

The Normativity of the Intentional

Ralph Wedgwood

Many philosophers have claimed that *the intentional is normative*. (This claim is the analogue, within the philosophy of mind, of the claim that is often made within the philosophy of language, that *meaning is normative*.) In the first two sections of this paper, I shall give a brief clarification of what this claim that “the intentional is normative” actually means. In the third section, I shall consider a number of arguments that other philosophers have advanced in favour of this claim; as we shall see, many of these arguments are inconclusive at best. However, in the last section, I shall give a sketch of a different argument, which I regard as a persuasive argument for this claim.

1. “The intentional is normative” – what does that mean?

What is meant here by ‘the intentional’? For present purposes, I shall take the “intentional” to be a subset of the mental. Many mental states—indeed, some philosophers would say, all mental states—are *about* something; they are “concerned with” or “directed towards” something. For example, some of my beliefs are *about Norway*; some of my intentions are *about returning to Oxford*; and some of my fears are *about snakes*. As many philosophers put it, these mental states all have *intentional content*, and so may be called “intentional states”.¹

What does ‘normative’ mean? For the purposes of this paper, I hope that it will

¹ Compare John Searle, *Intentionality* (Cambridge: Cambridge University Press, 1983), esp. chap. 1.

be sufficient if we just start with the assumption that the paradigmatic normative terms are ‘ought’, ‘should’, ‘right’, ‘wrong’, and their closest equivalents in other languages. In addition to terms like ‘ought’, ‘should’, ‘right’ and ‘wrong’, I shall assume that terms like ‘correct’ and ‘incorrect’, ‘rational’ and ‘irrational’, are also normative terms.

Of course, the interpretation of these terms is philosophically controversial.² My own view is that as most formal semanticists would agree,³ most of these terms are context-sensitive, and express different concepts in different contexts. Thus, there is not just the ‘ought’ of morality, but also the ‘ought’ that is relative to some particular goal or purpose, the ‘ought’ of rationality, and many others. It may be possible to give a deeper account of what unifies all normative concepts.⁴ For present purposes, however, it will probably be enough if we base our grasp of what normative concepts are simply on our intuitive understanding of terms like ‘ought’.

What then does the claim that the intentional is normative amount to? The answer to this question may depend, in part, on what is the correct sort of semantics for normative concepts. First, suppose that a broadly truth-conditional or factualist semantics for normative concepts is correct. According to this sort of semantics, whenever propositions involving these normative concepts are true, they are true in virtue of corresponding *normative facts*; and it is convenient to suppose that these

² For example, Robert Brandom argues that normative truths are in some sense “imposed” by “human practices”, in his book *Making It Explicit* (Cambridge, Mass.: Harvard University Press, 1994), chap. 2—whereas I would myself insist that there must be some fundamental normative principles that are simply necessary truths, and so wholly independent of any contingent facts about human practices. But I shall have to remain neutral about these philosophical controversies here.

³ See especially Angelika Kratzer, “The Notional Category of Modality”, reprinted in *Formal Semantics: The Essential Readings*, ed. Paul Portner and Barbara Partee (Oxford: Blackwell, 2002).

⁴ Indeed, I have attempted to give such a unifying account elsewhere. Roughly, my idea is that what unifies all these normative concepts is that it is a *constitutive* feature of these concepts that they play a *regulative* role in certain rational practices; see my paper “The Aim of Belief”, *Philosophical Perspectives* 16 (2002), 267-297, esp. 268-69; for a more detailed account of a particular normative concept, see “Conceptual Role Semantics for Moral Terms”, *Philosophical Review* 110 (2001), 1-30.

normative facts must involve corresponding *normative properties and relations*.⁵ Then we can interpret the claim that the intentional is normative as the claim that in giving an account of the essence or nature of intentional states, we must mention these normative properties or relations.⁶ Equally, in giving an account of the nature of an intentional fact—an account of what the intentional fact consists in, as some philosophers would say—we must state some normative fact. In that sense, intentional facts are partially constituted by normative facts.⁷

On this understanding, then, the claim that the intentional is normative is a metaphysical claim concerning the nature or essence of intentional states, or concerning the constitution of intentional facts. It is not a semantic thesis about the meaning of intentional terms, or a conceptual thesis about what is built into our concepts of intentional states. Such semantic or conceptual theses could be true even if the corresponding metaphysical thesis is false. It could be that, even if we ordinarily *think* of intentional states as states that we “ought” to be in under such-and-such circumstances, the *nature or essence* of those states can be explained in wholly non-normative terms.

The claim that the intentional is normative may have to be understood a bit differently if an *expressivist* account of normative concepts (instead of a broadly truth-conditional or factualist account) is correct. According to the expressivist, although it is perfectly all right for everyday conversation to talk about “normative

⁵ Philosophical of broadly *nominalist* sympathies may be reluctant to countenance such normative properties and relations. But the debate between nominalists and realists about properties and relations is clearly independent of the debates about the normativity of the intentional that we are focusing on here. For our purposes, speaking about properties and relations is useful for achieving generality; but it could in principle be dispensed with if it were necessary to do so.

⁶ To fix ideas, we may think of an “account of the essence or nature” of something in roughly the same way as Kit Fine would conceive of a statement of the “constitutive essence” of the thing; see Fine, “Senses of Essence”, in Walter Sinnott-Armstrong, ed., *Modality, Morality, and Belief* (New York: Cambridge University Press, 1995), 56.

⁷ Compare the interpretation of this claim given by Gideon Rosen, “Brandom on Modality, Normativity and Intentionality”, *Philosophy and Phenomenological Research* 63 (2001), 611-623.

facts” or “normative properties and relations”, no such normative facts, properties or relations are capable of playing a real explanatory role in an account of the essence or nature of our intentional mental states. So, on this expressivist understanding of normative concepts, the claim that the normative is intentional would have to be understood as a claim about concepts, rather than as a claim about properties and relations. Specifically, it would have to be the claim that it is an essential feature of our concepts of intentional mental states that competent users of these concepts must think of intentional mental states in partly normative terms.⁸

If the claim that the intentional is normative is true (according to either of these two ways of interpreting this claim), then it will be possible, at least in principle, to give an account of the nature of our intentional mental states (or concepts) in partly normative terms. In fact, such normative accounts of the nature of intentional states could take many forms. For example, some of these accounts may claim that the intentional is actually *reducible* to the normative—that is, that there is a correct reductive account of the nature of intentional states, and any such account must mention normative properties or relations.⁹ Other accounts may claim merely that the normative and the intentional must each be mentioned in giving any account of the other, so that the two domains of properties and relations are essentially interdependent, without either of the two being reducible to the other. These

⁸ For an exploration of the possibility of combining an expressivist account of normative concepts with the claim that the intentional is normative, see Allan Gibbard, “Meaning and Normativity”, in Enrique Villanueva, ed., *Truth and Rationality* (Atascadero, Calif.: Ridgeview, 1994). Gibbard also interprets Brandom’s *Making It Explicit* as articulating a view of this sort; see Gibbard, “Thought, Norms, and Discursive Practice”, *Philosophy and Phenomenological Research* 56 (1996): 699-717. (Other philosophers have interpreted Brandom differently; see Gideon Rosen, “Who Makes the Rules Around Here?”, *Philosophy and Phenomenological Research* 57 (1997): 163-171. I am not myself certain which interpretation of Brandom is correct.) In fact, I believe that there are very serious problems with this attempt to combine expressivism about normative concepts with the claim that the intentional is normative; but I cannot attempt to explain what these problems are here.

⁹ Compare e.g. Brandom, *Making It Explicit*. (At least, this is how Gideon Rosen interprets Brandom in “Who Makes the Rules around Here?”) See also Michael Morris, *The Good and the True* (Oxford: Clarendon Press, 1992).

normative accounts could also differ in many other ways as well. In what follows, however, I shall abstract from all these differences between different normative accounts of intentional states. I shall also not be able to explore the metaphysical and methodological implications of the claim that the intentional is normative. Instead, I shall focus on examining the reasons that philosophers have put forward in favour of this claim.

2. What sort of account is called for, anyway?

In the last two sections of this talk, I shall consider various arguments that philosophers have given in favour of the claim that the intentional is normative. But before doing that, I shall consider a rather abstract objection that can be made against this claim. Why should there be any informative “account” of the intentional at all? Perhaps intentional mental states are wholly irreducible properties of thinkers, of which no informative “account” can be given. Some philosophers are committed to a strong form of naturalism, according to which it must be possible, at least in principle, to give a reductive account of the intentional in wholly naturalistic, non-normative terms. But the claim that there is a reductive account of this sort hardly supports the claim that there is also a *normative* account of the intentional. Thus, such normative accounts of the intentional may seem to be a curious compromise, half way between a fully reductive naturalistic approach, and a doggedly quietist non-reductive approach.

It seems to me that there is a problem with the doggedly quietist approach. This is because of a plausible point about the relation between the intentional states that we are in and our *dispositions*. (When I refer to our “dispositions” here, I do not mean to focus exclusively on our behavioural dispositions; I mean to include our

mental dispositions as well—such as our dispositions to revise our beliefs, intentions and other attitudes, in response to various conditions.) This point is an analogue of the point that is sometimes expressed by the slogan “meaning is use”: just as the meaning of a linguistic expression is determined by the ways in which the members of the relevant linguistic community are disposed to use the expression in discourse, so too the identity of a concept is determined by the ways in which we are disposed to use that concept in thought.

Suppose that a community has a concept in their repertoire that they are all disposed to apply, quite systematically, to cats and only to cats; they base their judgments involving this concept on various sorts of evidence, but most often on precisely the sort of experiences that we would use to recognize something as a cat; and they also treat everything that falls under this concept as belonging to a single natural kind. Now consider the supposition that in this case, this concept is in fact the concept of a *bicycle*, so that almost all of these thinkers’ beliefs involving this concept are irrational and false. This supposition appears absurd.¹⁰ In general, it seems that if there is a community the members of which possess a concept that they are disposed to use in their reasoning in exactly the same ways as we are disposed to use a certain concept in our repertoire, then their concept must be the same as ours. So it seems that it must be something about the thinker’s dispositions—together perhaps with certain facts about his environment and the dispositions of other thinkers in his linguistic community—that determines which concepts figure in the thinker’s thoughts.

Similarly, if there is a community that has *type of attitude* that they are disposed to form and revise in just the same way as we are disposed to form and revise our beliefs, it would be absurd to suppose that this attitude is in fact the attitude

¹⁰ Compare Elizabeth Fricker, “Analyticity, Linguistic Practice and Philosophical Method”, in Klaus Puhl, ed., *Meaning Scepticism* (Berlin: De Gruyter, 1991).

of *disbelief*—so that almost all of their attitudes of this type are incorrect and irrational. Again, something about our dispositions with respect to an attitude-type must determine exactly which attitude-type it is that figures in our thoughts.

However, if it is true that facts about the thinker's dispositions determine which concepts and attitude-types figure in the thinker's thoughts in this way, then there surely must be some account of the precise way in which these dispositions determine this.¹¹ The bare claim that the thinker's dispositions determine which concepts and attitude-types figure in the thinker's thoughts is compatible with infinitely many different precise ways in which these dispositions might determine this. Surely there must be some general principle that captures the specific role that these dispositions play in determining this. Otherwise it will remain utterly mysterious why these dispositions determine this in the specific way in which they do, rather than in any of the infinitely many other possible ways.

So it seems that there must, at least in principle, be some explanatory account of how the thinker's dispositions determine what intentional states he is in. There is no reason why this account should take the fully reductive form of giving an account of intentional states in terms of dispositions that are characterized in wholly non-intentional, non-normative terms. On the contrary, there is no objection to characterizing these dispositions in *general* intentional terms—that is, precisely as *mental* dispositions, dispositions to form certain attitudes involving certain concepts in certain circumstances—so long as our characterization of these dispositions does not presuppose the identity of the particular concepts and attitude-types that they involve. It may well be a completely irreducible feature of us that we have attitudes and use concepts, or a completely irreducible feature of some state that we are in that

¹¹ Compare Paul Boghossian, "Naturalizing Content", in Barry Loewer and Georges Rey, eds., *Meaning in Mind: Fodor and His Critics* (Oxford: Blackwell, 1991).

it is a state of standing in a mental relation to some conceptual content. However, it cannot be an irreducible feature of me that I am using a concept that refers to *cats* (as opposed to referring to *dogs* or to *cows*, for example); that must be explained in terms of the dispositions that I have with respect to this concept.¹²

So the explanatory account that seems to be called for could take the following form. It could start out by presupposing that we are dealing with a thinker who has attitudes towards contents that are composed of concepts.¹³ But it would start by treating these attitudes and concepts as initially *uninterpreted*. That is, it would start with a “neutral” characterization of the thinker’s intentional dispositions—a characterization of the thinker’s dispositions with respect to these attitudes and concepts that does not in any way presuppose the exact identity of these attitudes and concepts. Then this account could articulate certain principles concerning the nature of various attitudes and concepts, which would enable us to identify the thinker’s attitudes and concepts—say, as the attitude of *belief*, or as the concept of *water*—on the basis of such a neutral characterization of the thinker’s intentional dispositions (along with further facts about the thinker’s natural and social environment). In what follows, I shall assume that all adequate accounts of the nature of our intentional mental states take this form.

Certain normative accounts of the intentional could certainly accommodate this point. For example, such an account might consist of two parts. The first part would give an account of the different attitudes and concepts in terms of certain

¹² In effect, an account of what it is for a thinker to be using the concept *cow* may employ both the term ‘concept’ and the term ‘cow’, but it may not employ the phrase ‘the concept *cow*’ or ‘refers to cows’ or any variant thereof. It should conform to Christopher Peacocke’s constraint, that an adequate account of a concept should take “the A(C) form”; see his *A Study of Concepts*, 5-16. Compare Saul Kripke’s insistence that an account of why a given name refers to a particular individual must not be “circular”, in *Naming and Necessity* (Cambridge, Mass.: Harvard University Press, 1980), 68-70.

¹³ It may also be necessary to start out with presuppositions about the syntactic category that the various concepts belong to—whether they are general or predicative concepts or singular concepts or propositional operators, and so on.

normative principles governing them. The attitude of belief, for example, might be identified as that attitude that it is *correct* or *appropriate* to take towards a proposition just in case that proposition is true; the concept ‘if’ might be identified as that concept for which certain sorts of inference are *rational* (such as inferences that are instances of *modus ponens*); and similar accounts could be given of other concepts and types of attitudes. Second, there would have to be an account of what has to be true of a *thinker* if she is to possess those concepts or to be capable of having those attitudes. Here the account could say, very roughly, that the right way to interpret the thinker must be the most *charitable* reading of his dispositions, the way that makes thinkers’ dispositions emerge as most rational and most sensitive to the norms that apply to the intentional states that they are in.¹⁴

So a normative account of the intentional could satisfy the demand for an explanatory account of how exactly thinkers’ dispositions (and environment) determine which intentional states they are in. But what reason is there to think that any such account is true?

3. Some questionable arguments for the normativity of the intentional

3.1 The entailment argument

Some philosophers try to argue for the claim that the intentional is normative by pointing out that facts about one’s intentional states *entail* normative facts. Then they

¹⁴ For the idea that interpretation must be guided by a principle of charity, see Donald Davidson, “Radical Interpretation”, in his *Inquiries into Truth and Interpretation*, 2nd edition (Oxford: Clarendon Press, 2001). A normative theory could avail itself of this idea without accepting Davidson’s full-blown “interpretivism”. As David Lewis suggests, the reference to interpretation could just be taken as a way of dramatizing what is objectively constitutive of the intentional states in question; see Lewis, “Radical Interpretation”, *Synthese* 23 (1974): 331-344.

argue that this entailment shows that the intentional is in itself essentially normative.¹⁵

For example, the argument might be elaborated as follows. Necessarily, anyone who believes *p* is *committed* to believing ‘*p or q*’. That is, it is a necessary truth, applying to everyone, that one *ought not* simultaneously to believe *p* and consciously to refuse to believe ‘*p or q*’ (that is, to consider whether or not ‘*p or q*’ is the case, and yet to fail to believe ‘*p or q*’). But how else are we to explain *why* this is a necessary truth, unless this truth logically follows from principles that are essential to or constitutive of all beliefs whose content is of the form ‘*p or q*’? And if a normative truth follows from these principles, mustn’t these principles themselves mention normative properties or relations?

One might start by questioning whether it really is true in general—let alone necessarily true—that one ought not simultaneously to believe *p* and consciously to refuse to believe ‘*p or q*’. What if a powerful demon will destroy the world unless you simultaneously believe *p* and consciously refuse to believe ‘*p or q*’?¹⁶

The best response to this sort of objection, I think, is to insist that there are different sorts of *oughts* involved here. It is still a “mistake” of some kind simultaneously to believe *p* and consciously to refuse to believe ‘*p or q*’—even if the demon will destroy the world unless you commit this mistake. But now the proponent of this argument must say more about what exactly this sort of ‘ought’ means. It may be that it just means: “ought, in order to achieve the goal of believing the truth and nothing but the truth (about any question that one has actually considered)”. If that is right, then we can give a simple explanation of why it is necessary that we “ought”, in this sense, to avoid simultaneously believing *p* and consciously refusing to believe ‘*p*

¹⁵ This is very close to an argument that Kripke used to argue for the normativity of meaning. See Kripke, *Wittgenstein on Rules and Private Language* (Cambridge, Mass.: Harvard University Press, 1982), 11, 21, 24, 37.

¹⁶ Compare Alex Byrne, “Semantic Values?”, *Philosophy and Phenomenological Research* 65, no. 1 (July 2002): 201-207.

or q ': this is necessary because it is necessary that if p is true, so is ' p or q '. This explanation it appeals only to the meaning of this sort of 'ought', and to the connection between these intentional states and their truth conditions. It seems to make no appeal to the claim that the intentional is in itself intrinsically normative.¹⁷

Moreover, even if the sort of 'ought' that is involved here is not susceptible to any such deflationary interpretation, we still do not have a sound argument for the normativity of the intentional. As Paul Horwich writes:¹⁸

[A] property may have normative implications without itself being normative. For example, killing is *prima facie* wrong; none the less one can presumably characterize ' x kills y ' in entirely non-normative terms. ... It might be possible for our most basic normative principles to have the form

$$(x)(Dx \rightarrow Nx)$$

where " D " describes some state of the world and " N " specifies what ought or ought not to be done in that situation. Thus the normative implications of a meaning property leave it entirely open that its nature is completely non-normative.

This "entailment" could be explained by the existence of certain fundamental normative axioms; but there is no reason to suppose that these normative axioms are *constitutive* or *essential* to the intentional states that they apply to. So, the entailment argument seems not to be a persuasive argument for the normativity of the intentional.

3.2 *The interpretation argument*

Some philosophers argue for the claim that the intentional is normative on the basis of considerations concerning *interpretation*. This argument is based on two premises.

According to the first premise, the central or canonical way to ascribe mental

¹⁷ This objection and the next are due to Paul Horwich, *Meaning* (Oxford: Clarendon Press, 1998), chap. 8.

¹⁸ Horwich, *Meaning*, 188-193.

states to others is to attempt to ascribe those mental states that make best sense of them and of their behaviour—where making best sense of them requires following a *principle of charity*, ascribing those mental states that give the most charitable possible interpretation of them and their behaviour.¹⁹ Furthermore, according to this first premise, this principle of charity must be understood in normative terms: to interpret someone charitably is to interpret her as forming beliefs and intentions in a *rational* way, and as having at least a disposition to have experiences, desires and emotions that really are *correct* or *appropriate*, and so on.²⁰

According to the second premise, the canonical method of ascribing mental states to others must proceed by attempting to track those facts about them that really make it the case that they have the mental states in question. The first premise appears to imply that this canonical method attempts to track certain normative facts about the subjects who are being interpreted—namely, the normative fact about what is the most charitable interpretation of them. So it must be this normative fact that makes it the case that they have those mental states; and so a version of a normative account of the intentional is true.

Both premises of this argument can be questioned. The second premise seems too quick to assume that what is fundamental to our canonical method of coming to know some fact must also be fundamental to the nature of that fact itself. The “canonical method” could still provide us with true beliefs about the mental states of others, even if it were only a *contingent* fact about us that we tend to form beliefs and intentions in rational ways, and to have experiences and desires that are correct or appropriate. In that case, the normative facts that guide our interpretation would

¹⁹ See esp. Donald Davidson, *Essays on Actions and Events* (Oxford: Clarendon Press, 1980), 221, and *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984), 153.

²⁰ One philosopher who insists that the principle of charity (at least as it applies to the ascription of preferences), must be interpreted in evaluative terms is Susan Hurley, *Natural Reasons* (Oxford: Clarendon Press, 1989), chap. 6.

generally lead us to the truth about the mental states of thinkers in the actual world, but they would not be in any way constitutive of the mental states themselves.

The first premise could also be questioned. Perhaps the principle of charity can be formulated in entirely non-normative terms. For example, perhaps we can define which interpretation is most charitable in purely logical terms, as the interpretation that makes the subjects emerge as much as possible as believing the *truth*, and conforming to certain standards of *consistency* and *probabilistic coherence* in their beliefs and preferences. Or perhaps interpretation should follow, not a principle of charity, but a principle of *humanity*, ascribing those mental states that make the subjects emerge as much as possible like statistically normal human beings.²¹ No doubt, if we have *independent* reasons for accepting that the intentional is normative, then it will become much more plausible that the canonical method of interpretation involves following a normative principle of charity. But in the absence of such independent reasons, we have not been given much reason to accept the first premise of the argument. For this reason, the interpretation argument seems dialectically ineffective as an argument for the claim that the intentional is normative.

3.3 *The argument from the alleged failure of non-normative dispositionalism*

Many of the philosophers who claim that the intentional is normative seek to support their claim by arguing against the obvious alternative—what I shall call the “non-normative dispositionalist” view, according to which the nature of our intentional states is determined by features of our dispositions that can be specified in wholly non-normative terms.

²¹ See Richard Grandy, “Reference, Meaning, and Belief”, *Journal of Philosophy* 70 (1973): 439-452.

Here I shall just discuss the two most challenging objections that any non-normative dispositionalist theory must face. Both objections stem from the point that if the nature of a given concept is determined by our dispositions to form beliefs and other attitudes involving that concept, then those dispositions must surely also fix the *extension* or *semantic value* of that concept. But it may seem that there are at least two reasons why our dispositions cannot fix the extension of our concepts.

First, as Kripke famously insisted, it appears, at least offhand, that our dispositions are *finite*, whereas the extension of many concepts—at least if we include their extensions across all possible worlds—is infinite.²² For example, the extension of the symbol ‘+’ is a function whose domain and range are both infinite. In consequence, the arguments and values of this function include numbers that are so vast that I have no dispositions whatsoever with respect to them. So how can my finite dispositions for using this symbol fix on a unique infinite function—the addition function—out of all the infinitely many such functions that there are?

Second, we all have dispositions to make *mistakes*. Our mistakes are not random; on the contrary, our mistakes are typically quite systematic. The non-normative dispositionalists face a dilemma here. On the one hand, they could be “holists” and claim that *all* our dispositions for using a concept are equally involved in determining the concept’s extension. Alternatively, they could reject holism and try to draw some distinction between those dispositions that fix the concept’s extension and those that do not. But according to Kripke, the first horn of the dilemma is hopeless. It is natural to assume that if one’s dispositions for using a concept determine the concept’s semantic value, then they must determine it in such a way that one’s disposition to form beliefs involving the concept in certain circumstances

²² Kripke, *Wittgenstein on Rules and Private Language*, 26.

guarantees that those beliefs are *true* in those circumstances. But then if *all* one's dispositions for using a given concept are equally involved in determining the concept's extension, then one cannot be disposed to form false beliefs at all. Indeed, if it is true, as some philosophers assume, that however one uses a concept, one must have been disposed to use it in that way, then as Kripke put it this "holistic" dispositional account amounts to "an equation of performance and correctness".²³ Thus, the first horn of the dilemma seems to lead to the absurd Protagorean claim that "all beliefs are true".

However, the second horn of the dilemma looks almost equally daunting. How exactly is the non-normative dispositionalist to distinguish between the extension-determining dispositions and all the other dispositions that we have with respect to a given concept? Perhaps if we could appeal to a normative distinction—say, between rational and irrational dispositions—we could answer this challenge. But it is quite unclear how to formulate this distinction in wholly non-normative terms, without simply presupposing what the extension of the concept is.

As challenging as these objections are, it is not obvious that non-normative dispositionalists cannot answer them. As for the first objection, it is not obvious that our dispositions really are "finite" in the way that the objection assumes. Until recently, many philosophers assumed that dispositions could be analysed by means of counterfactuals: according to this sort of analysis, to be disposed to ϕ in circumstances C is to be such that if one were in circumstances C , one would ϕ . But recently, this view has been shown to have many counterintuitive consequences.²⁴ Dispositions are more plausibly regarded as being closer to *ceteris paribus* laws, where laws may be regarded as nomic connections between properties. For an object to have a disposition

²³ Kripke, *Wittgenstein on Rules and Private Language*, 24.

²⁴ See for example Alexander Bird, "Dispositions and Antidotes", *Philosophical Quarterly* 48 (1998), 227-234; and Michael Fara, "Dispositions and Habituals", forthcoming in *Noûs*.

is for it to fall within the domain of the *ceteris paribus* law in question. If the law is a connection between properties, then the disposition is just as “infinite” as those properties are. So, for example, the relevant disposition with respect to the concept ‘+’ may be the disposition to accept an inference precisely in response to the fact that one has considered an inference of a certain *general form*. For example, in the case of the concept ‘+’, perhaps the relevant dispositions are as follows. First, one is disposed to respond to the state of considering a proposition of the form ‘ $n + 0 = n$ ’ by accepting that proposition. Second, one is disposed to respond to the state of considering any inference of the form ‘ $n + m = k$; therefore, $n + m' = k'$ ’ (where n' is the successor of n) by accepting that inference. Now, these “forms” are exemplified, not just by those propositions and inferences that one is capable of considering, but by infinitely many propositions and inferences.²⁵ It is, one might claim, these *general forms*, to which one’s disposition responds, that fix the extension of the concept.

As for the second objection, several different responses are possible. First, proponents of the “holistic” form of non-normative dispositionalism may reject Kripke’s assumption that if our dispositions to use a concept determine a concept’s semantic value, then our disposition to form a belief involving the concept in certain circumstances must guarantee that the belief is *true* in those circumstances. Holists influenced by Davidson might say that the correct interpretation of the concept is the one that “*maximizes*” the true beliefs that the thinker is disposed to form using that concept; but it does not have to guarantee that *every* belief that the thinker is disposed to form using that concept is true.²⁶

Alternatively, non-normative dispositionalists might reject holism, and claim

²⁵ For this response to the “finitude” objection, see Christopher Peacocke, *A Study of Concepts* (Cambridge, Mass.: MIT Press, 1992), chap. 5. It may genuinely be a certain property that one’s disposition is responding to, even if one’s dispositions are not perfect and do not respond to that property in many cases.

²⁶ See Davidson, “Radical Interpretation”, 137.

that it *is* possible to distinguish between (i) the dispositions that are essential to possessing the concept, and (ii) the dispositions that are not essential to possessing the concept, in wholly non-normative terms. For example, they might claim that the disposition that fixes the concept's extension is the disposition that is causally the most *basic* to one's use of the concept.²⁷ This notion of the disposition that is the "most basic" to one's use of a concept is hard to explicate precisely. But it may be that something like Jerry Fodor's notion of "asymmetric dependency" helps here.²⁸ The other dispositions, which do not fix the concept's extension, depend upon the extension-fixing dispositions in a way in which the latter dispositions do *not* depend upon the former. It may be then that the non-normative dispositionalist can give a wholly non-normative distinction between the extension-determining dispositions and the other dispositions that we have with respect to a given concept, without simply presupposing what the extension of the concept is. For these reasons, then, this Kripke-inspired argument also seems not to be a persuasive argument for the normativity of the intentional.

4. An argument for normative dispositionalism

In this final section, I shall give a sketch of a different sort of argument for the claim that the intentional is normative. I have argued above (in § 2) that some sort of dispositionalism must be true. I shall now argue that only a normative sort of dispositionalism (not non-normative dispositionalism) can give an adequate account of the nature of our intentional mental states.

²⁷ This seems to me a much more promising route than the route that Paul Boghossian assumes the dispositionalist will take, of focusing on the thinker's dispositions under "optimal conditions"; see Boghossian, "The Rule-Following Considerations", *Mind* 98 (1989), 507-549.

²⁸ See Fodor, *A Theory of Content and Other Essays* (Cambridge, Mass.: MIT Press, 1990), 90-92.

Suppose that one of your dispositions with respect to the use of a concept *F* is a disposition to use the concept in an *irrational* way, such as a disposition to engage in some fallacious form of reasoning. Then the holist would have to admit that it is partly in virtue of this disposition that you possess the concept *F*. This disposition is part of what it is about you that makes it the case that you possess the concept *F*. In effect, this disposition is an essential part of what *realizes* your possession of *F*.²⁹

However, it seems doubtful to me that one's possession of a concept can rest on an irrational disposition in this way. If this really could be the case, then there could be a thinker who possesses a concept at least partly in virtue of an irrational disposition of this sort. But then the thinker's possession of the concept would be at least partly founded in irrationality, and any use that the thinker makes of this concept would depend on or be underlain by this irrationality. It seems most doubtful whether this could be the case. The possession of a concept is a cognitive *power* or *ability*—not a cognitive defect or liability.³⁰

Thus, it seems to me that if one possesses a concept, the dispositions in virtue of which one possesses it must be rational dispositions. In general, one of the crucial features that mark out the dispositions that determine the identity of the concepts (and of the types of attitude) that figure in one's thoughts is that these dispositions must be rational dispositions.

More specifically, it may be that the disposition that determines the identity of a concept that figures in one's thoughts is whatever is causally the most *basic* of one's

²⁹ More precisely, if the holist is right, it is *essential* to possessing the concept *F* that one's possession of the concept *F* *supervenes* on one's total set of dispositions for using the concepts in one's repertoire; and the *minimal* supervenience base for possessing *F* that you actually instantiate is one that includes your having this irrational disposition. This seems to imply that your having this irrational disposition is part of what *realizes* your possession of this concept.

³⁰ Compare Gareth Evans' insistence that understanding a linguistic expression is a species of knowledge, and so cannot be based on false belief: "Truth is seamless: there can be no truth which it requires acceptance of a falsehood to appreciate"; see Evans, *The Varieties of Reference* (Oxford: Clarendon Press, 1982), 331.

rational dispositions with respect to the concept. Thus, if we are told that there is a concept *C* that figures in a thinker's thoughts, and we are also told (in terms that do not presuppose the identity of *C*) what is the most basic of the rational dispositions that the thinker has with respect to *C*, then this information will be enough to determine which concept *C* is.

This makes it plausible, it seems to me, that for each concept there is some *specific* rational disposition that is essential to possessing the concept. For example, in the case of the concept 'if', it seems impossible for there to be a thinker who possess the concept 'if' without having any disposition for rationally accepting *modus ponens* inferences (whereas it certainly seems possible for there to be superhuman thinkers who possess the concept 'if' but have no disposition to engage in any fallacious forms of reasoning involving this concept). If this is so, then any adequate account of a concept (or of an attitude-type) must mention this specific rational disposition, which is essential to possessing that concept or being capable of that attitude-type. If this disposition is essential to possessing the concept, then it will be *impossible* for any thinker to possess the concept without having this disposition. As I have argued, the dispositions in virtue of which we possess a concept must be rational dispositions. So, if this disposition is essential to possessing a concept, it must be impossible for this disposition not to be a rational disposition: it must be an *essentially rational disposition*.

Now some philosophers may think that this still does not show that an account of what is essential to a given concept will have to mention normative properties or relations (such as the properties or relations that are referred to by the term 'rational'). According to these philosophers, it is possible for such an account to single out the relevant essentially rational disposition by some purely non-normative feature—

where this other feature *guarantees* that the disposition in question is a rational disposition, but is not *equivalent* to its being a rational disposition.

However, there is a reason for thinking that these dispositions cannot in fact be specified without mentioning normative properties or relations.³¹ A disposition can be specified by means of a function from stimulus-conditions to response-conditions. To say that an object has the disposition is to say that it has some intrinsic features in virtue of which it is a *ceteris paribus* law that whenever the object is in one of these stimulus-conditions, it goes into the response-condition onto which the relevant function maps that stimulus-condition. As I shall now argue, a disposition cannot be an essentially rational disposition unless the stimulus-conditions to which it responds are conditions that essentially involve normative properties and relations.

We can illustrate this point by means of examples. It seems plausible that the relevant rational dispositions are all dispositions to form or revise one's mental states in some rational way. In a broad sense of the term, they are dispositions to engage in certain forms of *rational reasoning*. For example, one sort of rational reasoning might lead one from having a visual experience that presents an object in a certain distinctive way (in the absence of any positive reason to believe one's experiences to be unreliable), to one's forming a belief that predicates the concept '... is yellow' of the object in question. A disposition for this sort of reasoning might be essential to possessing the concept '... is yellow'. Other sorts of rational disposition might be dispositions to accept rational deductive or inductive inferences, or to engage in rational sorts of practical reasoning. For example, one simple sort of rational practical reasoning might be the kind that leads from having a rational belief of the form 'All things considered, I ought to ϕ ' to forming an intention to ϕ . It may be that a

³¹ When I speak of "specifying" a disposition, I do not mean an exhaustive complete account of how this disposition works. I simply mean "picking out" the disposition, by one of its essential or non-accidental features.

disposition for this sort of reasoning is essential to possessing the concept that is expressed by the relevant kind of ‘ought’.

As these examples suggest, there seems not to be any way to specify the relevant sorts of rational reasoning without mentioning normative properties or relations. The specification that I gave of the sort of reasoning that leads from visual experiences to beliefs of the form ‘ x is white’ included a proviso “in the absence of any *positive reason* to believe that one’s experiences are unreliable”. But in specifying this sort of reasoning in this way, I mentioned the normative relation of *having a reason to believe* something. Similarly, my specification of the sort of reasoning that leads to an intention to ϕ is restricted to cases in which one has a *rational* belief that all things considered, one ought to ϕ ; and in specifying this type of reasoning in this way, I have mentioned the normative relation of *rationally believing* something.

The reason for specifying these sorts of reasoning in this normative way is clear. These sorts of reasoning, like most if not all others, are *defeasible*. That is, even if one is in the input mental states for this sort of reasoning, certain defeating conditions must be absent if it is to be possible to form a rational belief or intention by means of this form of reasoning. But the nature of defeating conditions is precisely that they are those conditions that make it *irrational* for one to regard it as reliable in the circumstances to form the belief or intention in question in response to these input mental states. There seems to be no way of specifying these defeating conditions without mentioning normative properties or relations.

For a sort of reasoning to be essentially rational, then, it seems that its input conditions must include the absence of defeating conditions *as such*. It would not be enough if its input conditions included only the absence of certain defeating conditions, but not the absence of defeating conditions as such. Then there would be

cases in which this sort of reasoning would be irrational: in these cases, this sort of reasoning would involve treating certain conditions as if they constituted a reason for a certain revision to one's beliefs or other attitudes when in fact, on account of the presence of defeating conditions, they do not. Equally, it would also not be enough if this sort of reasoning were markedly oversensitive, so that its input conditions included, not just the absence of defeating conditions, but also the absence of various additional factors that were in fact quite irrelevant to the rationality of the sort of reasoning in question. Then this sort of reasoning would be irrational in a different way: it would involve treating certain conditions as relevant when in fact they are not relevant at all.

For this reason, it seems that the relevant forms of rational reasoning cannot be specified without mentioning the absence of defeating conditions—which in effect involves mentioning certain normative properties and relations. But as I have argued, the dispositions that are essential to possessing concepts, and to being capable of the various types of attitude, must be dispositions to engage in various essentially rational forms of reasoning. So it seems that the dispositions that are essential to possessing concepts (and to being capable of attitude-types) cannot be specified without mentioning normative properties or relations.

Some philosophers might suggest that one could still be disposed to reason in accordance with these essentially rational forms of reasoning, even if the normative properties or relations do not have to be mentioned in specifying these dispositions themselves. For example, one might have a host of *separate* dispositions, such that the net effect of all these separate dispositions is that one forms and revises one's beliefs and other attitudes in just the same way as one would if one had a disposition that actually responds to the normative facts themselves. However, if all these dispositions

really are quite separate and causally independent of each other, then it seems that it could just be a *fluke* that the net effect of all these dispositions is that one forms and revises one's attitudes in a rational way. But in that case, it actually seems to me quite doubtful whether the manifestation of these dispositions can really count as rational thinking at all. Indeed, this case seems in effect to be the case that was originally described by Ned Block and has since come to be known as the "Blockhead", which seems not be engaged in rational reasoning at all.³²

So it seems most plausible that if there are any dispositions that are essential to possessing concepts (and to being capable of attitude-types), these dispositions must respond precisely to normative facts; hence these dispositions cannot be specified without mentioning such normative properties or relations.

Some other philosophers might insist that it is still bound to be possible in principle, by using some highly complex psychological predicates, to specify these rational forms of reasoning in wholly non-normative terms. But I do not have to deny this.³³ I have said that we cannot specify these rational dispositions *without mentioning normative properties or relations*; I have *not* said that we cannot specify these rational dispositions *without using normative terms*. It may well be possible, in principle, to mention these normative properties or relations by means of some complex psychological predicates. Even if these normative properties or relations can be specified in purely psychological terms, it would still be an interesting and surprising fact that any adequate account of the nature of intentional mental states has

³² Block's original example was of a machine that passes a "Turing Test" simply by means of a huge "look-up table". See Block, "Troubles with Functionalism", *Minnesota Studies in the Philosophy of Science* IX, ed. C. Wade Savage (Minneapolis: University of Minnesota Press, 1978), pp. 294–96.

³³ I do not *have* to deny this; but there may still be reasons for denying it. For an argument for the claim that the conditions of rational belief and rational choice are "uncodifiable", see T. W. Child, "Anomalism, Uncodifiability, and Psychophysical Relations", *Philosophical Review* 102(1993), 215–45.

to mention these normative properties or relations—so long as these normative properties or relations can only be specified by means of highly *complex* psychological terms. This is because it would still be an interesting and surprising fact if this particular complex psychological term has to appear in any adequate account of the nature of intentional mental states.

It certainly seems that if there are any purely psychological predicates that refer to the absence of defeating conditions, they will be immensely complex. Defeating conditions, after all, can come from *anywhere*—from experience, from memory, from the testimony of others, or even from elaborate deductive reasoning or scientific theorizing. So the only *simple* way of specifying what these conditions all have in common is in normative terms—as conditions that make it irrational to reason in certain ways. Even if there is also some more complex way of specifying these defeating conditions in non-normative terms, my argument would not be undermined if the only simple way to specify the dispositions that are essential to possessing certain concepts, or to being capable of certain attitude-types, is by using normative terms.

As I have argued, however, there must be some account that explains how the thinker's dispositions determine which concepts and attitude-types figure in his thoughts. As I have also argued, the prospects for a purely non-normative dispositionalism are dim. So since some sort of dispositionalism must be true, then (at least so long as it seems that a normative sort of dispositionalism could avoid the problems that other theories face) we have a good reason to accept the claim that I have been considering here—the claim that the intentional is normative.