Modality, Normativity, and Intentionality

A striking feature of the contemporary philosophical scene is the flourishing of a number of research programs aimed in one way or another at making intentional soup out of nonintentional bones[1]—more carefully, specifying in a resolutely nonintentional, nonsemantic vocabulary, sufficient conditions for states of an organism or other system to qualify as contentful representations. This is a movement with a number of players, but for my purposes here, the work of Dretske, Fodor, and Millikan can serve as paradigms. The enterprise in which they are jointly engaged is not so much one of conceptual analysis as it has been traditionally understood as one of conceptual engineering. That is, instead of thinking about what ordinary people or even sophisticated philosophers already mean by terms such as 'representation', they appeal to the tools of the special sciences (for instance, information theory and evolutionary biology) to describe abstractly, but in criticizable detail, how one might craft a situation in which some state arguably deserves to be characterized as 'representationally contentful' in various important senses. Insofar as the theories are good ones, they may shed light on how human knowers actually work. But their immediate aim is a broader one: to say what would count as doing the trick, rather than how we manage to do it.

---

[1]   Steven Turner's adaptation of Dretske's phrase. [ref.]

It is not hard to get into a position from which intentionality[2] can seem philosophically puzzling. One need not, for instance, be an atoms-in-the-void physicalist in order to want to understand better the nature of the transition from a world devoid of semantic properties to one in which they are at least locally rife. One way of dispelling the sense of mystery that threatens to surround such a transition (indeed, where such a story is possible, the very best way to do that) is to explain in terms applicable in principle already to phenomena before the transition conditions under which the new vocabulary would be applicable.

This is an exciting and invigorating sort of project—not least for the prospects it holds out of doing careful, detailed, ground-level work that contributes directly to the clarification of fundamental, large scale, longstanding philosophical problems. The goal of this collective undertaking is typically taken to be the production of a *naturalistic semantics*: showing how intentionality can be made intelligible by deploying the concepts and methods characteristic of the natural sciences, exhibiting semantics as a special natural science.[3] Certainly that is the way Dretske, Fodor, and Millikan, for instance, think of it, as is evident not only from the motivations they avow, but equally from the raw materials they permit themselves to employ and the criteria of adequacy to which they submit the resulting accounts. The project of naturalizing the mind and cognition

---

[2] In this essay, I'll speak very roughly, and use the notions of <u>intentionality</u>, <u>representation</u>, <u>concept-use</u>, and <u>semantic content</u>, interchangeably—not because I believe that there are no important distinctions to be registered by the use of these different locutions, but because those differences don't matter at the level of abstraction at which I am telling my story.

[3] Hartry Field presents an interesting account of one way into this project, early in his classic essay "Tarski's Theory of Truth" (*Journal of Philosophy* 69:347-375). The difference between the sorts of conceptual tools that Dretske, Fodor, and Millikan, each in his or her own way, propose to deploy to

has a long philosophical history. But it is possible to see these projects as episodes in a

still broader movement of thought, in which naturalism is only one strand—indeed, an

optional one. What I want to do in this essay is to sketch that broader philosophical

tradition, and to locate some of these contemporary semantic projects within it.

I

It is helpful to begin with to focus on two phases in the development of contemporary

philosophical thought about intentionality: an early one and a later one. (Though I don't

mean to pretend that nothing else of significance for this tradition was going on. I'll return to some other

strands in the debate later on.) One feature of intentionality—the of-ness, about-ness, or more

generally the representational contentfulness characteristic of thought and talk—that first

caught semantic theorists' attention is the non-extensionality (inten_s_ionality or

specification relativity) of the locutions used to ascribe intentional states. Already in

"Über Sinn und Bedeutung," Frege had pointed out the failures of substitutivity within

propositional-attitude ascribing contexts that disqualify them as extensional. Whereas in

extensional contexts, substituting one coreferential term for another (or one co-truth-

valued sentence for another) preserves the truth of the whole sentence in which they are

embedded, such substitutions can alter the truth values of sentences of the form "S

believes (or thinks) that $\phi(t)$." Again, one can believe that $\phi(t)$, for instance that the

present king of France is bald, even if t (that king) does not exist. But one cannot stand in

a nonintentional relation—say, kicking—to what does not exist. And one can want to buy

---

address this common problem, on the one hand, and those Field appealed to, on the other, I think mark an

a sloop, even though there is no particular sloop one wants to buy—even though one wants only, as Quine puts it, relief from slooplessness. But one can only sail or sleep on some particular sloop.[4]

 By the 1950's and early '60s, a broadly Tarskian, model-theoretic approach to the semantics of extensional locutions (descended from Frege's work) was widely agreed to provide the best available paradigm of a successful formal semantic theory. Frege's own approach to the semantics of intensional contexts, such as the propositional attitude ascriptions by which intentional states are attributed, was by contrast both not adequately formalized and much less widely accepted philosophically. The result was a distinctive philosophical problematic, centered on the inten_sionality (non-extensionality) of inten_tionality.

One influential response, of course, was Quine's philosophical rejection of the ultimate intelligibility of the notion of intentional state, motivated by the recalcitrance to extensional semantic analysis of the locutions by which they are attributed. Others looked for conceptual raw materials that would support a more positive response. By far the most popular candidate seized upon was another sort of inten_sional idiom that was being philosophically domesticated at around the same time (from the mid '60s through the '70s): alethic modalities. Thus was born the fruitful and still ongoing project of

---

important dimension of change in the analytic tradition over the twenty five years spanning these projects.
[4] Significantly, from the point of view of the story to be told here, many of these intensional features of intentional idioms are shared by *normative* phenomena. Thus one can promise to give someone a blue ox, even though there is no blue ox, and owe someone an ox, though there is no ox in particular one owes (one is obliged only to relieve his oxlessness).

understanding the intentionality characteristic of mind and language by deploying a

naturalistic but modally rich vocabulary centering on the sorts of *counterfactuals* used to

codify *causal* relationships. This explanatory project forms one of the central strands of

the story I want to tell here.

The second episode in the development of contemporary philosophical thought about

intentionality that I want to highlight is the heightened appreciation of the *normative*

character of meaning and concept use aroused in part by Kripke's discussion of

Wittgenstein, beginning in the early '80s. The core idea is that anything recognizable as

an intentional state (for present purposes, we can think of these as propositional

contentful states or as conceptually contentful representations) must underwrite normative

assessments as to whether things are as they *ought* to be, according to that state—whether

the state is *correct* or *successful* according to the standards determined by its content.

Believing includes *committing* oneself, undertaking a *responsibility* concerning how

things are (how they might be found to be). Intending also includes committing oneself,

undertaking a responsibility concerning how things will be (how they must be made to

be). Beliefs are essentially, and not just accidentally, things appropriately assessed as to

their correctness in the sense of their truth. And desires and intentions are essentially, and

not just accidentally, things appropriately assessed as to their success, in the sense of their

fulfillment. (Even less overtly committive intentional states such as conjectures, hopes,

and wishes are contentful only insofar as they settle how things ought, according to them,

to be.) Similarly, speech acts such as claiming and commanding essentially involve

adopting normative statuses, including as they do the undertaking of responsibility (for

how things are) and the assertion of authority (over how things are to be). Using a term with a determinate meaning (using it so as to express a particular concept) is binding oneself to a norm that determines the correctness or incorrectness of that use (along with that of many other possible uses). It is this normative dimension of intentional content that makes it possible to distinguish two complementary "directions of fit" that beliefs and intentions, or declaratives and imperatives can have, depending (intuitively) on where the fault is taken to lie if things are not as they are supposed to be according to the norm articulated by their contents.

For what an organism is doing to be intelligible as representing, there must be room also for *mis*representing, for representation that is *in*correct. One of the hallmarks of the normativity of intentionality is that what one commits oneself to in applying a concept outruns in principle what one *takes* oneself to be committed to. The norm of correctness one thereby binds oneself by goes beyond both the dispositions of those undertaking those commitments and what they consciously envisage themselves as committing themselves to thereby. Because it does, a question arises about how to understand the features of the intentional state or meaningful utterance that settle which determinate conceptual norm one is bound by—exactly which standards for the assessment of correctness or success one has implicitly put in play—by being in that state or producing that utterance. If what one is committed to is not settled by what one consciously envisages (because one never so envisages enough), nor by what one is disposed to accept as such (because one can be wrong), how is it settled?

These observations collectively give rise to another distinctive philosophical problematic: offering an account of the distinctive normative character of intentionality. Combining this new challenge with the expressive resources first called into play in response to the recognition of the intensionality of representational content—namely the vocabulary of

3/16/1999--6

alethic modalities, of possibility and necessity, of subjunctive or counterfactual conditionals—yields a satisfyingly detailed research program with reasonably clear criteria of adequacy. It is the project of explaining conceptual normativity in terms of counterfactual dependencies, paradigmatically of representing events and states of affairs on represented ones. This is the project Dretske and Fodor have, each in his own way, pursued in a number of careful and justly influential works produced over the past twenty years.[5]

[Lecture: Disjunctivitis, the problem of leaving room for *error*, shows how work is needed to apply modal conceptions to this problem.

The *prima facie* difficulty that its normativity raises for attempts to understand intentionality directly in modal terms, by appealing to counterfactually robust correlations, appears in a number of different, but related guises. Crudely put, one cannot take what is represented by a state or performance to consist simply in whatever stimuli the system in question is disposed to respond to by entering that state or producing that performance. For that would leave no room for mistaken responses, for misrepresentation. Whatever the user of a concept takes to be correct would count as correct.[6] A more careful formulation would acknowledge that talk of dispositions and counterfactual dependencies makes sense only as applied to types of states and events, not for tokens. The shift to *types* of occurrence, understood as indicating other *types* of occurrence, in terms of the counterfactual reliability of the correlation holds out the possibility that individual token occurrences could fail to correlate, and so to count as mistaken, as *mis*representations. But

---

[5] It is *one* description of that project. I don't mean that they have typically thought of what they are doing in just these terms. I'm offering a *de re* specification of the content of their project, not a *de dicto* one.
[6] Indeed, I conjecture that one of the sources of Kripke's own interest in the topic of the normativity of intentionality (in the late 60's, when he wrote the prototype that was published only much later as *Wittgenstein on Rules and Private Language* [Cambridge: Harvard University Press, 1982]) was his realization that the modal apparatus that otherwise seemed so promising for dealing with the intensionality of intentionality raised exactly these problems.

even so, the possibility of gerrymandering the types involved presents a problem cognate to the original one. It is always possible in principle to draw boundaries around the stimulus types (perhaps seriously disjunctive ones) so that the correlation with the response type is *perfectly* reliable. Rather than misidentifying a particular echidna as a porcupine, one counts as correctly identifying it as a porcupine-or-echidna. Leaving room for the possibility of *error*, of *in*correct representation then requires somehow privileging some among all the possible types that each particular candidate representation-type correlates with. And that seems to require further conceptual resources.[7] Another guise in which appears what is, I think, recognizably a difficulty of the same general kind, concerns the conceptual resources required for picking out a *distal* stimulus kind as what is represented by a specified response kind, holding fixed all of the counterfactual dependencies. For given any reliably covarying causal chain of events leading from some stimulus to a response, later elements, more proximal to the eventual response, will always be more reliably correlated with the occurrence of a response of that kind than will earlier ones. Thus in the chain leading from the ringing of a bell through motions in the intervening air molecules, motions in the tympanic membranes of a dog, electrochemical changes in its auditory nerve, and so on ultimately to stimulation of its salivary glands, each intervening stage can also be brought about by some nonstandard causes, and so will less reliably elicit the response in question than its fellows downstream in the causal chain.

I am not claiming that the problem of misrepresentation is fatal to the enterprise. A number of strategies have been suggested for responding to the difficulties of reconstructing the normative dimension of concept use by means of modal concepts, and no doubt there will be more in the future. (One sociological reason for confidence that the enterprise will not soon be abandoned is the mutual conceptual and motivational support given to each other by taking subjunctive reliability of correlation of features of representings and representeds as the basis of a naturalistic semantics, on the one hand, and exploiting such reliable

---

[7] This is the point Dretske raises under the heading of "misrepresentation" (in his well-known article of the same name, pp. 17-36 in Radu Bogdan (ed.) *Belief: Form, Content, and Function*, Oxford University Press, 1986), and that Fodor discusses under the heading of the "disjunction problem" (*Psychosemantics* MIT Press, 1987, and the title essay of *A Theory of Content* MIT Press, 1990).

correlations as the basis of a naturalistic epistemology, on the other.[8]) The conclusion I am after is just that, while the use of modal vocabulary offers semantic theorists a direct purchase on the non-extensionality of intentionality, the use of that vocabulary does not in the same way serve *directly* to make its normativity intelligible. Another approach is to do just that: start with some familiar kind of normativity, and try to exhibit the normativity of intentionality as a species of it.

Some sorts of normativity do not provide promising raw materials for such an explanatory strategy. Soldier Schweik *ought* to march Northwest because his sergeant *commanded* him to do that. The painting *ought* to be hung this side up, because that is how the painter *intended* it. I *ought* to drive you to the airport tomorrow, because I *promised* to do it. These are all perfectly good senses of 'ought', so one might try out the idea of assimilating the sense in which a claim or belief ought to be true to one of them. But because commanding, intending, and promising are themselves all intentionally contentful states or acts, they already involve the kind of normativity characteristic of intentionality, and so cannot serve as non-question-begging unexplained explainers of it. (This is not to say that there is something wrong with explaining the normativity of speech acts by appeal to intentional states—as, say, Grice does—only that this is not a strategy for understanding the normative dimension of intentionality generally.)

A much more promising, and therefore more popular, strategy is to look to *teleology*, and understand the norms characteristic of intentional contentfulness as broadly *functional* norms. Some care is required with this strategy too. Appealing to norms of proper functioning of artifacts explained in terms of the intentions, beliefs, and purposes of their designers, will be unattractive in this methodological context because it raises circularity concerns of the sort just mentioned: it is hard to see how one could explain in the same way what counts as *properly* carrying out such an intention, what would make such a belief *correct* in the sense of true, how things *ought* to be according to the purpose in the sense of what state of affairs would fulfill it. On the other hand, we have learned that we need not think of *biological* functions—for instance, the function of the heart to pump blood—in these terms. Besides explanations of the proper functioning of

---

8   I express some doubts about the solidity of this support in "Insights and Blindspots of Reliabilism,"

subsystems of artifacts in terms of *design*, there are explanations of the proper functioning of subsystems of organisms in terms of *selection*. Understanding the normativity of intentionality on the model of norms of proper biological functioning has the structural virtue that understanding the modeling functional norms does not require even tacit appeal to intentional norms of the sort being modeled. Understanding the content of intentional states by likening the way a norm of truth is integral to beliefs (a norm of success integral to intentions, and so on) on the way a norm of blood-pumping is integral to hearts, involves no covert circularity. This thought generates another research program: teleosemantics, a kind of functionalism about intentionality that makes its normativity intelligible as an extension and elaboration of norms of proper biological functioning. Millikan's project is by far the most sophisticated and best worked-out exemplar of this kind of explanatory project.

Here, then, are two promising ideas: to understand the non-extensionality (inten<u>s</u>ionality) of inten<u>t</u>ionality by appealing to counterfactual dependencies of represent*ing* states of affairs on represent*ed* states of affairs, and to understand the normativity of intentionality by appealing to norms of proper functioning of represent*ing*s relating them to represent*ed*s, modeled on norms of proper functioning of reproducing biological systems that evolved by natural selection. In thinking about the relations between these ideas, it is important, I think, to keep in mind that the selectional explanations on which the second idea relies themselves depend essentially on the invocation of counterfactual dependencies that must be expressed in the sort of modally rich vocabulary on which the first idea relies. Maybe this is obvious. But even if it is, it is worthwhile to rehearse a couple of examples in order to make clear just what role the appeal to alethic modalities has in explanations of functional norms in selectional terms.

Dretske[9] talks about a species of Northern Hemisphere bacteria that contain magnetosomes, which reliably respond differentially to magnetic fields. The bacteria in turn reliably respond differentially to the orientation of their magnetosomes, and move toward geomagnetic north, which in that environment is also down, hence to water less rich in oxygen, which is toxic to these creatures. The state of these organisms' magnetosomes in this environment reliably counterfactually correlates with all of geomagnetic north, deeper water, and less oxygenated water. What should it be understood as representing? As Dretske points out:

> [T]his primitive sensory mechanism is…functioning perfectly well when under [a] bar
>
> magnet's influence, it leads its possessor into a toxic environment.[10]

Millikan counts it as a significant advantage of her selectional account over Dretske's that it privileges the latter of these in terms of its role in explanations of the evolution of the system.[11] For that responding in the way they do to the orientation of their magnetosomes Normally[12] leads organisms of this kind to less oxygen-rich water plays an essential role in evolutionary explanations of the persistence of this whole stimulus-response system.[13] And this essential role is articulated by counterfactuals concerning the development of the system: if it *had not* been the case that the mechanism (as we can now, retrospectively, say) *successfully* led the ancestors of contemporary bacteria to less toxic waters, then they *would not* now respond as they do to magnetosomes (or perhaps, have them at all). Teleosemantic theories of intentionality are a special kind of modal theory, distinguished by the particular form taken by their appeal to counterfactual dependencies expressed in a rich modal vocabulary.

---

[9]   in "Misrepresentation," op. cit. [6], and again in *Explaining Behavior* [M.I.T. Press 1988] p. 63.

[10]   in "Misrepresentation," p. 29.

[11]   Retailed, for instance in both "Biosemantics" and "Compare and Contrast Dretske, Fodor, and Millikan on Teleosemantics" (at pp. 93ff., and pp. 125 ff. respectively, in *White Queen Psychology and Other Essays for Alice* [M.I.T. Press 1993].

[12]   The capital letter indicates Millikan's technical sense: see LTBOC, e.g. p.96.

[13]   The difference that in this case makes a difference between Millikan and Dretske is that her account allows her to look downstream, to the response a system makes to a representation, and not just upstream, to its antecedents.

That this is so depends not on specific features of representational systems, but on the general nature of selectional explanations of function (and hence of functional norms). Consider (to use an example adapted from Sober[14]) a device that filters marbles depending on their size. Marbles of assorted sizes and colors are placed in the bin at the top, and a subset is selected by finding their way to the bin at the bottom. If size and color of the marbles are correlated, it may happen that all the marbles selected are red, as well as being small. The claim that it is size, rather than color that is selected *for*, even though both are *selected*, must be cashed out in terms of counterfactuals: *if* a marble *were* small and of some other color than red, the mechanism is such that it *would* be selected, while *if* it *were* large and red, it *would* not be. In the same way, Millikan can say that what the magnetosomes are *for*, their *function*, is to indicate gradients of oxygenation, rather than of pressure or magnetic field, so that they are functioning *properly* only when in fact they lead the organism to less toxic environs, because that is what was selected *for*. In fact *all* of Millikan's fundamental technical terms from the notion of "reproduction"[15] with which she begins, through the definitions of the various sorts of "direct proper functions,"[16] and of "Normal explanations"[17] are (and must be) specified in a modally rich vocabulary, in terms of laws and counterfactual conditionals.

In pointing out this feature of teleosemantic biofunctionalist approaches to intentionality, I by no means want to suggest that their conceptual reliance on modal notions is in any way a defect or a drawback. I point to it only to situate them in the context of counterfactual dependency theories of intentional content. This kind of functionalism about the normativity of intentionality should be seen as adding significant resources to modal approaches, rather than as simply a different kind of theory. It offers a *functional* account of the *normativity* of intentionality, but seeks to understand *functions* themselves in *modal* terms, explicating that normativity by means of

---

[14]   [ref.]
[15]   Her condition (3) is explained (*LTBOC*, p. 20): "Roughly, the law in situ implies that *had A been different…B would have differed accordingly*" [italics in original].
[16]   Beginning at *LTBOC* pp. 27ff..
[17]   Beginning at *LTBOC* pp. 33ff..

the sorts of counterfactuals first invoked in discussions of intentionality in response to the recognition of its intensionality.

It is against the background of this broad brush-stroke, highly selective rational reconstruction of some of the main strands in the tradition of recent thought about intentionality that

I want now first to consider exactly why it is appropriate to appeal to modal vocabulary in this explanatory context, and then to sketch the motivations for a somewhat different approach.

II

We can ask a very general programmatic or methodological question about the explanatory strategy of using modal notions to account for the non-extensionality and normativity of intentionality. Suppose we bracket the question of whether these conceptual raw materials are *sufficient* to account for the phenomena in question—even supplemented and articulated in the way the best teleosemantic accounts do. What makes us think that using these resources is so much as legitimate? The question becomes more pointed in historical perspective. For at least the first half of this century, philosophers (especially English speaking ones) were extremely suspicious of modal idioms. What such talk expressed seemed at best obscure, and at worst perhaps in principle unintelligible. During this period one of the outstanding philosophical problems and projects was precisely to render respectable the use of counterfactuals and appeals to what was possible or necessary. Conceptually fastidious philosophers in the broadly empiricist tradition felt themselves obliged to either to demonstrate the legitimacy of these notions

or to show how to do without them. Thus Peirce worried about what it meant to say of a diamond that spent its entire existence, from initial crystallization to final incineration, swaddled in cotton wool, that it was hard or fragile—that is that it *would have* scratched glass or shattered *had* it had a different career.[18] Such claims seemed to have a status wholly different from those attributing a certain shape, size, or velocity to the diamond. Modal vocabulary does not seem to be (and in fact is not) reducible to ordinary, nonmodal, descriptive vocabulary. We know (we think) what it is for the cat to be on the mat. But what is it for it to be *possible* for the cat *not* to be on the mat, but *necessary* that she is a mammal or that she is gravitationally attracted to the mat in inverse proportion to the square of the distance between them and in direct proportion to the product of their masses? What sort of properties, relations, or arrangements of objects constitute these facts, or the fact that *if* a mouse came within range she *would* leap on it? What are we committing ourselves to when we make this sort of modal claim, and what counts as sufficient evidence for them? The central role played by the notion of *habit* in pragmatist thought, as a natural phenomenon discernible in the inorganic world no less than the organic, but that could nonetheless be deployed to account for discursive and intentional phenomena, is in no small part a response to this issue. Humean concerns about the ultimate intelligibility of necessary connections, and so of the idea of causation itself are pervasive in the empiricist tradition, and are given added impetus by the development of extensional formal languages for the codification of empirical claims. First order quantificational languages could express *regularities* and *generalizations* with hitherto undreamed of precision. But for philosophers such as Russell and Carnap, that just made

---

[18] Sober, E. *The Nature of Selection*. Cambridge: MIT Press, 1984.

all the more urgent to explain or explain away *laws* and *lawlike generalizations*, whose

content somehow extended beyond what could be captured with these expressive

resources.

I think it is worthwhile reminding ourselves how surprised philosophers who lived and

moved and had their being in this milieu would have been to discover that by the end of

the century, when their successors found the intensional and normative character of

intentional idioms problematic, their first impulse and dominant strategy would be to

appeal to modal notions to explain them. Just how did that which seemed most in need of

philosophical explanation and defense become transformed so as to be unproblematically

available to explain other puzzling phenomena? The most cursory familiarity with

intellectual history accustoms us to the spectacle of explananda becoming explanans. But

in this particular case, what did we come to understand in the interim that made the

transition legitimate, rather than its being merely a product of change of fashion, fatigue,

or amnesia? Quine has kept the old questions alive. He finds the inten_s_ionality even of

modal vocabulary ultimately and in principle unintelligible, and has continued to argue

forcefully for the wisdom of eschewing all such nonextensional tropes. The world spirit

has not moved his way, but why not? Why are we *right* not to be bothered by his

scruples, hallowed as they are by tradition?

Here are three answers. I think the first two, which are probably the most popular

responses, will not do as they stand. They contain elements of a more adequate analysis,

but can be understood properly only in the context of the third, which I take to be the

deepest of the three. The first answer is that what happened is a revolution in modal logic and semantics. Here's a quick sketch. C.I. Lewis began a process that yielded formally adequate axiomatic ways of expressing various senses of 'necessary' and 'possible', and then Kripke developed a complete model theoretic semantics for those systems. With the extension of Kripke's semantics from modal logics to modal languages more generally, the process was complete. In the light of the tradition I have just been talking about, it has seemed particularly significant to many that Kripke showed us how to do semantics for modal languages using a first order *extensional* metalanguage.[19] Modal operators are interpreted by first order quantifiers over possible worlds. Given a structure of possible worlds and accessibility relations among them, all we need to do is to *describe* that structure in the ordinary, *nonmodal* language of first order logic, in order to understand the use of distinctively modal vocabulary. So here was an answer to the question, what sort of facts are modal facts? And the answer was extensional, using only ordinary descriptive vocabulary. Thus modal vocabulary became demystified and respectable. Now it no longer seemed to present a problem that laws and the dispositions expressed by subjunctive or counterfactual conditionals loom so large in the natural sciences. And for this reason modal vocabulary becomes available to philosophers to do explanatory work in its own right—not just as an *object* of explanation, but as a *tool* of explanation. So when certain potentially puzzling features of *semantic* vocabulary are noticed—for instance, that the vocabulary in which we describe inten*t*ionality is inten*s*ional—it is natural to appeal to the newly respectable modal vocabulary to explain them. This is one

---

[19] What has been called "California semantics" is second order (in both the version due to Kaplan and that due to Montague), in quantifying over functions, predicates, and so on. But in another sense it is still extensional, even in its treatment of intensions.

of the conceptual inheritances of our generation of philosophers: the shiny new toolkit of

modally rich philosophical metalanguages. So it has been one of our cardinal tasks to use

such metalanguages to reduce or explain what is expressed by specifically *semantic*

vocabulary: talk of what *represents* what, of what *expresses* what *content*, of what

something *means*.

There is obviously something importantly right about this line of thought, but it is

important to be careful about just what it is. It cannot be that, faced with the choice

between giving up the use of suspect modal locutions and showing them to be respectable

by reducing them to ordinary extensional nonmodal ones, we did the latter. The Kripke

semantics is not even a candidate for providing such a reduction, because it owes it's

extensional character to the introduction of new primitive notions, possible worlds and

accessibility relations (and in the case of quantified modal languages, further apparatus

permitting re-identification of individuals and across worlds) that are themselves richly

modal (and whose deployment turns out to require further metaphysically nontrivial

commitments concerning about what is essential and what accidental). Any probative

reasons to doubt the legitimacy of talk of necessity and possibility are just going to be

passed off and transformed into corresponding reasons to doubt the legitimacy of appeal

to such primitives. Thus an appeal to advances in formal semantics for modal logic is not

really responsive to the original conceptual challenge. The new semantics gives us much

greater control over our use of modal vocabulary, helping us to get clearer about what we

commit ourselves to by using it, how different modal notions related to one another and

to some nonmodal ones, articulating the fine structure of modal concepts. But this

clarification is resolutely internal to a language deploying the modal concepts, and does not address in any direct way more global worries about the legitimacy in principle of all such languages.

Here is a second answer to the question of what justifies the radical shift in attitude from treating modal notions as the paradigms of philosophically problematic and suspect concepts eminently in need of explanation to treating them as prime raw materials to employ in the explanation of other puzzling concepts. It, too, makes an important point, but one that is prone to misunderstanding. Its basis is the very observation that made the analysis and justification of modal concepts seem so urgent to the empiricist tradition in the first place: their ubiquity in scientific theory and practice. If the natural sciences generally help themselves generously to modal notions such as dispositions, if they distinguish what is possible from what is not, treat some but not all regularities as lawful, and endorse counterfactual conclusions, why should not a scientific semantics be entitled to do the same? What justifies putting more restrictive requirements on the scientific explanation of meaning than one puts for instance on the scientific explanation of the distribution of different sorts of animals, plants, or minerals over the earth? Once the demand that explanations in semantics be *a priori* and available from the armchair is recognized as an unwarranted relic of exploded philosophical prejudices, they can take their place alongside other scientific explanations, and help themselves to the same sorts of conceptual tools.

The trouble with this argument is that semantics is not just one more special science. It is (also) a philosophical discipline. Among its tasks is precisely the understanding and if need be criticism of the concepts employed by other disciplines. Part of its business is adjudicating disputes over the legitimacy of various sorts of attempts to mean something. It is accordingly methodologically required to be much more self-conscious and critical about the concepts it itself relies upon than are the disciplines whose concepts it studies. The geologist and the evolutionary biologist are entitled to ignore challenges to the ultimate coherence and legitimacy of talk of dispositions and laws, and just get on with the business of talking about the movements of tectonic plates and the adaptive advantages of homeothermy. The semantic theorist is in a different position, subject to different explanatory obligations, and cannot justify the same attitude simply by appeal to their example. At most the indispensability of modal locutions for the special sciences would show that there must *be* a philosophically satisfactory response to the challenge to their legitimacy forwarded by empiricism in both its traditional and its distinctively logical twentieth century forms. It does not by itself provide one.

Now, I think it is entirely legitimate for the contemporary semantic theorist to employ counterfactuals, appeal to dispositions, and in general to deploy a richly modal vocabulary in formulating theories about intentionality and meaning. But I also think it is important to understand why, and exactly what role the two sorts of consideration just sketched play in justifying our current level of comfort with the use of modal vocabulary. With the wisdom of hindsight, I think we should see that there were two different sorts of philosophical challenge regarding modalities. One of them was legitimate, but is

adequately responded to by the developments in formal logic and semantics of the last half century. Those developments do not provide a responsive answer to the other sort of challenge—but that challenge is illegitimate, for reasons that are gestured at, though not made explicit, by appeals to the indispensability of modal locutions in science.

 The empiricist tradition that stretches from Hume to Quine offers the friends of modality a stark choice: either show how to explain modalities in nonmodal terms or learn to live without them. But properly understanding the presuppositions of this challenge gives us good reasons to reject it. For it is predicated on the idea that the legitimacy of modal vocabulary turns on its reducibility to nonmodal vocabulary. And that presupposes a level of independently and antecedently intelligible discourse that is purely descriptive and nonmodal, as a background and model with respect to which the credentials of modal discourse can then be invidiously compared. An argument due to Kant and revivified by Sellars urges that this idea is a chimera. The ability to use ordinary empirical descriptive terms such as 'green', 'rigid', and 'mass' already presupposes grasp of the properties and relations made explicit by modal vocabulary. Sellars summed up the claim admirably in the title of one of his early papers: "Concepts as Involving Laws, and Inconceivable Without Them."[20] The incompatibility of the property of being green with that of being red, and so the *impossibility* of one color patch being at once green and red, is part of the content of the concept green. To call something 'rigid' is to commit oneself to various consequences concerning what *would* happen *if* one applied a potentially deforming force. That an object has a finite mass means that the application of force is *necessary* to

accelerate it. Apart from their involvements in material incompatibilities, counterfactual dependencies, and necessitations of the sorts these examples illustrate, bits of ordinary descriptive empirical vocabulary cannot have the meanings they do.[21]

Quine, consistently with his anti-modal stance, insists that a mature science must not appeal to 'dispositional properties'. He holds out as an example of how this revisionary project might proceed accounts of the solubility of sugar in terms of its microstructure. There might be some go to this thought if it is limited to overtly dispositional terms such as 'soluble' and 'rigid'. But if it is extended to include all terms whose use, like that of 'green' and 'mass', essentially, albeit implicitly, brings with it commitment to counterfactual consequences, then it is incoherent. For 'structural' appeals to properties such as valence and features such as molecular orbitals explain physical or chemical behavior only in virtue of the modally robust consequences they have, e.g. for what *would* happen if two molecules having a valence of +1 combine with one molecule having a valence of –2. One can perhaps have an empirical language in use that lacks the expressive power to make explicit modal relations among its terms and predicates. But one cannot have an empirical language the use of whose terms is intelligible apart from prohibitions and requirements concerning their relations which would be made explicit by the use of modal vocabulary. This is one of the central points Kant was arguing for in insisting that commitment to the applicability of the categories (the pure concepts of the understanding), which underwrite necessities, is implicit in the applicability of all empirical concepts. Bracketing the particulars of his account, it is an argument we should endorse today. The appeal to the fact that even the special sciences that are not much in the business of formulating necessary laws nonetheless routinely employ modally rich counterfactual-supporting idioms has force in justifying a similar policy for semantics, insofar as it does have force, only as backed up by an argument of this sort.

---

[20] Reprinted at pp. 87-124 in J. Sicha (ed.) *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars* [Ridgeview Publishing Company, Reseda CA, 1980]—hereafter *PPPW*.

[21]   Semantic theories of the sort I call " strongly inferentialist" (in *Making It Explicit* [Cambridge: Harvard University Press, 1994, pp. 131-2] treat this sort of relation as *sufficient* to determine the meaning of the broadly inferentially related terms. But the current, weaker, claim depends only on the *necessity* of standing in such relations for having the meanings they do.

Now I have only sketched the conclusion of this argument, not done what would be required to establish it. But if and insofar as there is a good argument that the applicability of ordinary empirical descriptive nonmodal vocabulary implicitly presupposes the applicability also of modal vocabulary such as 'possibility', 'necessity', and 'counterfactual dependence', the empiricist dichotomy—either reduce modal concepts to nonmodal ones or resolve to do without the modal ones—has no bite. For one can never be in the supposed position from which the challenge is issued. One can never be in a position in which the use of ordinary nonmodal terms such as 'green', 'rigid', and 'mass', 'cat', and 'mat', is in principle legitimate and unproblematic, while the use of modal terms is in principle questionable and problematic. For we cannot make sense of what is expressed by nonmodal vocabulary apart from all consideration of counterfactually robust consequences and implications. If that is so, if modality is already implicitly intricated in the very meanings of nonmodal terms, then the only legitimate challenge is to make sense of the whole complex idiom that includes both modal and nonmodal descriptive vocabulary. So the only remaining task is to be sure that we know what we mean by using the modal vocabulary, where that is something to be understood in terms of having inferential control over it in relation *both* to the basic descriptive vocabulary, *and* to other modal vocabulary. *That* is the enterprise within which I think we should understand the conceptual progress that has been made by the development of model theoretic semantics for modal formal languages. It contributes to our understanding of the function of modal vocabulary *within* the whole idiom that comprises both modal and nonmodal vocabulary. It helps articulate the relation between ordinary

descriptive predicates and explicitly modal ones, without pretending to reduce the content

of the latter to the former, or explain the one in terms of the other.

This is what I meant when I said that we should distinguish two different philosophical

challenges regarding modalities. The legitimate challenge, to clarify the expressive role

of modal vocabulary and establish adequate awareness of and control over the inferential

consequences we are committing ourselves to by its use, is adequately responded to by

the developments in formal logic and semantics of the last half century. Those

developments do not provide a responsive answer to the other sort of challenge—the

challenge to the in-principle legitimacy of modal discourse *tout court*. But that challenge

is itself illegitimate, because predicated on an ultimately untenable commitment to the

autonomous intelligibility of what is expressed by nonmodal discourse, with respect to

what is expressed by modal discourse.

III

I have indicated what I take to be the real reasons why it is entirely in order for semantic

theorists to appeal to alethic modalities, typically in the form of counterfactual

dependencies, in explaining semantic phenomena such as the intensionality or the

normativity of intentionality. But I think the story I have gestured at has further lessons

for the way we think about normativity, and hence about intentionality. For I believe we

should tell an analogue of that story, *mutatis mutandis*, with *normative* vocabulary

playing the role originally played in it by *modal* vocabulary. For we might ask why there

is a philosophical problem about normativity (whether in the context of thought about

intentionality or not), and what sort of problem it is. Guided by the previous discussion,

we can see that one sort of problem takes the form of a challenge to its legitimacy: either

explain the use of normative vocabulary reductively in terms of the use of nonnormative

(including classical modal) vocabulary, or undertake to do without it (for instance, in

accounts of semantic content or intentionality). On the other hand, if the legitimacy and

in-principle intelligibility of normative vocabulary is not in this way globally challenged,

a very different sort of enterprise can still be in order: explicating and clarifying the

expressive role played by normative vocabulary in a context that includes both normative

and nonnormative vocabulary.

I alluded above to Kant's argument for the conclusion that (to put it in my words) what is

made explicit by modal vocabulary can be see already to be implicit in the use of

ordinary, nonmodal vocabulary. In fact, I think his argument is much broader, and

concerns *normativity* generally, rather than just *modality*. The scope of his claims is easy

to overlook for two reasons. First, he couches it in terms of 'necessity' (Notwendigkeit),

which we automatically but I think anachronistically read back in terms of alethic

modalities. Second, we tend to think of Kant (at least in this context), as responding to

Hume's worries about induction, causation, and necessary connection in *theoretical*

philosophy. But on the first point, Kant uses the same term, 'Notwendigkeit', to talk

about *moral* necessity. Indeed, for him, natural necessity and moral necessity are species

of a genus—and the *genus* clearly is not alethic modality. And on the second point, Kant

was equally concerned to respond to Hume in *practical* philosophy. He sees the issues

about the relation between *must be* and *is* as of a piece with those about the relation

between *ought* and *is*.[22]

What ties the two together for Kant is precisely *normativity*. For him 'necessary' just

means *according to a <u>rule</u>* (his jurisprudentially influenced term for the form of norms).

Objectively binding rules he calls 'laws'—on the practical, as well as the theoretical side.

Specifically moral norms and those invoked in a scientific setting are just two species of

this genus. The most basic sort of rule he calls 'concepts', and his overarching concern is

to understand the bindingness or validity (Gültigkeit) of concepts. Perhaps Kant's most

basic insight is that what it is that *judgments* and *actions*—the subject matters of

theoretical and practical philosophy respectively—have in common that distinguishes

them from the responses of mere animals is that they are things we are in a distinctive

way *responsible* for, that they express *commitments* of ours. Central to what we are

responsible *for* (doing) is giving (or at least having) *reasons* for them. And what counts

as a reason for a given judgment or action is determined by the *rules* we govern our

judging or acting—that is, the *concepts* that we thereby count as *applying*.[23] So for Kant

the discursive realm ('discursive' for him just meant 'of or pertaining to concept use'),

that is, the intentional, is picked out by its being articulated according to a certain kind of

norm: norms governing what is a *good reason*, what else one counts as *committing*

---

[22]   Kripke implicitly rediscovers this connection in modeling his skeptic's worries about 'ought' on Hume's inductive skeptic's worries about 'necessarily'. In fact in English (and German) we are quite comfortable using 'must' to express both Kantian natural necessity and Kantian moral necessity.

oneself to by a certain judgment or action, and what would count as *justifying* those commitments. Grasping the conceptual or intentional content of a state or performance requires mastering those inferentially articulated norms that determine when it would be *correct*, and what else one has *obliged* oneself to thereby.

Kant goes on to draw conclusions about the specific forms of judgments in which concepts must be able to participate in order to play this normative role. Leaving aside the details of his argument, what connection can be recovered between this general view about the inferentially articulated normativity of intentionality as such and discerning alethic modal involvements as implicit in the use of ordinary descriptive concepts? In an early paper, Sellars says "I shall be interpreting our judgements to the effect that A causally necessitates B as the expression of a rule governing our use of the terms 'A' and 'B'."[24] Elsewhere, he develops this idea, (following out a thought of Carnap's[25]), according to which "the language of modalities is interpreted as a 'transposed' language of norms." The transposition is from the formal to the material mode, from talk about talk about things to talk about things. The idea is that endorsing the claim that A necessitates B is endorsing the propriety of a certain kind of inference from ⌈A⌉ to ⌈B⌉. Of course what the claim that A necessitates B *says* is *not* that this inference is good. It doesn't mention expressions, proprieties, or inferences. But *understanding* the claim requires being able to distinguish the inferences one would become committed to by endorsing it. Those normative consequences for what inferences one who applies the

---

[23]   In the latter case, this is bringing the performance under a maxim.
[24]   "Language, Rules, and Behavior" footnote 2 to p. 136/296 in *PPPW*, op. cit.
[25]   "Inference and Meaning," pp. 279/331-284/336 in *PPPW*, op.cit.

concept of necessary consequence thereby becomes committed to are in that sense

implicit in the concept. Making them *explicit* (the reverse transposition from material to

formal mode) requires concepts pertaining to the *use* of expressions that are not made

explicit by that use—concepts such as expression, inference, and (most importantly for

our purposes here) *normative* concepts such as propriety, commitment and entitlement,

obligation and permission.[26] That a given bit of vocabulary expresses causal

necessitation, an alethic modality, is a matter of the sort of inference it implicitly involves

(that one implicitly endorses by applying the concept). The inferences its use commits us

to are counterfactually robust. All the coins in my pocket are copper. It follows that if I

had heated one of them to 1082° C., it *would* not have melted, but if I *had* heated it to

1084° C., it *would* have, (though I didn't and it didn't). But it does *not* follow that if this

nickel *had* been in my pocket, it *would* have been copper.

I said in the previous section that Kant and Sellars think there is an argument that the use

of any ordinary empirical descriptive concept commits us to the correctness of at least

some counterfactually robust inferences (though I didn't pretend to present such an

argument). Thus the use of any such concept presupposes the applicability in principle of

modal vocabulary, whose expressive role it is to make explicit the difference between the

two sorts of inferences illustrated by those concerning the coins. So if what is expressed

by the use of ordinary empirical descriptive predicates is intelligible in principle, then so

is what is expressed by modal vocabulary. My present point is that the background of

that argument is a more general one, to the effect that mastering the use of ordinary

---

[26] Sellars gestures at this point by his somewhat dark distinction between what one "asserts" by using a

empirical descriptive predicates, which is in practice understanding their content or the meaning they express, requires being able to distinguish some uses and inferences as *correct*, and others as *incorrect*. It requires knowing (in the practical sense of being able to distinguish, a kind of know-*how*) at least something about what one is *committing* oneself to by applying the concept, and what would *entitle* one to do that. Treating one descriptive predicate as applicable in a particular case *obliges* one to consider others (suitably inferentially related to it) as applicable, (normatively) *precludes* one from applying others, and *licenses* one to apply others. Since this essential dimension of the *use* of even ordinary, descriptive, nonnormative concepts (in belief and judgment no less than in linguistic assertion) is what is made explicit by normative vocabulary, it cannot be that ordinary empirical descriptive concepts are coherent and intelligible in principle, but normative concepts are incoherent and unintelligible in principle. The discriminations and significances that are made explicit by central kinds of normative vocabulary are implicit already in the ability to use and understand nonnormative vocabulary. One may not yet have the words to make them explicit, but just by using the nonnormative vocabulary one already has all the abilities needed to understand the basic normative concepts. This argument is strictly parallel to that for the implicit, practical mastery of what is made explicit by *modal* vocabulary being part of the background against which alone mastery of the use of *nonmodal* vocabulary is intelligible.

This means that no concept user can be in a position of being skeptical about the in-principle intelligibility of specifically *normative* concepts, without thereby being skeptical

---

sentence, and what one merely "conveys" by using it. Ibid. p. 281/333.

about the in-principle intelligibility of concept use (and so meaning, content, and intentionality) generally. It is a version of this thought that sets the stage for Kripke's Wittgenstein's skeptic's argument for precisely that paradoxical attitude: a *reductio ad absurdam* of the coherence of conceptual normativity, and so of the use of concepts, including those appealed to in mounting the argument. The form of that argument was a challenge: to say what it is about how a concept *is* or *has actually* been applied in the past that settles how it *ought* to be applied in the future. The force of the argument turns critically upon what vocabulary one is allowed to use in specifying the past use. Kripke's skeptic is willing to countenance *modal* vocabulary—for instance the invocation of dispositions. Now the rules for determining just what vocabulary one *is* allowed to use, and their rationale, are never explicitly discussed. But I think it is clear that Kripke's skeptic is *not* willing to countenance the use of *normative* vocabulary in describing the past use. For if one acknowledges that applying a concept *correctly* or *incorrectly* is also something we *actually* do and have done, then the problem takes on a very different shape. It becomes the problem of projecting future proprieties from past proprieties, of saying what it is for us to be bound *now* by a *prior* decision about what was correct and incorrect, about how an expression ought to be used. Such a question invites certain kinds of philosophical explication, but does not threaten to yield a paradox.

In asking how the actual use of concepts could determine how they ought to be used, how it would be correct to use them, Kripke's skeptic implicitly restricts the vocabulary used to specify that use to *nonnormative* vocabulary. He thus assumes that normative specifications of proprieties of concept use are in principle intelligible only if they can be reduced without remainder to specifications of nonnormative properties of concept use. But the more general Kantian argument we have been considering invalidates this assumption. We cannot be in a position where the specification of the use of concepts in nonnormative vocabulary is unproblematic, but the applicability of richer normative vocabulary in specifying that use is in principle questionable. For grasping the nonnormative descriptions requires already making the discriminations that are expressed by normative vocabulary. The implicit restriction to nonnormative specifications of the actual use of concepts begs the question, in the context of acknowledgments of the

essentially normative character of concept use. The credentials of normative concepts should no more be held hostage to their reducibility to nonnormative ones than the credentials of modal concepts should be held hostage to their reducibility to nonmodal ones, and for the same reasons.

If that is right, then there is methodological room for another sort of explanatory strategy for addressing philosophical issues about intentionality, besides the one considered in my opening remarks. For just as it is legitimate, where helpful, to use *modal* vocabulary in elucidating and explaining various features of intentionality, so it is in principle legitimate to appeal to *normative* vocabulary to do so. It is not circular or otherwise objectionable to appeal to non-extensional modal concepts to make sense of the kind of intensionality peculiar to intentional phenomena. And it is not circular or otherwise objectionable to appeal to normative concepts to make sense of the specifically *conceptual* normativity distinctive of intentional phenomena. In short, I think there is room to lay *normative* approaches to intentionality alongside the *modal* ones made familiar by Dretske, Fodor, and others (including, in her distinctive way, Millikan).

I have claimed that we should reject global skeptical challenges to the in-principle intelligibility of normative concepts, as we should reject corresponding challenges to the intelligibility of modal ones. They proceed from and depend on an ultimately unsustainable picture of the meanings of extensional, descriptive, empirical predicates as autonomously intelligible apart from counterfactual-supporting connections to other predicates, and normative features of their use. But to reject that picture, and with it the global skeptical challenges it sustains, is not to reject more local demands for explication

of the expressive and inferential roles played by vocabulary of the various sorts, and of interrelations among them—the sort of enterprise to which possible world semantics makes such a signal contribution in the case of modal vocabulary. An explicative response does not require the reduction of normative to nonnormative vocabulary, but its clarification as a component in a whole that comprises both, as well as modal vocabulary.

One can imagine many ways of approaching this undertaking. Two that have been worked out in some detail are explications of the norms internal to intentionality in *Darwinian* or *developmental* terms, as Millikan elaborates in *Language, Thought, and Other Biological Categories*, and explications of those norms in *Wittgensteinian* or *social* terms, as I do in *Making It Explicit*.

I am not claiming that these alternatives are exhaustive, or even exclusive.[27] Nor will I attempt here to assess their relative merits. I want rather to conclude the abstract methodological discussion I have presented here with a few considerations about the explanatory ecological niche that such accounts might be seen to occupy. One reason to want a story of the genus of which these two are species is that it is important that the *advent* of intentionality not be left seeming magical or mysterious. There was a time when there was no conceptually articulated activity—nothing bound by or subject to assessment according to conceptual norms. Later, there was. An important part of understanding ourselves is understanding the nature of this transition. The final point I want to make concerns the criteria of adequacy it is appropriate to impose on candidate contributions to such understanding.

For it is, I think, in the first instance in connection with these considerations that it comes to seem right and proper to insist on a *naturalistic* account of intentionality in all its aspects. Now if the alternative is to allow *super*natural explanations, it is hard to see how to object to such a commitment. But, granting that, I think it is useful to keep in mind that there are still options to be considered. One of them, to be sure (perhaps the

---

[27] I am very interested, for instance, in what I see as Hegel's way of combining social and developmental (though not, for pardonable chronological reasons, Wittgensteinian and Darwinian) considerations in his account of the nature and content of conceptual norms.

one that comes most easily to mind) is to offer an account that is naturalistic in the sense of being

reductive—showing how conceptual normativity takes its place as a species of a genus that is intelligible

entirely in terms of vocabulary already in use in the natural sciences. Millikan's story is a *tour de force* of

this kind. But there are other possibilities. One might be a nonnaturalist about conceptual normativity

simply in the sense of being nonreductionist about it. On the approach I have recommended, for instance,

normativity of the sort that articulates discursive practice (and so intentionality in the sense I address) is

seen as essentially a product of *social* interactions: there were no commitments and entitlements, no

authority and responsibility, until creatures started in practice *taking* each other to be committed and

entitled, *recognizing* or *attributing* authority, *holding* each other responsible. The normative statuses that

are products of such essentially social interactions are the subject matter, not of the *natural* sciences, but at

most of the *social* sciences. Explicating them may require a vocabulary that is in principle richer than that

employed by the natural sciences[28]—for instance, a normative vocabulary. *Non*naturalism of this

nonreductive kind is still safely removed from any sort of *super*naturalism.

In fact it appears that one can mix and match the two kinds of approach to conceptual norms with

naturalism or nonnaturalism. Thus—picking a few prominent examples—I take it that a broadly social

approach to conceptual normativity is pursued in a nonnaturalistic form by Wittgenstein, and in a

naturalistic form by Crispin Wright. A broadly developmental approach is developed in nonnaturalistic

form by Hegel, and in naturalistic form by Millikan. Within Millikan's own account, it is a worthwhile

enterprise to distinguish those features that are consequences of her general naturalistic commitments, those

that are consequences of the particular developmental-evolutionary way she follows out her program, and

those that are features simply of taking conceptual normativity seriously.[29] Prominent among the latter, for

---

[28]   In spite of the remark above, while Millikan begins her story with concepts available already to the natural sciences, it is much less clear that by the end she has not introduced social considerations of a sort that go substantially beyond those countenanced by, say, natural history or population biology. Insofar as her project depends merely on the a continuous development of concepts rooted in biology, there may be less difference between our approaches than initially appears.

[29]   There are many other separable threads in Millikan's rich tapestry, whose responsibility for various attractive features and consequences it would be interesting to investigate: the explanatory priority accorded to propositional representations, looking downstream to consumers of representations as well as upstream to their producers, the fact that her functionalism is not only normative (rather than simply causal), but also

instance, is her denial of meaning rationalism: the claim that the contents of our thoughts and our concepts is transparent to us. Part of the significance already of Kant's displacement of Cartesian certainty as the central phenomenon in concept use in favor of the rule-governedness he called 'necessity' was a shift from concern with our grip on concepts to concern with their grip on us. So he asks not "Are they clear and distinct to us?" but "Are they valid for or binding on us?"[30] According also to Wittgenstein's nonnaturalized social account of conceptual norms (or that of *Making It Explicit*), the commitments one undertakes by making a move in a language game may well outrun what the one undertaking the commitment appreciates. Again, any theorist who thinks of applying a concept primarily as undertaking a commitment or binding oneself by a norm ought to find the undertaking of incompatible conceptual commitments (and so having an incoherent constellation of commitments) no more mysterious in principle than is the possibility of binding oneself by incompatible promises. White Queen psychology too has nonnaturalistic varieties.

<p style="text-align:center">***</p>

A prominent and welcome feature of the philosophical landscape of our times is a set of research programs offering reconstructions of notions of intentionality and representation that are vastly more detailed and well worked-out than any previously available. They are standardly and properly seen as developments in two traditions. First, as regards its aim, the common enterprise epitomized by the work of Dretske, Fodor, and Millikan seeks a *naturalistic* treatment of the central phenomena of intentionality and semantics. This aspiration is recognizable as a version of the unity of science program that motivated

---

social (both diachronically and synchronically), rather than individualistic, and so on. (For instance, the superiority of her account of the magnetosome containing bacteria over Dretske's derives from her consideration of the consumers of representations. See pp. 125ff of *White Queen Psychology*.)

[30]   This feature of the normative conception of conceptual content is of the utmost importance in making palatable *holism* about such contents. For the problem of how it is possible to *share* such contents (and so

Tarski, Carnap, and Quine. Second, as regards its explanatory raw materials, this constellation of approaches is distinguished by its reliance on and comfort with the use of alethic *modal* vocabulary: appeals to dispositions, counterfactuals, and so on. These are idioms that their naturalist predecessors regarded as at best themselves too much in need of explanation for them to be eligible to serve in explaining intentional or semantic idioms, and at worst as in principle defective and incoherent. I think the contemporary attitude towards the legitimacy of modal locutions is entirely in order. But I also think that understanding *why* it is in order opens up a broader intellectual context within which it is illuminating to locate these semantic projects. So I have said next to nothing about explanatory virtues and vices of particular undertakings in this tradition. My concern in this essay has rather been to try to understand better the philosophical landscape they occupy, and so their relation to other approaches to intentional and semantic phenomena, which share with them neither their motivating naturalism nor the central methodological place they accord to recently rehabilitated modal notions.

---

to communicate) takes on a completely different shape when one understands it as a question of how individuals with different beliefs can come to be bound by a shared set of norms.