

The Acquaintance Argument for Intrinsic Intentionality

Philip Woodward

ABSTRACT: Naturalistic theories of intentionality identify intentional properties with complex functional/dispositional profiles within cognitive systems. Whatever the merits of these theories, they cannot account for the intentional properties instantiated in our conscious mental states. Here's why: such functional/dispositional properties are not intrinsic features of conscious mental states. We can only be acquainted with the intrinsic features of conscious mental states, and the intentional properties of our conscious mental states are among the intrinsic features with which we can be acquainted—as evidenced by the fact that introspective acquaintance guides our ability to recognitionally sort conscious mental states with respect to their content.

1. Introduction.

We creatures with minds are able to instantiate mental states that have intentional content.¹ Examples of intentional states include: believing that roses are red; contemplating the Pythagorean Theorem; intending to pick the children up from school; registering a stranger's gesture as a threat; imagining an ostrich; fearing one's mortality. For a state to be intentional is for it to be *about* or *directed at* something, such as a proposition, a property, an object, an event, or a state of affairs. Intentional directedness is one of the two traditional “marks of the mental” (the other being consciousness).

Intentional properties, as I will understand them, are those properties of intentional states that individuate such states by their contents, rather than according to their psychological modes (believing, wondering, hoping, imagining, thinking, etc.).

Examples of intentional properties include: *meaning that p*; *representing F*; *being about*

¹ I use the term “state” to refer to a property-instantiation at a time or to multiple property-instantiations at a time, by the thing in question. A *mental state* is the instantiation of *mental properties* by a thing at a time. The total mental state of a thing at a time is the cluster of mental properties instantiated by that thing; parts of this total state—proper subsets of the cluster of instantiated properties—are also states. I'm not picky about times: they can be momentary or not.

o, referring to *o*. The task of the present paper is to make progress in understanding the nature of intentional properties. In particular, I argue that some intentional properties are, essentially, *intrinsic* (in a sense to be spelled out) to the mental states in which they are instantiated.

I will focus on intentional properties that are (a) *purely qualitative*, (b) *potentially conscious*, (c) *subjectively presentable* intentional properties. By *purely qualitative* intentional properties, I mean those that are not individuated in terms of concrete particulars. For example, the property of being about unicorns and the property of being about stubbornness are purely qualitative intentional properties, whereas the property of being about *that* object, or being about *The Metropolitan Museum of Art* are not.

By *potentially conscious* intentional properties, I mean intentional properties that can be instantiated in a mental state for which there is something it is like to be in that state, for the subject whose state it is. Hence I am not directing my inquiry at those intentional properties instantiated in standing belief-states, non-accessed memories, suppressed desires, or representational states implicated in pre-conscious neural processing. It is of course a live possibility that the very same intentional properties are instantiated in conscious and non-conscious states.

By *subjectively presentable* intentional properties, I mean intentional properties to whose instantiation in a state a subject can have complete first-personal access. There are many merely third-personal ways that mental contents can be characterized. For example, a subject's perceptual state might be directed at H₂O, but if the subject does not *take* the stuff perceived as H₂O, this way of characterizing her perceptual state does not reflect its

subjectively presentable content. It is in virtue of instantiating subjectively presentable intentional properties that mental states present intentional contents *to the subject* in the particular way that they do.

Call such properties “paradigmatic intentional properties.” Why focus my inquiry on intentional properties of just this sort? The reason is that they are the only intentional properties whose instantiation we are aware of from *within*, as follows. When, in contrast, we attribute intentional properties third-personally on the basis of an inference to the best explanation, the following questions can arise: Might a different, non-intentional explanation turn out to *better serve* the data? Might explanations in terms of intentional properties perform a practically indispensable but ultimately *merely heuristic* function? Might there be nothing more to *which* intentional properties are instantiated than the successful honoring of some *charity principle* or other? I contend that these challenges to realism do not and cannot come up for intentional properties whose instantiation we are aware of from within. To say that we are aware of instantiations of intentional properties from within is not to say that we have infallible introspective knowledge of the content of our mental states. It is to say that we could not be wrong about *whether* our mental states have content, *or whether there is a fact of the matter* about the content of our mental states.

2. Naturalistic Theories of Intentional Properties

If you want a theory of intentional properties that invokes only concepts that enjoy good standing in current natural science, you need a strategy for “naturalizing” intentionality. This has been a central philosophical project in the philosophy of mind and cognitive

science at least since the middle of the 20th century, and especially since the work of Fred Dretske, Jerry Fodor and Ruth Garrett Millikan in the 1980s. We can identify four families of strategies that have emerged. All of these accounts characterize intentional properties as essentially *non-intrinsic* to the states in which they are instantiated. If the arguments of the present paper are correct, it follows that all of these accounts are false, at least when it comes to characterizing paradigmatic intentional properties.

*1. Causal covariance theories.*² I begin with the following two observations. First, we often exploit causal relations in order to transmit information about our environment and in order to extract information about our environment.³ (Whenever we communicate over the phone we're so exploiting.) Second, we often attribute contents to cognizers on the basis of the way her environment is causally affecting her. (Whenever we interpret someone's behavior in terms of beliefs/desires regarding information perceptually available to her, we're so attributing.) The covariance theorist goes a step further and *identifies* intentional properties with the holding of certain causal relations between a cognizer and the environment. Roughly, when the instantiation of some property covaries with the activation of a certain sort of structure within the cognitive system, that structure counts as a representing of that property. We might think of the relevant type of activation as a state that plays the functional role of an existentially-quantified perceptual belief, e.g.: "Lo, a ____." (I will assume for present purposes that it's possible to give a complete characterization of perceptual belief in functional terms.) To use a familiar example from this literature, we could say that a frog's mental state represents *flyhood* if

² See e.g. Dretske (1981) and Fodor (1990).

³ I'll be framing naturalistic theories in terms of causal relations, though some theorists (e.g. Dretske 1990) might prefer nomic relations. Pretty much everybody wants to reduce one to the other, so I don't think anything turns on my choice.

a part of that mental state would properly activate in the (right sort of) causal presence of flies. Put more formally: let ‘F’ stand for a subpropositional content; let ‘O’ stand for the relevant class of cognitive system;⁴ let ‘M_O’ stand for a mental-state type that the relevant systems can instantiate; let ‘\$’ stand for some content-bearing structure within a system; let ‘C’ stand for some causal relation; and let ‘A’ stand for the type of activation in ‘O’ that realizes the functional profile of perceptual belief:

Causal Covariance Theory (CCT): M_O has the content F if: \$ is part of M_O, and \$ would A were an F-instance to bear C to O.

Activation A will get cashed out differently by different models of cognitive architecture. Perhaps the clearest picture is that supplied by the Computational Theory of Mind, on which A-ing is the “tokening” of \$ in the system’s perceptual-belief box. \$ itself could be a physical structure (a particular neuron, say, or the firing of that neuron) or something a little more abstract, such as a pattern of neural activation, or something *much more* abstract, such as a symbol in the language of thought (where *being a symbol in the language of thought* can itself be naturalized).

In short, according to CCT, a mental state has the content it has because it is built out of content-bearing structures; and these structures have the content they have because they are disposed to be activated (as components of perceptual beliefs) by environmental triggers. There are a number of complicating factors for such theories, but perhaps the most conspicuous such factor is this: sometimes perceptual beliefs *misrepresent*: perceptual beliefs can be caused by things or states of affairs that they are not about. So the contents of perceptual beliefs cannot just be given by what they covary with.

⁴ ‘O’ is for Organism, though I don’t propose any particular way of delineating that category; I will take for granted that there are such things as cognitive systems whose individuation-conditions can be spelled out in non-circular, realist terms.

Otherwise, we would not be able to mistake sugar for salt. Since we do sometimes token the thought “that’s salt” in response to sugar, the simple covariance theory would have to say that the thought’s content is really “that’s salt-or-sugar (since both substances trigger type-identical perceptual judgments). So a viable version of CCT needs to specify those circumstances in which bearing C to an F-instance fails to be determinative of content—when, as Fodor puts it, “wild tokenings” of \$ occur. Error-conducive circumstances (or “E-circumstances,” for short) could be built directly into the specification of C, or they could be tacked onto CCT in the form of a qualifying clause (“except when...”).⁵

2. *System-role theories.*⁶ The theories I have in mind are versions, or at least close cousins, of covariance theories, in that they identify intentional properties in terms of causal relations that hold between states of a system, on the one hand, and instances of properties in that system’s environment, on the other. But they deal with the problem of misrepresentation otherwise than by adding a clause specifying E-circumstances. Instead, what makes it the case that a system can misrepresent—that its G-caused A-ings do not automatically endow \$ with a disjunctive content $F \vee G$ —is a matter of the system’s itself having certain representational *purposes*. \$, that is, has the *function* of representing F.

As best I can tell, two strategies have been proposed in the literature for grounding the relevant teleological properties: one built around a system’s *homeostatic needs*, and another around a system’s *evolutionary ancestry* (i.e. the selection pressures that caused ancestor-systems to evolve in the way they did). System-role theorists take

⁵ Several accounts of E-circumstances can be found in the literature. According to Fodor (1990), E-circumstances are those in which something other than an F-instance causes \$’s A-ing—a G-instance, say—but would not have, were F-instances not apt to do so as well (i.e. G-caused A-ings are “asymmetrically dependent” on F-caused A-ings). Dretske (1988) has suggested that E-circumstances are those in which G-instances cause \$’s A-ing after the culmination of O’s “learning period.” Hill (ms) thinks we can just construct a list of E-circumstances.

⁶ See e.g. Dretske (1986) and (1995); Bickhard (1994) and (ms).

the first strategy. According to system-role theories, a frog's mental state represents *flyhood* if part of that mental state is such that the frog *needs* it to properly activate in the causal presence (of the right sort) of flies. More precisely,

System-Role Theory (SRT): M_O has the content F if: $\$$ is part of M_O , and it would contribute to O's homeostasis were $\$$ to A in the event that an F-instance bore C to O, and it would detract from O's homeostasis were $\$$ to A in the event that a G-instance bore C to O.

SRT thus ties intentionality to usefulness; a state of an organism has a content when that state guides the organism toward its own upkeep.

3. *Adaptive-role theories.*⁷ SRT understands a structure's representational purposes in terms of the *needs* of a system; a second way to understand such purposes is in terms of the *design* of a system. According to such a view, a frog's mental state represents *flyhood* if the frog's ancestors gained a selection advantage by having a part of type-identical mental states properly activate in the (right sort of) causal presence of flies. More precisely,

Adaptive-Role Theory (ART): M_O has the content F if: $\$$ is part of M_O , and it contributed to O's ancestor's selective advantage when ancestral correlates of $\$$ A'ed in the event that F-instances bore C to O's ancestors, and it did not contribute to O's ancestor's selective advantage when ancestral correlates of $\$$ A'ed in the event that G-instances bore C to O's ancestors.

ART thus ties intentionality to adaptational history; a state of an organism has a content when the organism's capacity for instantiating that state is a pretty direct result of a natural-selection process.

⁷ See e.g. Millikan (1984) and (2009).

4. *Conceptual-role theories.*⁸ The three families of theories we have just looked at all pin down content in terms of property-instances instantiated in the environment: how an organism would respond if stimulated by those property-instances (CCT), how an organism needs to respond if stimulated by those property-instances (SRT), and how an organism's ancestors responded when stimulated by those property-instances (ART). Conceptual role theories set aside this strategy entirely. Instead, they pin down content in terms of intra-systemic relations. According to these theories, a frog's mental state represents *flyhood* if that state's functional relations with other mental states map onto the inferential relations that hold between various fly-related propositions. To capture this idea, it will be helpful to change our notation slightly. Let 'B₁', B₂' and 'B₃' stand for types of activation that play the functional role of belief in O, and let 'B(\$)' (and the like) stand for an activation-type that includes \$ in such a way as to functionally realize a belief about or involving whatever \$ denotes. Let 'P₁', 'P₂' and 'P₃' stand for propositions, and let 'P(F)' (and the like) stand for propositions about or involving Fs or F-ness.

Conceptual-Role Theory (CRT): M_O has the content F if: \$ is part of M_O, and were C-relations to hold between B₁(\$), B₂(\$), B₃(\$), etc., these relations would be isomorphic with I-relations between P₁(F), P₂(F), P₃(F), etc. (except in E-circumstances.)

Precisely because CRT does not make mention of environmental states of affairs, it is uniquely suited to explain how mental states can be directed at anything with which the system has no causal interface, such as the identity-relation or the addition-function. So, for example, according to CRT, a structure expresses *plus* only if certain counterfactuals

⁸ See e.g. Harman (1982) and Block (1986).

are true of it, such as: were it functionally connected in the right sort of way with structure that express numerical contents, it would causally contribute to the activation of a state that expresses the sum of those numerical contents.

The problem of misrepresentation is particularly tricky for CRT. In fact, I suspect it is unsolveable, when CRT is taken on its own, though there is no space here to explain why.⁹ The obvious solution is *not* to take CRT on its own, but rather to combine it with one or more of the above theories. It would not be fruitful here to explore all the possible ways that the four theories could be combined, to construct a comprehensive theory of content. I will briefly sketch a hybrid version that looks to me as strong a version as any. Here is the basic idea: a cognitive system has representational *modules* with particular representational purposes, where these purposes are fixed either by way of evolutionary history or homeostatic needs. The purposes of the modules will be somewhat general, e.g. *to indicate objects, to indicate causes, to make conceptual inferences*, etc. These modules will generate particular representations, whose content is not fixed by the general representational purposes of each module, but rather by (a) patterns of causal dependence between representations and properties in the environment (per CCT), or (b) patterns of causal dependence among representations (per CRT). Presumably these meager representational resources can be combined to form new representations and to encode new representational purposes, on the basis of which an intentional structure of arbitrary complexity and power can be iteratively built.

⁹ Here's the basic idea: irrational transitions are not content-fixing. But what makes a transition between two states rational or not is at least partly a matter of the content of those states. There is no way to distinguish content-fixing from non-content-fixing transitions that does not make mention of the system's intentional states as such.

Our brief exposition of the program of naturalized intentionality is almost complete, but not quite. We have framed the four strategies in terms of *sufficient conditions for a mental state's having some content*. But intentionality has not been “naturalized”—shown to be locatable within the ontology of the natural sciences—until the *natures* and not just the *instantiation-conditions* of intentional properties have been framed naturalistically. That is, it is consistent with the characterizations so far given of the four theories that intentional properties are supervenient, non-natural properties. So, how can we extract a naturalistic metaphysics of intentionality from the characterizations given in the present section? Here is one way: *identify* intentional properties with the functional/dispositional properties delineated by the right hand side of one or more of the schemas above. Going this route would have the odd result that *there is no property* of primitively being about something; that is, not even God can think about F-ness without instantiating the right sort of functional/dispositional property. A second approach would be to say that the relevant functional/dispositional properties merely *constitute* intentional properties, thus amounting to one of multiple ways that intentional properties can be “realized.” But absent a clear understanding of the realization-relation in this case (we understand how functional states are realized in physical states, but here the question is how intentional properties are realized in functional/dispositional properties without being identical to them), such talk is unilluminating. I propose a third option: advocates of naturalized intentionality should understand their views as characterizing a *species* of intentionality, i.e. one way among multiple ways (the set of which comprise a strongly unified family of properties) for an item to be about another item. Call intentional properties of this species “N-I” properties. To advocate for one or the other naturalistic

theories is to contend that the intentional properties instantiated by *us* (or whatever cognitive systems amount to our theoretical target) are N-I properties— i.e. one or the other, or a combination, of the following four types (characterized more schematically than above):

- (1) *N-I(CC) property* = the presence (in a state of a O) of \$ such that \$ would A were O causally stimulated in certain ways.
- (2) *N-I(SR) property* = the presence (in a state of a O) of \$ such that O needs \$ to A when O is causally stimulated in certain ways.
- (3) *N-I(AR) property* = the presence (in a state of a O) of \$ such that O's ancestors needed their \$-correlates to A when they were causally stimulated in certain ways.
- (4) *N-I(CR) property* = the presence (in a state of a O) of \$ such that B₁(\$), B₂(\$), B₃(\$), etc. would cause each other's activation in O in certain circumstances.

My purpose in this section has not been merely been to summarize the literature on naturalized intentionality, but also to map the relevant conceptual space. That is, there is good reason to think that any naturalistic theory of intentionality is going to amount to a version of one of these four and/or a combination of them. Consider the resources these theories draw on: patterns of causal dependence that hold (a) among internal states of a system; (b) between internal states of system and external states; (c) current states of a system, states of ancestral systems, and external states. If the goal is to locate intentional phenomena within the ambit of causal phenomena (actual and counterfactual), it does not look as though there are any raw materials that the naturalizers have simply skipped over. No, the materials are all out on the table; remaining philosophical work will amount to

refinement and assembly. Hence, the force of the arguments of the present paper is not only to show that *current* naturalistic theories are inadequate, but that *all possible* naturalistic theories are inadequate.

3. The Acquaintance Argument

In the remainder of the paper I argue that paradigmatic intentional properties are not N-I properties. That is not to say that N-I properties are not intentional properties, but only to say that some of our intentional states instantiate intentional properties that are *not* N-I properties. I lay out my argument briefly here and then proceed to argue at length for its premises.

- (1) Some paradigmatic intentional properties are such that subjects can be introspectively acquainted with them.
- (2) Subjects can be introspectively acquainted only with the intrinsic features of their conscious mental states.
- (3) N-I properties are not intrinsic features of conscious mental states.
- (4) Hence, some paradigmatic intentional properties are not N-I properties.

Call this “The Acquaintance Argument.” I will address its premises in reverse order—from least controversial to most controversial.

Premise 3: N-I properties are not intrinsic features of conscious mental states.

For my purposes, a property is intrinsic to a conscious mental state only if it supervenes on the essential properties of that state—i.e., the properties that make it the very type of conscious mental state that it is. So the question at issue is whether N-I properties supervene on the essential properties of conscious mental states. Here is one way we

might answer the question: No—the essential properties of conscious mental states are *phenomenal* properties, and there is no necessary connection between phenomenal properties and the functional/dispositional properties relevant to N-I properties. Whatever the virtues of this response, it isn't dialectically appropriate here, because what drives it is the explanatory gap that holds between phenomenal states and functional states *generally*.

In order to evaluate the question in a dialectically appropriate way, we will need to begin not from a phenomenological conception of conscious mental states but from a functional conception of conscious mental states.¹⁰ There is disagreement over how best to do this, but I gather that at least three components will be involved: (1) a functional specification of the content of conscious mental states (which we have at hand in the form of N-I properties); (2) a functional specification of psychological modes; and (3) a functional specification of a state's *being conscious*. Likely (3) will be framed in terms of a state's *occurrence*—i.e., what makes it the case that the state of a system is *activated* in a special way.¹¹ So: suppose a structure \$ is part of an occurrent state, where \$ has the content *being a snake*, per one of the four naturalized theories; and that content of the occurrent state as a whole is the fear that one will see snakes at the zoo. Our question is whether its N-I properties supervene on its essential features.

It is immediately clear that any N-I(AR) properties it has do not so supervene. Such properties partly supervene on the *actual phylogenetic histories* of systems. No one thinks that a system's tokening some particular occurrent state is a matter of or depends

¹⁰ It is open to the functionalist to maintain that the two conceptions converge on the very same type of state, despite our being unable (at present, or maybe ever) to connect the two conceptions in an explanatorily perspicuous way.

¹¹ Occurrence simpliciter isn't enough, since there are plenty of brain-activities that do not correspond to conscious states. But perhaps occurrence of a certain sort *will* correspond in the right way with consciousness.

on its having a certain evolutionary history. Matters are not so obvious regarding the other three types of N-I property, since those properties are not defined in terms of actual causal histories, but rather in terms of dispositions of \$ to be activated by the system under certain conditions.¹² Nevertheless, I think it is evident upon reflection that N-I properties of the other three sorts also fail to supervene on the essential features of states that instantiate them. Here's the basic idea: there is no necessary connection between \$'s being tokened in an occurrent state, on the one hand, and its playing the content-fixing role that it plays in the system, on the other. The grounding for its content-fixing profile—i.e., whatever it is in the system that makes it the case that \$ plays the role it plays—is included in the supervenience-base of N-I properties. But it's not in general the case that an occurrent state that includes \$ is going to ground \$'s functional/dispositional profile. For example: according to CCT, an occurrent state is about snakes if it contains a structure that would be tokened within a perceptual belief in the event that the system were to stand in the right sort of causal relation to snakes. But what makes it the case that \$ has precisely that functional/dispositional profile turns on the system's having the right sort of sensory apparatus in the right sort of environment. If \$ were embedded in a very different system and/or a very different environment, it would *not* be disposed to be tokened within a perceptual belief in the causal presence of snakes. Similar

¹² Any version of naturalized intentionality that *does* appeal to actual causal histories has the immediate implication that intentional properties are never intrinsic (thus falling prey to the Acquaintance Argument). Though naturalistic theories of intentionality are sometimes characterized in terms of causal histories (owing, I gather, to their resemblance to causal theories of reference), I have not done so, in part to avoid this immediate implication. Here are two other reasons that one might give (but that I reject) for counting N-I properties as extrinsic: (1) N-I properties are dispositional, and dispositions never locally supervene, but instead are instantiated only relative to the laws of nature. (2) N-I properties are dispositional, but only categorical properties are instantiated in conscious mental states. I reject these lines of argument because I accept that there are primitively dispositional properties and because I think that some dispositions are phenomenally given (for example: disposing me to scratch is part of what itches essentially feel like). Advocates of naturalized intentionality who reduce dispositions to laws and who deny that phenomenal properties are essentially dispositional will feel the pull of (1) and (2).

considerations apply when it comes to SRT (which brings in wider organismic needs as part of the supervenience-base for N-I properties) and CRT (which brings in rational transitions between occurrent states).

I can think of two ways that a defender of the intrinsicity of N-I properties might push back. First, she might say that since a mental state's intentional content is essential to that state, the supervenience-base for its intentional properties must be essential to that state; and hence all broader systemic/environmental states of affairs that form part of the supervenience-base *are*, after all, somehow constitutive of the state. One's conscious fear that one will see snakes at the zoo turns out to be metaphysically dependent on one's having the right sort of sensory apparatus, being suitably embodied, having a cognitive system with right sort of inferential capacities, etc. In short, tokening an occurrent state implicates much more of the world than we may have thought it did.

In reply: it is one thing to say that some state's occurring entails much else besides (with nomological, metaphysical, or even logical necessity). It is another thing to say that all of these entailments are *essential to* the state. For example, the lifting of my arm at noon on Tuesday might entail both that I have an arm at noon on Tuesday and that atmospheric oxygen/carbon dioxide levels are within a certain range at noon on Tuesday; plausibly, the former but not the latter is essential, in my sense, to the event in question. Similarly: for all I know, I could never have acquired the concept of a snake were I not suitably hooked up to a snake-saturated environment; nevertheless such causal embeddedness is not *what it is* to have a conscious fear of snakes.¹³ Now, I am not sure

¹³ But aren't semantic externalists committed to saying exactly this? No, they are not (though they *could* say it; Dretske (1996) is one who *does* say it.) Rather, semantic externalists are committed to saying that intentional properties do not supervene on internal states of an organism. And that's consistent with

what criterion a functionalist ought to apply in order to distinguish between what is and is not essential to an occurrent state. Note, though, that some such distinction has to be made on pain of preserving the occurrent/latent distinction. An occurrent state of a system is *what the system is doing*, not *what the system is structured to be able to do*—even if what it is doing is deeply dependent on its having a certain structure.

A second way of defending the intrinsicity of N-I properties would be to alter the way that N-I properties are characterized. For example: instead of saying merely that an N-I(CC) property is the presence (in a state of a O) of \$ such that \$ would A were O causally stimulated in certain ways, one could add: *were \$ embedded in a system of such-and-such a sort and were O embedded in an environment of such-and-such a sort*. (We might say that the strategy works by attributing to \$ a *context-relative intrinsic disposition*.) The idea here would be to render \$'s content-fixing functional/dispositional profile necessary to it, not by expanding the class of items essential to states that include it (per the previous objection) but by precisifying its functional/dispositional profile so that it has that profile *regardless* of which sort of system embeds it.

The strategy has the ring of cheating to it: if it works, then I have the *intrinsic* property of being disposed to be-thought-highly-of-by-the the-President-were-the-President-to-think-highly-of-me. But whether or not someone would think highly of me is surely not intrinsic to me! But there is a deeper problem with the strategy. If a structure has one context-relative disposition, it has countless context-relative dispositions. For example, relative to some contexts \$ would be activated by snakes; relative to others, by frogs; and so on forever. So it will turn out that a state that includes \$ will have *all of*

everything I have said here, so long as mental states can have intentional content non-essentially—which is exactly what we should say with respect to wide content.

these contents. That's the cost of making it such that \$ has its context-fixing functional/dispositional profile necessarily.

I conclude that N-I properties are not features of the conscious mental states that instantiate them.

Premise 2: Subjects can be introspectively acquainted only with the intrinsic features of their conscious mental states. Introspective acquaintance is that direct cognitive relation that a subject bears to her conscious mental states when, and because, she attends to them.¹⁴ Acquaintance is not reducible to any propositional attitude or to any ability, but is rather a form of thing-knowledge, of *knowing-what*. Acquaintance is a species of cognitive relation, of which there are many (imagining, sensing, thinking about, and so on), but it is unique among cognitive relations in that its cognitive object must actually exist. If S thinks about x, it does not follow that x exists. If S is acquainted with x, it *does* follow that x exists.¹⁵

While Premise 2 strikes me as obvious, I doubt it will strike everyone that way. But I think it can be derived from some facts about introspection that are widely accepted—facts which shed light on why it seems so strange to deny it. It is well known that one's judgments about one's own mind are not infallible.¹⁶ There is, nevertheless, a certain form of fallibility that a certain type of introspective judgment cannot have: when

¹⁴ I will sometimes drop the modifier 'introspective'. As a matter of fact I do not think there are any other types of acquaintance (unlike e.g. Russell, who thinks we can be acquainted with universals).

¹⁵ This is because the cognitive relation of introspective acquaintance is always underwritten by a *metaphysical* relation, which we might neutrally characterize as presence-in-consciousness or phenomenal presence. That is, S can be acquainted with x only if x is phenomenally present to S.

¹⁶ There are many sources of fallibility. One can self-attribute motivations for one's behaviors, where such self-attribution is more or less confabulatory, a game of self-interpretation. One can attribute to oneself the belief that p on the basis of one's judgment that p, despite the fact that one lacks a sufficiently stable credence level with respect to p. One can fail to attend sufficiently carefully to the state one proceeds to report. One can lack concepts of sufficient fine-grainedness to accurately capture one's mental states. One can forget what one has just attended, during the time-lag between introspection and the forming of an introspective judgment.

one forms an introspective judgment on the basis of an act of introspective acquaintance with a conscious mental state, one cannot go wrong *by being misled by the appearances*. In introspective acquaintance, appearance and reality go together. Hence acquaintance-based introspective judgments are not subject to what Terry Horgan (2012) calls “appearance/reality fallibility”.

By contrast, judgments about non-intrinsic features of our conscious mental states *do* admit of appearance/reality fallibility. This is because non-intrinsic features of a mental state supervene on states of affairs external to that state, i.e. states of affairs that are no more closely connected to that state than via a causal connection. And appearances can mislead, when it comes to judgments about items that are no-more-closely-than-causally connected to one’s mental states: the familiar effects of a familiar cause can be replicated by a *distinct*, non-familiar cause (even if it would take an evil demon to pull off the stunt.) It follows that one can only form acquaintance-based introspective judgments about items that are more-closely-than-causally-connected to one’s mental states—i.e. their *intrinsic* features, those features that supervene on nothing outside of the states in which they are instantiated.

Premise 1: Some paradigmatic intentional properties are such that subjects can be introspectively acquainted with them. I defend Premise 1 via the following sub-argument:

- (a) It is possible for a subject to cross-modally recognitionally sort her mental states with respect to their intentional content.

- (b) The only adequate epistemic explanation of such a capacity is in terms of a subject's introspective acquaintance with her intentional states.¹⁷
- (c) Such a capacity calls out for epistemic explanation (it is neither epistemically trivial nor epistemically brute).

Here's an illustration of the capacity invoked in (a). A boy works as an apprentice to a potion-master. The potion-master makes concoctions of sundry ingredients, samples of which he has meticulously collected over his many years. The demands on his time prevent him from ever cataloguing his collection. Instead, he stuffs his findings in whatever receptacles he can get his hands on and then piles the receptacles in the cupboards in his pantry, never forgetting which ingredients are in which receptacles and where they are located in the pantry. While the potion-master is at work, he calls out identifying descriptions of the receptacles to his apprentice. "Bring me the glass vial on the third shelf with the blue liquid in it! Bring me the wooden crate in the second cabinet with the black beetles in it!" The apprentice dutifully runs to the pantry, retrieves the items that satisfy his master's descriptions, and delivers them.

My focus is on the apprentice's cognitive process that allows him to succeed at the task. First, he has an auditory experience as of the potion-master's vocalizations, and understands those vocalizations as linguistically encoding a certain content.¹⁸ Next, when he enters the pantry, he has visual experiences as of the cupboards and all they contain. Finally, he visually identifies those items that satisfy the linguistic content of his master's

¹⁷ By "epistemic explanation" I mean an explanation whose explananda invokes relations that are explicitly epistemic. A purely causal explanation is not an epistemic explanation in my sense (even if epistemic relations turn out to be special sorts of causal relations).

¹⁸ Here I remain neutral with respect to whether linguistic content is part of the apprentice's auditory experience, is a distinct but concomitant experience, or is inferred by him (however rapidly and automatically) on the basis of his auditory experience.

vocalizations. That is: when he finds the item his master sent him into the pantry to find, he reidentifies the intentional contents of his visual experience as of the same sort as the intentional contents of his auditory experience or, at any rate, of his conscious understanding of the linguistic content that experience. He is able to tell that one of his experiential states shares intentional properties with another of his experiential states: he can *cross-modally recognitionally sort* conscious mental states according to their intentional properties.^{19,20}

Why think that the right way to characterize the apprentice's success is in terms of his being introspectively acquainted with intentional properties? The basic thought here is that if the apprentice's recognition judgments are in any sense *guided*, then he must be acquainted with the items mentioned in those judgments, because only states with which he is acquainted could do the guiding. Now, it is not in general the case that if a subject can sort two items X and Y for sameness or difference, then the subject is introspectively acquainted with X and Y: after all, we have the capacity to *perceptually* recognize an item as the same or different from another item, and we are certainly not introspectively acquainted with the items we routinely perceptually recognize! I think, however, that any epistemic explanation of such capacities will either appeal to acquaintance at some level, or else will dispense with epistemic explanation altogether. Consider, by way of

¹⁹ Note: it is consistent with my way of telling the story that (a) the apprentice's understanding of the master's instructions and (b) his visual experiences within the pantry have different types of content: say, conceptual vs. nonconceptual. (Indeed, I'm committed to their having different types of content, inasmuch as content is typed in terms of experiential mode—auditory vs. visual, for example.) What matters for my purposes is that the two types of states have intentional properties in common, and that the apprentice can recognize the commonality.

²⁰ "The apprentice doesn't compare his visual experience to the Master's descriptions; he compares *the items in the pantry* to the Master's descriptions." He does so, yes, *by way of* recognitionally sorting his experiences. (Compare: he *sees* the items in the pantry by *visually representing* them.) Here are two reasons to think the two epistemic tasks (sorting experiences, sorting experienced items) are not in competition. (1) Phenomenal experiences present worldly items, but also present themselves. (For interesting discussion, see Nida-Rumelin (2011) and Siewert (2012).) (2) Were the apprentice hallucinating the whole affair, he could still perform an epistemic task of the same type.

illustration, our ability to perceptually recognize faces. For most of us, facial recognition is totally automatic—it is as obvious to us *whose* (familiar) face we are looking at as it is *that we are looking at a face*, and if we were pressed to justify our recognitional judgment, we might have nothing to say. Our capacity for facial recognition need not be *guided*. For prosopagnosics, by contrast, facial recognition is an epistemically substantive capacity. Their perceptual experiences of faces do not, by themselves, represent *whose* face perceptually appears. When a prosopagnosic recognizes a face, she does so by inference from the features that *are* represented in perception. I contend that she will need to be introspectively acquainted with such perceptually represented features in order for her inference to be guided by them. So even in the case of perceptual recognition, either epistemic explanations are inappropriate, or if they are, acquaintance will figure centrally in them.

Cross-modal recognitional sorting with respect to content *is* an epistemically substantive capacity, I will now argue. I do so by looking at two alternative conceptions of our semantic self-knowledge. According to the first, the “redeployment strategy”, semantic self-knowledge is a trivial form of propositional knowledge. According to the second, the “reidentification strategy”, semantic self-knowledge is a brute form of ability-knowledge. Neither can account for our ability to cross-modally recognitionally sort.

4. The Redeployment Strategy

The literature in the philosophy of language has included extensive debate as to whether semantic externalism is consistent with self-knowledge (e.g. knowledge of the meanings

of our words).²¹ According to one influential defense, suggested by John Heil (1988) and Tyler Burge (1988) and developed more fully by Christopher Peacocke (1996 and 1999, ch. 5), knowing the meanings of my words is a matter of believing certain things about them and about myself. These beliefs will “re-deploy” the very same terms whose meanings are in question. Here is a simple example: suppose that environmental factors (at least partly) fix the content of my word ‘water’. Do I know what ‘water’ means? Sure, I do: I know that ‘water’ refers to water—and the very same environmental factors that fix the meaning of ‘water’ in my everyday usage of it fix its meaning when I reflect on the word itself. I can go on to say more about what I mean by ‘water’ and what I believe about water, all the while depending on environmental factors outside of my ken as the partial determinants of what my words refer to.

Heil, Burge and Peacocke have understood the redeployment strategy to apply to mental contents as well as language. So let us see if the strategy can serve present purposes, viz. whether one’s *beliefs about* one’s intentional mental states can make sense of one’s ability to recognitionally sort them. When the potion-master’s apprentice recognizes his visual experience as having the same content as his previous auditory-linguistic experience, what does the redeployment strategy have to say by way of explanation? Something like the following: the apprentice (a) forms the belief that he has heard his Master say thus-and-so and the belief that he is visually experiencing in such-and-such a way; and then he (b) recognizes that the beliefs have something in common—

²¹ Two of the most influential papers on the side of the *inconsistency* of self-knowledge and semantic externalism come from Paul Boghossian (1989) and William McKinsey (1991). On the other end of the spectrum, Gary Ebbs has argued that if semantic externalism is true, skeptical worries about self-knowledge cannot be so much as intelligibly articulated. See Brueckner and Ebbs (2011).

i.e. that thus-and-so = such-and-such. But it's murky how either one of these steps is supposed to occur.

What is it for the apprentice (a) to form the requisite beliefs? It's tricky to see what *redeployment* could come to in the perceptual examples we're looking at. Notice a contrast with language: redeploying a word is a matter of using a new token of the same word-type. But what is it to redeploy a new token of the same *experience*-type? I suspect that an advocate of the redeployment strategy has two options. When the apprentice believes that he is visually experiencing in such-and-such a way, 'such-and-such' will amount to a demonstrative, filled in either by his actual visual experience (the selfsame token), or else an *imaginative recreation* of that visual experience (a new token of—sort of—the same type).

Consider an analogy with monetary currency. Coins and bills (and seashells and saltcakes and whatever else) have both intrinsic and relational properties, and their monetary value is of course among the latter. Suppose you put a coin in front of me and ask me to how much it is worth. (Let's suppose there isn't any writing on the coin that I can understand.) I tell you that it is worth *this much*, while nudging the coin forward. I think it should be clear how insubstantial my belief is. And that is as robust a belief as the redeployment strategist can reach for in accounting for the apprentice's abilities.

But now, if it is beliefs of this flavor that the apprentice forms about his two experiences, how is he supposed to (b) recognize that the beliefs have something in common? If you set a second coin in front of me, I can form a similar belief about its value. But I am not in a position to tell you, on the basis of the intrinsic²² features of the two coins, whether the two coins have the same denomination. The trouble here is that

²² Not so much intrinsic as readily perceptually available, but close enough.

the apprentice's two mental states—his auditory-linguistic experience and his visual one—appear to have nothing in common *save* their intentional properties. If intentional properties are relational, then there is nothing on the basis of which sameness of intentional properties could be recognized, from the first-person perspective.

It seems, then, that the redeployment strategist is committed to there being some intrinsic property in common between mental states that have the same intentional content. But I do not see how this move is available to redeployment strategist. Here is the trouble: *there is no guarantee that distinct content-bearing structures and distinct intentional contents are one-to-one*. Just as two words can mean the same, or two coins can denominate the same, so distinct content-bearing structures can instantiate the same content-fixing functional/dispositional profile. If S_1 is part of a mental state, and S_1 has the right functional/dispositional profile, then that state has the content F. But if S_2 , a different structure, has the same functional/dispositional profile, then the presence of S_2 in a state makes it the case that that state also has content F. And that means that even if S *could be* part of all of one's mental states (in any mode) that have the content F (which, as I say, is already an odd suggestion: it does not *seem* that there are intrinsic similarities between the apprentices' two experiential states distinct from their shared intentional properties), there is no guarantee that S *will be* part of all of one's mental states that have the content F.²³

²³ Further, nothing so far guarantees that a type of content-bearing structure could only admit of *one* functional/dispositional profile. Type-identical intrinsic properties could herald the presence of distinct intentional properties in distinct mental states. Just as one word can mean distinct things in distinct contexts, or one coin could denominate differently in differently social groups, so two tokens of the same structure-type could instantiate distinct content-fixing functional/dispositional profiles. It seems to me, however, that it is a constraint on naturalistic theories of intentionality that they rule out this possibility. (Not that I see *how* to rule out the possibility.)

Perhaps the redeployment strategist will say something like the following: there is an as-of-yet undiscovered a priori reason that content-bearing structures and intentional contents *must be* one-to-one, and hence that the apprentice is rationally justified in treating sameness of structure as an indication of sameness of intentional type. But there could not be such a reason, because it is easy to imagine systems in which structure and intentional content are not one-to-one. For example, we could set up a system such that, whenever a content-fixing state involving S_1 is tokened, a second content-fixing state involving S_2 is tokened in parallel. Hence S_1 and S_2 have the same the content-fixing functional/dispositional profile, and so any mental state that has *either one* as a part thereby has the very same content.

It seems that the best a redeployment strategist can say is that there is an as-of-yet undiscovered *empirical* reason that structures and intentional contents *happen to be* one-to-one, and hence that the apprentice is reliabilistically justified in treating sameness of activation-pattern as an indication of sameness of intentional type. (Maybe cognitive systems such as ourselves are built so that only one content-fixing state can be tokened at a time.) What I want to point out is that we have pretty much abandoned the redeployment strategy at this point. Our question was: what accounts for the apprentice's ability to recognitionally sort his mental states according to intentional type? The redeployment strategist invoked the apprentice's same-contented beliefs about his two mental states. But we observed that the beliefs in question actually just demonstrate the mental states themselves; so the question remains how the apprentice is able to tell that the demonstrated states have the same content. The best answer so far is that the two mental states share a content-bearing structure, and that *it so happens that* sameness of

structure is a reliable indicator of sameness of intentional type. But the apprentice's *beliefs* about the presence of these structures are irrelevant to his recognitional ability. Furthermore, his transition from recognition of sameness of structure to sameness of intentional content is not a *rational* transition (since the one-to-one pairing of activation-pattern with intentional content is at best a contingent matter of fact). Far from semantic self-knowledge amounting to trivial propositional knowledge, such knowledge turns out (at least when it comes to knowledge about sameness and difference across mental states) to require primitive recognitional capacities whose reliability cannot be established a priori. In short: the sense in which the apprentice is "redeploying" anything is immaterial to his ability to recognize sameness and difference of intentional type, and that ability has to be epistemically brute, built into the wiring of his cognitive apparatus. The redeployment strategy has collapsed into an appeal to epistemically brute recognitional abilities. So let us look head-on at a strategy built around such an appeal.

5. The Reidentification Strategy

According to the sub-argument for premise 1, the epistemic relation that a subject bears to her intentional mental states can explain her ability to recognitionally sort them according to intentional type. It is her knowledge of them—an instance of knowing-*what*—that provides epistemic grounding for her cognitive abilities with respect to them. The advocate of the redeployment strategy rejects this: she says instead that it is a subject's beliefs about her intentional mental states (i.e. an instance of knowing-*that*) that trivially account for her cognitive abilities with respect to them. But, as we have seen, the redeployment strategy must rely on primitive epistemic abilities (and does not rely much

on the propositional knowledge it invokes). A second alternative appeals to primitive epistemic abilities quite self-consciously. According to this view, neither knowings-what nor knowings-that explain introspective recognitional abilities. Rather, semantic self-knowledge *just is* a species of ability, a knowing-*how*. Here is Millikan's statement of a view of the sort I'm describing:

The closest thing to the yearned-for ideal that actually makes some sense, I suggest, is a confrontation of one thought of an object with another thought of that same object, taking place within thought itself, and effecting a recognition of the sameness of the object.... Knowing what one is thinking of is, just, having the capacity to recognize when two of one's thought tokens are thoughts of the same.²⁴

This view, which I'll call "the reidentification strategy," strikes me as phenomenologically inadequate. Here's why: my conscious attention to my intentional mental states is clearly epistemically relevant to my ability to recognitionally sort them. It is not an accident that we attend to our mental states in making these judgments. Similarly for the apprentice: if he is worried about making a mistake, he is not at a loss; he need not just shrug his shoulders and say, "My recognitional abilities are brute; I either have them or I do not"; he can attend extra carefully to his memory of what his Master said and to his visual experience.

The point is closely related to what John Campbell has observed with respect to the cognitive role of phenomenology in his discussion of perceptual demonstratives:

Suppose you say to me, 'What is that mountain over there?' To understand your question I have to know which mountain you are taking about... You might

²⁴ Millikan (1993), p. 96-97.

acknowledge that ordinarily we would use visual information to interpret the [visual] demonstrative, but question whether it has to be conscious. The idea of visual information that is not conscious is made vivid by cases of blindsight.... The issue is whether the blindseer has the very same way of interpreting the demonstrative as the ordinary subject has. That is, the question is whether for the ordinary subject, consciousness of the object is not completely idle in an understanding of the demonstrative.²⁵

The point is that if a blindseer could have the same purchase on the referent of a visual demonstrative as a normally sighted person, then visual phenomenology is explanatorily idle. But it is not: as Campbell goes on to say, “knowledge of the referent of a demonstrative comes via conscious attention to an object.” My perceptual phenomenology *does explain* my ability to think about what I have seen; the comparison with blindsight serves to bring out the fact.²⁶

Suppose that the potion-master’s apprentice were blindsighted but nevertheless reliable in his retrieval-duties. He hears and understands his master, enters the pantry with his eyes open, and successfully reaches for the right object. Is his reliability to be explained in exactly the same way as the reliability of a normally-sighted apprentice? Clearly not: the normally-sighted apprentice is guided by his attentiveness to when his visual-phenomenal state has a content that satisfies the description given by his master. But if the reidentification strategy were on target, the apprentice’s phenomenality would make no such explanatory contribution.

²⁵ Campbell (2002), pp. 7-8.

²⁶ See Smithies (2011) for an attempt to account for the modal force of Campbell’s observation. Smithies argues that not only does conscious attention *actually* play the role of delivering knowledge of the referent of a demonstrative; it is the *only possible* medium for such knowledge. The weaker, actualist observation made by Campbell is sufficient for my purposes, however.

The reidentification strategy does not accord with our experience of forming comparative introspective judgments, then. But phenomenological inadequacy is not the only source of trouble for the reidentification strategy; there are third-personal, theoretical reasons to reject it as well. While I have been arguing that the epistemic relation that a subject bears to her intentional mental states can serve to explain her recognitional abilities with respect to them, these are by no means the *only* abilities that are naturally explained by knowledge of the content of one's mental states. Here are some other abilities that are epistemically explained by my knowing the content of my mental states:

- the ability to knowledgeably judge that one has recognitionally sorted one's intentional mental states
- the ability to bring to mind mental states with the same content as previous mental states
- the ability to express the content of a mental state in words (e.g. describing what I want)
- the ability to express the content of a mental state in images (e.g. drawing what I see)

I now propose a dilemma for the reidentification strategist. On the one hand, she can alter her initial identification. Rather than saying that semantic self-knowledge is identical to having the single ability to recognize states with the same content, she can say that semantic self-knowledge is identical to having a *conjunction* of abilities, including those just listed. But this is an awkward adjustment. The list is disunified; nothing accounts for our including these abilities rather than others—besides the fact that we would intuitively

take these abilities to all admit of the same explanation (viz. that they are rooted in one's semantic self-knowledge). And motley lists make for highly inelegant explanations. Suppose we observe that several symptoms tend to cluster together—nausea, dizziness, body aches, etc. Calling the symptom-cluster a “syndrome” is a way of naming a phenomenon that has yet to be explained; inquiry only ceases once we have discovered the underlying disease that explains the presence of all of them together. The list of capacities the reidentification strategist would have to appeal to is like this—a syndrome in need of explanation. The list is precisely the explanandum, not the explanans.

On the other hand, the reidentification strategist can retain her initial identification (viz., the identification of semantic self-knowledge with the ability to recognitionally sort mental states with respect to content), and then claim that *this* ability explains all the others on the list. While strategically better, the challenge is still substantial. It is not apparent that the abilities on the list can be explained in terms of a single recognitional ability. Then again, it is not apparent that they cannot.

But there is more. According to a venerable rationalist doctrine, consciously grasping certain propositions is sufficient for knowing them, or at any rate for justifiedly believing them. Now, plenty of applications of this rationalist principle have turned out to be false. Nevertheless, it's plausible that in some cases, an adequate grasp of one's mental states does put one in a position to have justified beliefs of a certain sort—a sort we might call *conceptual* knowledge. For example, a subject's knowledge of the content of relevant mental states is sufficient to justify the following:

- the belief that twice two is four (*justified by knowledge of the content of numerically-contented mental states*)

- the belief that round things are shaped (*justified by knowledge of the content of geometrically-contented mental states*)
- the belief that poems are more like prayers than like essays (*justified by knowledge of the content of literarily-contented mental states*)

Note that if a subject were disposed to deny any of these propositions, we would consider her not just wrong but confused: we would suspect an inadequate grasp on her part of the concepts contained in them.

If conceptual knowledge as I have construed it is possible, it renders both horns of the dilemma more troublesome for the reidentification strategist. On the first horn, she adds such beliefs to a list of items (a subject's possession of which items counts as semantic self-knowledge). But now the motley nature of the list has gotten truly out of hand: it consists of an assortment of mutually irreducible abilities *and* an assortment of mutually irreducible beliefs. The list contains elements in two different epistemic categories. Not only is there no apparent reason that the particular items are to be found on the list rather than others, there is no apparent reason why such items fall under two categories, rather than just one or several.

On the second horn, the reidentification strategist identifies semantic self-knowledge with a single recognitional ability, and then treats the abilities *and beliefs* on the list as epistemically grounded in that single recognitional ability. But it's very difficult to see how conceptual knowledge could be epistemically grounded in recognitional abilities. Supposing I am able to tell when identity-relations hold between the content of my mental states, this does not mean that I am able to tell when *other*

relations hold between the contents of my mental states, relations such as determinate-determinable (as holds between roundness and shapehood).

Now, conceptual knowledge is controversial, and I haven't supplied much by way of argument in defense of my assertions about it. Specifically, I have claimed (a) that we have conceptual knowledge and (b) that our semantic self-knowledge epistemically explains it. In order to *demonstrate* the theoretical inadequacy of the reidentification strategy, I would need to defend these claims, and so doing would extend beyond the scope of the present inquiry. Consequently, the advocate of the reidentification strategy is dialectically free to deny that we have any conceptual knowledge (maybe all beliefs are empirically justified if justified at all). But even if she were to dodge the argument from conceptual knowledge, her view would remain theoretical vicious—it would remain a list in lieu of an explanation—in addition to remaining at odds with our experience as introspective knowers.

6. Conclusion

In his 1999 book *Being Known*, Peacocke writes the following:

Many of the philosophical problems involving the first person and the self are, in one guise or another, instances of the challenge of ... reconciling two apparent truths. The first apparent truth, an epistemological truth, is that thinkers are able to know the intentional contents of many of their own attitudes without first checking on their environmental relations. The other apparent truth, a truth of the metaphysics of mental states with intentional content, is that for a wide range of intentional contents, a thinker is, as a constitutive matter, able to have attitudes

with those contents only if he stands in certain relations to an environment of a certain kind.

In this area, the option of revising the metaphysics is unattractive.

...Anyone moved by these considerations will look for a reconciliation of our two apparent truths by reconceiving the epistemology of self-knowledge.²⁷

I agree with Peacocke's description of the philosophical situation but disagree with his reconciliation-strategy. The problem is not that I *prefer* an acquaintance-based theory of self-knowledge to a belief-based or ability-based theory. The problem is that those theories are inadequate. They must either ignore certain sectors of our knowledge (i.e. those that are plausibly explained by our knowledge of the content of our mental states, such as conceptual truths), or treat as brute that which needs explaining.

I conclude that paradigmatic intentional properties—the sorts of intentional properties that are purely qualitative, potentially conscious and subjectively presentable—are intrinsic features of the mental states in which they are instantiated. It follows that naturalistic theories of intentional properties cannot provide a complete metaphysics of intentionality. If this is right, it means that intentionality cannot be fully located within the ontology of the natural sciences, as that ontology is currently understood.

A final observation: I do not intend simply to dismiss the program of naturalized intentionality. True, the naturalizing project fails according to its stated aims. But it might not fail according to more modest aims: perhaps one or another of the four naturalistic theories has successfully landed on *sufficient conditions* for a mental state's having a certain content. (The mistake lies in proceeding to *identify* these conditions with having a

²⁷ Peacocke (1999), p. 203.

certain content.) Or, there might be scientifically useful generalizations to be made about the properties that the four naturalistic theories identify. These properties are not the ones out of which *our* conscious, cognitive lives are built. Nevertheless it might be accurate to characterize such properties as intentional after a fashion—as the building blocks for a scientifically interesting phenomenon that we might call “proto-cognition.”

References

- Block, Ned. (1986). "Advertisement for a Semantics for Psychology". *Midwest Studies in Philosophy* 10: 615-78.
- Bickhard, Mark. (2004). "The Dynamic Emergence of Representation". In Hugh Clapin (ed.), *Representation in Mind*. Elsevier. 71-90.
- (ms). "Interactivism: A Manifesto."
- Boghossian, Paul. (1989). "Content and Self-Knowledge". *Philosophical Topics* 17: 5-26.
- Brueckner, Anthony & Ebbs, Gary (2012). *Debating Self-Knowledge*. Cambridge University Press.
- Burge, Tyler (1988). "Individualism and Self-Knowledge". *Journal of Philosophy* 85: 649-63.
- Campbell, John (2002). *Reference and Consciousness*. Oxford University Press.
- Dretske, Fred. (1981). *Knowledge and the Flow of Information*. MIT Press.
- (1986). "Misrepresentation." In R. Bogdan (ed.), *Belief: Form, Content, and Function*. Oxford University Press. 17-36.
- (1988). *Explaining Behavior: Reasons in a World of Causes*. MIT Press.
- (1995). *Naturalizing the Mind*. MIT Press.
- (1996). "Phenomenal Externalism: Or, If Meanings Ain't in the Head, Where Are Qualia?" *Philosophical Issue* 7: 143-58.
- Fodor, Jerry (1990). *A Theory of Content and Other Essays*. MIT Press.
- Harman, Gilbert (1982). "Conceptual Role Semantics". *Notre Dame Journal of Formal Logic* 28: 242-56.

- Heil, John. (1988). "Privileged Access". *Mind* 98: 238-51.
- Hill (ms). "How Concepts Hook Onto the World".
- McGinn, Colin. (2005). *Mindsight: Image, Dream, Meaning*. Harvard University Press.
- McKinsey, Michael (1991). "Anti-Individualism and Privileged Access". *Analysis* 51: 9-16.
- Millikan, Ruth. G. (1984). *Language, Thought and Other Biological Categories*. MIT Press.
- (1993). "Knowing What I'm Thinking Of". *Proceedings of the Aristotelian Society* 67: 91-108.
- (2009). "Biosemantics". In Brian P. McLaughlin & Ansgar Beckerman (eds.), *The Oxford Handbook of Philosophy of Mind*. Oxford University Press. 281-97.
- Nida-Rümelin, Martine (2011). "Phenomenal Presence and Perceptual Awareness: A Subjectivist Account of Perceptual Openness to the World". *Philosophical Issues* 21: 352-83.
- Peacocke, Christopher. (1996). "Entitlement, Self-Knowledge, and Conceptual Redeployment". *Proceedings of the Aristotelian Society* 96:117-58.
- . (1999). *Being Known*. Oxford University Press.
- Siewert, Charles (2012). "On the Phenomenology of Introspection". In Declan Smithies & Daniel Stoljar (eds.), *Introspection and Consciousness*. Oxford University Press. 129-168.
- Smithies, Declan (2011). "What is the Role of Consciousness in Demonstrative Thought?" *Journal of Philosophy* 108: 5-34.