in which calculating is something one does, not something that happens to one. It is these mental acts that are supposed to be excluded by the notion of **overt** behavior. We may think of them as **covert** behavior.

50. As Davidson (1970: 97) puts it, in constructing a theory of someone’s behavior, “we will try for a theory that finds him consistent, a believer of truths, and a lover of the good (all by our own lights, it goes without saying).”

51. By saying that the person takes the word “tiger” to express the concepts of a tiger (or pigeon), I mean that she takes the word to express the concept **we** take the words “tiger” (or “pigeon”) to express.

52. A quick comment on the goodness of the desire: the desire is good not in the moral sense, but in the **prudential** sense. In that sense, a desire is good when it is a desire for something that is genuinely good **for the desirer**. Consider the robber’s desire to get out of the bank as quickly as possible. This is a prudentially good but not morally good desire.

53. There are also descriptive principles of ascription at play here, which favor both these options over, say, ascribing to the person the belief that the moons of Jupiter are nice and the desire to fly backwards. Descriptive principles ensure that this assignment is inferior, on this occasion, to the options considered in the text. But the options in the text are discriminated by normative principles, not descriptive ones.

54. We might say, using terminology from Searle (1983 Ch.1) among others, that it involves one force that is an attitude with a mind-to-world direction of fit and one with a world-to-mind direction of fit.

55. Here too I follow Davidson (1963), now in his account of a reason for  $\Phi$ -ing as a combination of a desire that  $\mathfrak{p}$  obtain and a belief that  $\Phi$ -ing would make the obtaining of  $\mathfrak{p}$  more likely.

56. There may certainly be cases where there are several possible ascriptions that a competent interpreter might choose among while remaining competent. But my concern here is with other cases.

57. Davidson’s reasoning here may be represented as an argument by inference to the best explanation. What needs explaining is that competent interpreters manage, **qua** competent interpreters, to converge on a single interpretation (and can recognize that that is the correct interpretation) in most typical situations, even though the data at their disposal grossly underdetermine interpretation. The best explanation of this is that competent interpreters, **qua** competent interpreters, employ principles of charity that rule out possible interpretations otherwise consistent with the data. Alternative explanations would cite some other device(s) competent interpreters might employ, in their capacity as competent interpreters, to rule out those interpretations. But it is not clear what other device(s) there might be, so no plausible alternatives present themselves to the Davidsonian explanation that competent interpretation must employ, and is to that extent governed by, principles of charity. (Needless to say, all this applies only to **competent** interpretation. Faced with the pigeon-pointer who exclaims “This is a tiger,” we can, if we want, interpret her as believing that she is looking at a tiger. We can also interpret her, if we very much want, as believing that she is looking at a three-headed kangaroo. What we cannot do, however, is interpret her in this way **competently**.)

58. We obtain the four-way distinction by crossing two simpler distinctions, one between ascription of experiential-intentional states (“exp-intentional ascription”) and non-experiential-intentional states (“nexp-intentional ascription”) and the other between ascription of an intentional state to oneself (first-person ascription) and ascription of an intentional state to another (third-person ascription).

59. In this respect, my argument will use Davidson's own vector-of-two-forces observation to argue for a very non-Davidsonian approach to first-person ascription, and thus self-knowledge. For Davidson's own views on the matter, see Davidson 1987.

60. There certainly exist much more complex instances of intentional ascription. But my contention is that the complexity in those cases does not change the basic mechanics of ascription, though it may well serve to becloud those mechanics. This is a reason to focus on simple examples in which the mechanics come through clearly.

61. Recall that, as we saw in the previous subsection, these are the sorts of belief and desire that we need to ascribe in this case to subsume the case of verbal behavior under the more general case of behavior.

62. In other words, there are different desires that could be ascribed together with the same belief, and different beliefs that could be ascribed together with the same desire, but it is impossible to ascribe the belief in isolation from **any** desire or the desire in isolation from **any** belief.

63. A non-mental analogy may be useful. Suppose you are presented with a color patch projected on a window pane. You are told that the color is produced by projection from two different light sources, one from the left and one from the right. You can see the patch, but not the light sources. Thus, when you see a purple patch on the pane, your data are compatible with at least three hypotheses: that the left light is blue and the right one red; that the left light is red and the right one blue; that both are purple. Since the color of the patch is the vector of two forces, and you do not have direct access to those forces themselves, ascribing colors to the light sources must involve an element of **decision**. (In the present case, that the decision must be arbitrary; in the case of *nexp*-intentional ascription, it is based on charity.) Compare now a case in which you **can** see not only the patch, but also the light sources—at least **sometimes**. Here you do have direct access to the lights, and so ascribing colors to them is not the vector of two forces. This is what happens, I will argue below, in the central case of *exp*-intentional ascription.

64. There are places where Dennett seems to suggest that in many of our intentional ascriptions, we are simply ascribing to others beliefs in whatever we happen to think is true (see especially Dennett 1971). Thus, we may ascribe to someone a belief that **p** not on the basis of their behavior, but on the basis of it being the case that **p**. Such ascription would certainly not be based on data that are the vectors of two forces, but I hope it is clear that it would be imbued with charity in an even more straightforward way.

65. The notions of cognitive experience will be discussed more fully in §1.3.3. See also Kriegel forthcoming.

66. We may stipulate, since this is a thought experiment, that your colleague does not in addition have a conflicting desire to evince anger etc. There is no internal division inside her will. Her will is one, so to speak.

67. It should be stressed that, given that what we have been concerned to establish is that some *exp*-intentional ascription is based on data that are not the vectors of two forces, without committing on the question of whether **all** such ascription is, there is no need for us to worry overmuch about third-person *exp*-intentional ascription. For our existential claim can be established simply by examination of first-person *exp*-intentional ascription, and regardless of how third-person *exp*-intentional ascription turns out to work. Nonetheless, for the sake of completeness, let me consider briefly this kind of ascription.

68. The causal structure of the hidden mechanism is a little different when the ascribed state is not a belief, but the kind of perceptual state that causes belief. In the former case, the envisaged mechanism involves two states, the belief and the desire. In the latter, it involves three states, the belief, the desire, and the perceptual state that causes the belief. In both cases, however, it is impossible to ascribe a single, non-conspiratory state.

## 58 THE SOURCES OF INTENTIONALITY

69. One might object as follows. To ascribe to oneself an experience as of a table, one would have to possess the concept of a table, but possession of the concept of a table requires the possession of many other concepts, for example those that combine with one's concept of a table to constitute one's beliefs about tables. This seems to me a *non sequitur*. First, a concept is not a mental state, but a constituent of a mental state, so the symmetry between attributions holds even if first-person exp-intentional attribution requires the attribution of certain concepts. Secondly, and even more importantly, the fact—if it is a fact—that first-person exp-intentional attributions require the existence of certain concepts does not entail that it requires the attribution of those concepts. So again this would not undermine the asymmetry with other forms of intentional ascription (which do not require attribution of more than one mental state).

70. Thus although the experience, being perceptual, has a mind-to-world direction of fit, its ascription need not be accompanied by the ascription of any intentional state with a world-to-mind direction of fit.

71. This may invite the misguided objection that a perceptual experience is itself a combination of two states, say a sensation and a belief, and so first-person exp-intentional ascription also ascribes only two states at once. Clearly, however, even if we accept this view of perceptual experiences as combinations of two states, there are differences between the reason we can only ascribe two states at once in this case and the reason in other kinds of intentional ascription. The fact (if it were a fact) that we can only ascribe two states at once in first-person exp-intentional ascription would not entail that there are competing interpretations fully consistent with the data, as is the case with the other kinds of intentional ascription, and therefore it would not entail that principles of charity must be operative in choosing among these ascriptions.

72. I use the term “endorsement” to cover not only conscious acts, but also automatic sub-conscious processes that involve ratification of the content of one's experience. I am thus not wedded to a model according to which one consciously and deliberately goes through two distinct steps in performing a perceptual judgment. It may feel immediate to the subject when the subject judges that there is a table before her. Nonetheless, the cognitive processes involved in the production of such a judgment divide into different steps, one of which can be described as endorsement.

73. I will discuss first-person exp-intentional ascription momentarily.

74. It is not a straightforward matter to elucidate the sense in which the epistemic relation you bear to your experience can be described as “observation.” What I have in mind is the thought that there is an epistemic relation we sometimes bear to our own mental states that is analogous in some respects to perception of external entities. There is a difference between believing that it is raining on the basis of the weather report and believing this on the basis of *seeing* the rain. The latter case involves a kind of direct contact with the state of affairs believed to obtain that the former case does not. In an analogous way, there is a difference between the way you know that what you are visualizing right now is a smiling octopus and the way I know this. I know it on the basis of testimony, you know it on the basis of direct contact with the believed state of affairs. It is this kind of contact that I think would not be misleadingly described as “observation.” As I will argue in §1.2.3, however, this perceptual model of introspection is not necessary (but only useful) for the claims I am making using it.

75. I use the term “yields” in a way that is neutral between a causal and a constitutive reading, to accommodate different views of the nature of introspection. I will discuss some of these views more explicitly in §1.2.3.

76. I put “hypothesis” in quotation marks because, as I mentioned earlier, the view that we have such observational contact strikes me as phenomenologically obvious in a way that makes calling it a hypothesis sound funny. To me, it is almost a datum. As noted, however, I am offering an argument for those who see things differently, and in the context of the argument treat the claim as a hypothesis.

77. This statement of self-intimation is sloppy. What is claimed to be self-intimating is not introspection, actually, but that which is introspected. Also, it is not for **every p** that a doctrine of self-intimation counts as self-intimating; only **mental** propositions are relevant.

78. As it happens, I happen to believe that a strong kind of epistemic privilege does attach to the deliverances of introspection, and have argued so elsewhere (Horgan and Kriegel 2007), but that view is logically independent of the claims made here.

79. Importantly, conscious occurrent beliefs and desires do not qualify as **nexp-intentional** states, but as **exp-intentional** ones. As will be discussed in §1.3.3, I hold that beliefs can constitute conscious experiences—“cognitive experiences”—so I am assuming here that **exp-intentional** ascription is not restricted to states with a specifically **sensory** experiential character. It has sometimes been claimed that beliefs cannot be conscious and occurrent (Crane 2001). I am somewhat sympathetic to this claim, understood as a claim about how the English word “belief” is normally used, but it is clear that there is no substantive issue at stake here, inasmuch as a tacit belief could still transform into a conscious, experiential intentional state; it is just that that state would not qualify as a belief the way the word is normally used. A natural term to pick out the kind of state a tacit belief transforms into when it becomes conscious is “thought.”

80. To be sure, sometimes the evidence is so overwhelming that no process of “seeking” is needed. This seems to be the case with the belief that Michael Jordan is not a three-headed kangaroo. But even in this case the evidence for ascribing the belief is nothing but the evidence for the proposition believed.

81. It is true that just as when I ask you whether you are having a table experience, you reply immediately that you do, so when I ask you whether you believe that there are more than four countries in Africa, you reply immediately that you do; and the immediacy of reply might suggest direct observation in both cases. But given Evans’ model, the immediacy of reply is not to be explained by the fact that you observe your belief, but in some other way. My guess is that the right explanation of the immediacy is that when the proposition you are asked whether you believe is obviously true, in that the evidence you have for it is overwhelming and readily available, you will answer the question immediately because the question is so easy. The explanation for this immediacy is therefore not that you have observed the believing, but rather that the proposition believed is obvious (again, in the sense that the evidence for it is overwhelming and readily available). Thus there is no argument from immediacy-of-reply to direct observation.

82. There is a sense in which the word “charity” is a little strained for this kind of case, but this seems to do with the pragmatics of normative talk when self-applied. At bottom the cases are of the same general kind. Thus just as we often ascribe to others beliefs in whatever we realize is true, by way of charity, so we ascribe to ourselves dispositional beliefs in whatever we realize is true—also a form of charity.

83. Another possible motivation might be that there is something incoherent about the notion of a dispositional state—any respectable ontological assay of states would cast them as occurrent entities. To say that a **state** is dispositional is really a confusion. This kind of argument may be harder to prosecute, but my hunch is that it is fundamentally sound.

84. It is a fair question whether there is indeed an intuitive cost. On the one hand, it does sound counter-intuitive to deny that most people want (right now) to stay alive. But on the other hand, it is not counter-intuitive to deny, and is in fact counter-intuitive to assert, that most Americans believe (right now) that Michael Jordan is not a three-headed kangaroo. Overall, I do think there is an intuitive cost here, and recognize this cost to be a liability on the view that there are no dispositional states. Nonetheless, I think the cost is worth paying for the sake of the economy in states posited.

85. I take this eliminativism about dispositional states to cover both tacit beliefs that were never occurrent and those that were once occurrent and then were stored in long-term memory. It may seem

## 60 THE SOURCES OF INTENTIONALITY

that with the latter type of state there is more reason to be realist. But again there is no explanatory gain in preferring stored dispositional states over stored dispositions to enter states. (The only new pressure presented by stored states is that the intuitive cost of denying their existence seems greater.) In the case of beliefs stored in *short*-term memory, there seems to be genuine explanatory gain in admitting their existence, but this is mainly because such states are most certainly occurrent (short-term memory being what it is).

86. I use “broadly experiential” because I hold that only experiential states, in the sense of “experiential” I am working with here (to do with having a warranted sense of an explanatory gap), are introspectible. But if one insisted that there are introspectible states that are not experiential in the sense of warranting explanatory gap worries, then there is surely some broadened sense of “experiential” such that in it only experiential states are introspectible.

87. It is possible that, as with the attribution to oneself of Freudian *nexp*-intentional states, there could be indirect introspective evidence for the existence of such sub-personal *nexp*-intentional states. But the remarks I made above about the case of Freudian *nexp*-intentional states would apply here too: since there is no direct observation of the state being ascribed, the vector-of-two-forces observation would still hold.

88. Block (1995, 2007) has argued that some sub-personal state may be, and probably are, phenomenally conscious. Since Block means by “phenomenal” what we mean here by “experiential” (see Block 1995: 382 in the 1997 reprint), if he is right there would be *exp*-intentional states that are sub-personal and thus introspectively inaccessible. There would therefore be some first-person *exp*-intentional ascription that would behave just as this form of first-person *nexp*-intentional ascription does. As before, this would not undermine the epistemological asymmetry between *exp*- and *nexp*-intentional states, since the asymmetry requires only that *some* *exp*-intentional state be ascribed in the special way mentioned in the text.

89. Either so, or else they are imbued even more directly with charity and normativity, as we have seen in some cases of third-person *exp*-intentional ascription.

90. Interestingly, this reaction to claims that all intentional ascription is governed by principles of charity is structurally similar to an early reaction due to McGinn (1977). McGinn’s discussion is complex, but let me crush the subtleties and present what I take to be its upshot: ascription of *de dicto* intentionality may be based on data that are the vectors of two forces, but ascription of *de re* intentionality is not. Consider the person who utters the words “This balloon is yellow” while pointing at a yellow refrigerator. As interpreters, we may be concerned to ascribe to her either beliefs *de dicto* or beliefs *de re*. Suppose we wish to ascribe beliefs *de dicto*. Our data are consistent with two interpretations: (i) she believes that the refrigerator is yellow and by “balloon” she means refrigerator; (ii) she believes that a balloon is yellow and uses words as we do. But now suppose we wish to ascribe to this person beliefs *de re*. While the data are consistent with saying that the person believes, of a refrigerator, that it is yellow, they are *inconsistent* with saying that the person believes, of a balloon, that it is yellow. If there is no balloon, it is false that the person believes anything of a balloon. This reaction is structurally similar to ours, in that it indicts the normativist about intentional ascription with failure to make certain distinctions, and subsequently failure to appreciate that the vector-of-two-forces observation applies only to some kinds of intentional ascription and not others. It is worth noting, however, that *de re* intentional ascription is extensional (Quine 1956), whereas *exp*-intentional ascription is intensional (i.e., features substitution failure and failure of existential generalization). So even if McGinn is right (as I think he is), I would still maintain that *exp*-intentional ascription provides the only case of intensional intentional ascription that is non-normative.

91. It is an interesting question how much of philosophical research can be cast in these terms, that is, as involving empirical claims supported by considerations that do not require very involved

empirical research but are available, in some sense, to the casual observer. In any case, it should be noted that the various clarifications of the anchoring–instance model in §1.1.1, in presenting a number of ways the model could go, make it clear that the level of specificity at which the model is presented is indeed very low, and thus susceptible to understanding in these terms.

92. The notion of transparency of experience will be discussed in greater detail in §2.1.2.

93. How to do justice to the transparency observation (which I accept) given this fact is a good question, but the displaced perception model cannot be right. For my part, I am persuaded that experience is transparent in the sense that its vehicular properties are not available to direct introspection. But it seems to me not to follow that experience is itself unavailable to direct introspection. On the contrary, an experience is directly introspectible due to the fact that its content properties are directly introspectible. For further discussion, see Kriegel 2009 Ch.5.

94. Furthermore, it is hard for me to believe that the model can be successfully applied across the board. Even if made to work for perceptual experiences, it faces severe difficulties as a model of our introspective access to certain *imaginative* experiences. When a visual image of a smiling octopus pops up in my mind arbitrarily and involuntarily, I can introspect it, but no displaced perception can take place. It is not as though I become aware that I am having an experience as of a smiling octopus on the basis of being perceptually aware (or seemingly-aware) of an octopus. It might be claimed that I *pretend* to be perceptually aware of an octopus, but of course pretending is itself an imaginative exercise, so this account would be circular and vacuous.

95. Presumably, this would have to be conjoined with the claim that thoughts, or intellectual states more generally, do not involve a kind of epistemic contact with their objects. This second claim is denied by some (e.g., Chudnoff forthcoming-b).

96. It might be objected that this intellectualist view of introspection can explain the asymmetry by noting that, even if introspecting an experience is just thinking about it, the relevant thought would be formed on the basis of data that are not vectors of two forces: it is possible to think that I am having a table experience *without* at the same time thinking that I am in any other intentional state. However, this is to confuse the *explanandum* for an *explanans*: it is precisely this asymmetry that is in need of explanation!

97. This is in fact my own view—see Kriegel 2009 Ch.5 and Horgan and Kriegel 2007.

98. Furthermore, it is not clear to me that acquaintance cannot be thought of as a form of observation: that the numerical identity of the state of acquaintance and the state acquainted rules out the appropriate descriptibility of the latter as “observed.” This is an entirely verbal matter, however, that we need not go into. The substantive issue, as stressed in the text, is that of direct epistemic contact.

99. This may be an instance of the more general fallacy of confusing *causal immediacy*, which is a matter of lack of intervening causal links, and *temporal immediacy*, which is a matter of lack of temporal lag (or, more accurately, *noticeable* temporal lag). The above ascription of distress to one’s friend may be immediate in the temporal sense, but not in the causal one. To show this, however, we would need an individuation of causal links and temporal interval to support this, making it the case that seeing the facial expression is a mediating causal link in the case of ascription of distress to the friend.

100. This does not require denying that a mature subject, possessed of a concept of intentionality, can be perceptually aware of next-intentional item *as intentional*, as per Siegel’s (2006) claim. It is only to deny that a subject *lacking* the concept of intentionality can.

101. It may be objected that the intentionality of paintings and photographs is not convention-based in the way that of language is. However, while it is not entirely clear to me whether the resemblance between a portrait and its sitter is conventional, for resemblance to amount to intentionality certain further conditions must be met, and those are likely to involve convention. Indeed, as we will see in Chapter 4, one quite plausible view is that the intentionality of pictures is structurally the same as

## 62 THE SOURCES OF INTENTIONALITY

the intentionality of language and the kind of Gricean basis required for linguistic intentionality is also required for pictorial intentionality (Abell 2005, Blumson 2006).

102. It is a separate question whether there is an intentionality conscious experiences have in virtue of being conscious experiences. If there is not, then there is no experiential intentionality. But what it takes for there to be experiential intentionality is for conscious experiences to have a kind of intentionality they have purely in virtue of being conscious experiences. The question of existence is taken up in §1.3.2.

103. I should clarify that I use “conscious experience” so that every experience is conscious and every conscious state is an experience. I prefer this term over “phenomenally conscious state” because it is less technical, and for other reasons that will come out later in this subsection.

104. By “blindsight perception” I mean a perceptual state of the sort that occurs in blindsight (a condition that issues in subjects being able to navigate the environment on the basis of visual processing of information about the environment but do not undergo corresponding visual experiences—see Weiskrantz 1986).

105. Another possible formulation of the same idea would be this: A priori, for any intentional property  $P$  and state  $X$ ,  $P$  is and exp-intentional property of  $X$  iff there is an experiential character  $E$ , such that (i)  $X$  is  $P$ , (ii)  $X$  has  $E$  and (iii) it is the case that (i) because it is the case that (ii). Note well, however: the “because” in (iii) should be read not causally but constitutively, that is, not in the sense in which I might say “I am a bachelor because I never met the right woman,” but in the sense in which I might say “I am a bachelor because I am an unmarried man.” For a well-behaved gloss on apriority, which might help in this particular context, we could appeal to the notion of epistemic possibility (see Chalmers forthcoming, Huemer 2007).

106. This is akin to defining color as the determinable of which green, yellow, black, etc. are determinates. The claim that there is such a thing as experiential intentionality then becomes the thesis that there is at least one exp-intentional property.

107. Which definition we go with depends on our metaphysics of states in general. Option (a) works with an account of states as dated property exemplifications, akin to the way events are understood in Kim’s (1976) account of events. Option (b) works with an account of states as bare particulars that **have** properties, rather than being partially **constituted** by properties, akin to Davidson’s (1969) account of events.

108. I have in mind here a notion of grounding with the kind of robust primacy content, along the lines developed by Kit Fine (2001).

109. We may say that where the doctrinal definition of experiential intentionality requires **meta-physical** dependence of the intentional on the experiential, the neutral one requires only **counterfactual** dependence.

110. The main issue not to prejudge is whether the experience is more basic than intentionality. While I am seeking an account of intentionality that gives pride of place to experiential intentionality, it is not part of my project to establish the primacy of experience over intentionality within experiential intentionality. More on this momentarily.

111. A term I am much fonder of is “subjective intentionality.” However, Loar (1987) used this term early on, and I do not wish to acquire the baggage he associated with it. I also prefer “experiential intentionality” to “conscious intentionality”—a term used by McGinn (1988), Searle (1992), Kriegel (2003a), and Georgalis (2006)—because the literature on consciousness these days features a variety of concepts of consciousness, some of which (e.g., “access consciousness”) have no recognizable experiential, or phenomenal, or first-personal dimension (see Block 1995).

112. Relatedly, the term “phenomenal intentionality” is sometimes heard as definitionally non-reductive, so that asserting the existence of phenomenal intentionality implies the existence of physically

irreducible intentionality. Some also hear it as denying the divisibility of the mind into two broad categories, the sensory and the cognitive. And so on and so forth. By contrast, the term “experiential intentionality” does not come with any doctrinal baggage.

113. If we want to give an experiential property E metaphysical primacy over an intentional property P within experiential intentionality, we must keep the two properties distinct and posit some intimate relation between them, such as logical supervenience, constitution, or realization. If we do not want to give any primacy to the experiential property, we are free to claim that the experiential property and the intentional property are one and the same (Pautz 2008, Mendelovici 2010). I must say that I find myself most attracted to this view of the matter. There may be a way to adopt this identity view while giving the intentional some **explanatory** or **conceptual** primacy, however. For example, one might hold that there is an experiential concept and an intentional concept that pick out one and the same (experiential) property, but that the former concept is in some sense prior to the latter (e.g., in that the former can be used in explaining the latter more easily than the converse). There are some obvious difficulties presented by these notions of conceptual and explanatory priority, but a full discussion will take us too far afield anyway.

114. Thus, Georgalis (2006) argues against the notion of phenomenal intentionality, even though his view of intentionality is very much in line with the one I wish to defend here (as we will see momentarily, as well as in Chapter 4). This seems to be due to his understanding “phenomenal” as “sensory.”

115. I suppose one could also claim that the mental states of understanding and realizing I just described do not qualify as experiences. But here we can appeal to the everyday usage of the term “experience” to reveal the implausibility of this claim, something we cannot do with a technical term such as “phenomenal,” since the latter has no everyday use.

116. I will say more about cognitive experiences in §1.3.3. Their existence is somewhat controversial, but is defended in Goldman 1993, Strawson 1994 Ch.1, Peacocke 1998, Siewert 1998 Ch.8, Horgan and Tienson 2002, Kriegel 2003b, Pitt 2004, Klausen 2008, and Tennant 2009. See also Bayne and Montague 2011. My fullest defense of it is in Kriegel forthcoming.

117. This sense may be revealed by conceptual analysis, perhaps through the Ramsey sentence for “experience.”

118. This seems to me overly optimistic, insofar as intuitions about what is such that there is something it is like to have it seem to align miraculously with philosophers’ theoretical commitment about what is such that there is something it is like to have it, and thus raise the suspicion that the intuitions lie downstream of theory.

119. In Kriegel 2009 Ch.1, I propose—following a cue from Block (1995)—to define “phenomenal” in terms of the explanatory gap, as suggested here for “experiential.” If this proposal is adopted, the two terms become equivalent of course. (But, also of course, some philosophers do not adopt this proposal, and instead hear “phenomenal” as having to do with the sensory/somatic.) Chalmers (1995, 1996 Ch.1) uses the term “experience” for whatever feature of conscious states presents the Hard Problem of consciousness. If the Hard Problem is the same as the explanatory gap, this would be exactly the definition proposed in the text. I happen to think they are slightly different (Kriegel 2009 Ch.8), but very close. It is possible to hold, of course, that the understandings of “experiential character” in terms of the everyday notion of experience and in terms of the explanatory gap are equivalent: plausibly, at least in one of its uses, the mundane notion of experience coextends with the explanatory-gap-based notion.

120. This requires a number of qualifications and clarifications, which I provide elsewhere (Kriegel forthcoming). The reference-fixing description I end up with there is this: For any property P, P is an experiential property iff in the actual world, there is at least an appearance of a non-derivative empirical explanatory gap, relative to the human intellect, between P and physical properties. The qualification of “empirical” explanatory gap is supposed to rule out potential explanatory gaps between normative

## 64 THE SOURCES OF INTENTIONALITY

properties and physical properties. The qualification of “non-derivative” explanatory gap is supposed to rule out potential explanatory gaps parasitic on the explanatory gap between conscious experience and physical matter; thus, it has sometimes been claimed that there is an explanatory gap between color and physical matter, but also that this latter explanatory gap exists only insofar as color is tied up with color consciousness in such a way that an explanatory gap arises. As for the relativization to the human intellect, it is mandated by the fact that the availability of explanations depends partly on the cognitive abilities of the explanation-seeking subjects. For fuller discussion, see Kriegel forthcoming.

121. This elucidation of “experiential character” allows us to formulate an even more explicit statement of the neutral definition of an exp-intentional property: A priori, an intentional property *P* is exp-intentional iff there is a mental state *X* and a property *E*, such that (i) *X* is *P*, (ii) *X* has *E*, (iii) in the actual world, there is (or appears to be) an explanatory gap between *E* and physical properties, and (iv) if *X* were not *E*, *X* would not be *P*.

122. A more explicit way to formulate the thesis, given the neutral definition of an exp-intentional property, would be this: (EI) There is a state *X*, an intentional property *P*, and an experiential character *E*, such that (i) *X* is *P*, (ii) *X* has *E*, and (iii) if *X* did not have *E*, *X* would not be *P*.

123. Typically, the thought is that the experiential character of a conscious experience is the **medium** of intentionality rather than **message**. This is what McGinn (1988) calls the “medium conception” of experience’s intentional role.

124. Certainly introspection suggests that some mental states have both intentionality and experiential character. Whether introspection reveals that sometimes the former counterfactually depends on the latter is a harder question. Admittedly, it is implausible that introspection presents any counterfactual dependence **as** counterfactual dependence. But it is much more plausible that it presents what is in fact counterfactual dependence under a simpler guise, but in such a way that a sufficiently sophisticated theoretician could justifiably conceptualize what is presented as counterfactual dependence.

125. One thought-experiment that convinces me of the existence of experiential intentionality is this. One can conceive of a disembodied soul in an otherwise empty world who is experientially indistinguishable from oneself. Intuitively, portions of the soul’s inner life are also **intentionally** indistinguishable from corresponding portions of one’s own inner life. The fact that some experiential duplication secures intentional duplication provides evidence (and according to some, may even entail) that some experiential properties are sufficient for intentional properties. This brings us very close to counterfactual dependence of the latter upon the former.

126. I have already mentioned Siegel’s claim that we can perceive the meaning of linguistic expressions, but she makes similar claims about causation (e.g., we can see that one billiard ball causes the motion of another) and belonging to a kind (e.g., we can see that something colorful and shapely is a parrot—see also Siewert 1998 Ch.7). A related and quite central debate concerns the presentation of perspectival properties in conscious experience. Suppose you look at a tilted coin. Does your perceptual experience present the coin as (having the non-perspectival property of being) circular or as (having the perspectival property of being) elliptical? Kelly (2004) argues that only the (non-perspectival) circularity is presented in experience, whereas Noë (2004 Ch.5) claims that both the (non-perspectival) circularity and (perspectival) ellipticity are presented.

127. It is sometimes argued that there is an experiential residue in emotion that is better understood in terms of intentional directedness at bodily states (see, e.g., Armstrong Ch.8). Thus, disappointment with anything may be accompanied by a feeling of “sinking heart,” and that feeling is best understood as nothing but intentional directedness toward the internal bodily event in the vicinity of the heart. This view strikes me as wrongheaded, insofar as the bodily experiences in question are not part of the **proprietary** emotional experiential character, instead being contingent accompaniments more accurately classified as self-standing somatic experiences.

128. Associated with this is the experiential character of agency (trying and/or doing). This is argued for most thoroughly by Horgan and collaborators (see Horgan, Tienson, and Graham 2003 for agentive experience in general, and Horgan and Timmons 2008 for moral experience more specifically). For my part, I accept the existence of agentive experience not only on introspective grounds, but also due to an argument from Merleau-Ponty (1944: 93). Merleau-Ponty notes that when we rub our hands together, we can experience, more or less voluntarily, first the right hand doing the rubbing and the left hand being rubbed and then the left hand doing the rubbing and the right hand being passively rubbed. The strictly tactile experience in those two moments is the same, but there is a residue in the **overall** experience that is **not** the same. That residue concerns which hand is agent and which is patient, so to speak.

129. Perhaps the strongest thesis of the second kind is Pitt's (2004) claim that thoughts and beliefs have an experiential character which is both **proprietary** and **individuating**, that is, an experiential character that both is different from all other types of experiential character and varies whenever the content varies. Weaker theses might suggest that there is an individuating but non-proprietary cognitive phenomenology (say, because every belief is accompanied by imagery, and the imagery varies with the belief's content); or that there is a proprietary but non-individuating phenomenology (say, because although beliefs in general feel different from hopes, suppositions, and other propositional attitudes, it is not the case that a belief that **p** feels different from a belief that **q** for any **p** and **q**). In Kriegel 2003b, I consider a number of theses in decreasing order of ambition.

130. Most existing arguments for cognitive phenomenology fall under one of these three categories. Nonetheless, some arguments do not fit comfortably into these. They include arguments by Lurz (2006), Klausen (2008), Tennant (2009), and Mendelovici (2010).

131. The passage is taken from Klein 1981: 83.

132. Similar arguments are presented in Goldman's (1993) and Mangan's (2001) discussion of tip-of-the-tongue phenomenology, Siewert's (1998 Ch.8, 2011) and Horgan and Potrč's (2010) discussion of delayed understanding, and Siewert's (2011) and Chudnoff's (forthcoming-a) discussion of intellectual Gestalt shifts. The tip-of-the-tongue experience is familiar to most of us and is claimed by Goldman and Mangan to involve a distinctively cognitive phenomenology. Siewert's delayed understanding concerns experiences in which we suddenly understand a piece of text we have been rereading and trying to understand for a while (imagine suddenly thinking of kites on your own when reading the above passage). Intellectual Gestalt shifts, meanwhile, concern experiences in which a phenomenon is construed one way but is suddenly reinterpreted another way; Siewert (forthcoming) illustrates this with an entertaining exchange he overheard in Miami on an unbearably hot day, in which one woman said to the other "I am so hot" and the other responded "You don't have to brag about it."

133. For other arguments of this form, see Peacocke 1998, Siewert 1998 Ch.8, Horgan and Tienson 2002, Kriegel 2003b, Pitt 2004, and Chudnoff forthcoming-a.

134. Schematically, the argument proceeds as follows: we have a special, immediate access to our cognitive states (and their intentional contents); only to conscious experiences (and their contents) can we have this kind of special access; therefore, our cognitive states (and their contents) are conscious experiences (and exp-intentional).

135. An earlier argument of a similar form is developed in Goldman 1993 and recently endorsed in Lycan 2008.

136. I would like to thank Galen Strawson and Angela Mendelovici for openly sharing this sentiment. Strawson once put things to me very succinctly: we know that there is cognitive experience because without it life would be boring. Mendelovici once wondered aloud whether opponents of cognitive experience do philosophy because they enjoy the accompanying imagery so much.

137. For details, see Kriegel forthcoming.

## 66 THE SOURCES OF INTENTIONALITY

138. For comments on a previous draft of this chapter, I thank Ned Block, Brian Fiala, Jesse Prinz, anonymous referees for Oxford University Press, and especially Angela Mendelovici, Stephen Pearce, and Lee-Anna Sangster. For comments on material from this chapter, I thank Michael Bruno, David Chalmers, and Farid Masrour. I have also benefited from conversations with David Bourget, Jordi Fernández, Lizzie Graf, Iris Oved, Declan Smithies, and especially Farid Masrour, as well as audiences at La Trobe University, New York University, University of Arizona, University of California—Riverside, University of Western Ontario, and a workshop on phenomenology and intentionality at the University of Modena e Reggio Emilia (I thank in particular Gillian Barker, John Bigelow, Frank Jackson, Marc Johansen, Benji Kozuch, Michelle Montague, Jack Reynolds, Eric Schwitzgebel, Charles Siewert, Alberto Voltolini, Meghan Winsby, and probably a number of other people I am forgetting).