

## Normativism Defended

Ralph Wedgwood

My aim in this chapter is to defend the claim that “the intentional is normative” against a number of objections, including those that Georges Rey has presented in his contribution to this volume. In the first section of this chapter, I shall outline a specific version of this claim; and in the second section, I shall give a quick sketch of the principal argument that I have used to support this claim, and briefly comment on Rey’s criticisms of this argument.

In the remaining sections of this chapter, I shall try to answer the main objections that have been raised against this claim. Broadly speaking, there are two main objections here (both of these objections are mentioned by Rey, but in one form or another they have been raised by many other philosophers as well). First, it may seem that the claim that “the intentional is normative” is just hopelessly *Panglossian*: this claim seems to imply that rationality is in some way essential to the capacity for intentional mental states as such; but doesn’t this just wilfully ignore all the mountains of evidence that we have for the sheer ubiquity and pervasiveness of human irrationality? Secondly, the claim that intentional mental states are essentially normative seems to be intended as a purely *philosophical, non-empirical* account of the nature of these mental states: but why should we think that purely philosophical reflection can tell us *anything* interesting about the nature of the mind; shouldn’t we look to empirical psychology to enlighten us about such matters? I shall take these two objections in turn.

## 1 A version of the claim that “the intentional is normative”

In general, the claim that “the intentional is normative” is the claim that any adequate account of the nature of intentional mental states must employ normative terms (or at least must mention the properties and relations that these normative terms stand for). But different versions of this claim will give very different accounts of the exact role that normative terms must play in adequate accounts of the nature of intentional mental states.<sup>1</sup> In this section, I shall briefly outline a specific version of this claim; it is this version of the claim, and not any others, that I shall try to defend here.

By an “intentional mental state”, I mean a mental state that is *about* something, or in other words, has some *content*. For example, beliefs, intentions, desires, suppositions, and the acceptance of inferences, all seem to be intentional mental states in this sense.

It is clear on reflection that there are infinitely many types of intentional mental state. (For example, for every natural number  $n$ , there is the belief that there are at least  $n$  atoms in the universe.) Thus, there is no possibility of giving an account of all these intentional mental states one by one. Instead, any account of these mental states would have to proceed by giving an account of the *basic components* out of which these mental states are built up. I shall assume here that intentional states are built up out of the following components. First, they have some sort of *content*, which I shall assume here to be composed of *concepts*. Thus, the content of the belief that snow is white (and of the hope or the fear that snow is white) is the content ‘snow is white’, which is composed by predicating the concept ‘white’ of the concept ‘snow’. Second, these intentional mental states involve a certain *mental relation* to that content; in many cases this mental relation could be called an *attitude* (such as the attitude of belief or hope or fear) towards that content.

As I am using the terms here, the term ‘correct’ and the term ‘rational’ both express normative concepts. That is, to say that an attitude is “incorrect” is to say that it is an attitude that — in a certain way — one *ought not* to have; and to say that a way of thinking is “irrational” is to say that it is a way of thinking that — in a somewhat different way — one *ought not* to engage in. The main difference between the concept of a “correct” attitude and the concept of a “rational” way of thinking is the following. Whether or not an attitude is *correct* is typically determined by some relation between that attitude and the external world. By contrast, whether or not a way of thinking is *rational*, for a thinker at a given time, is determined purely by the intrinsic features of that way of thinking and its relation to the antecedent mental states that the thinker has at that time.<sup>2</sup>

According to the version of the claim that “the intentional is normative” that I am outlining here, the nature of every concept is given both by the principle that specifies when beliefs involving that concept are correct, and also by certain basic principles of rationality that specify certain ways of using the concept as rational (or specify certain other ways of using the concept as irrational). Assuming that a belief is correct if and only if the content of the belief is true, the conditions under which beliefs involving a concept are correct would in effect define the concept’s *semantic value* — the contribution that the concept makes to the truth conditions of contents in which it appears. On the other hand, the basic principles of rationality that feature in the account of the nature of the concept would determine what we could call — to adapt a term from Frege (1892, 25) — the concept’s *cognitive significance*.

Thus, for example, the nature of logical concepts, like ‘if’ and ‘not’, might be given both by their semantic values — their contribution to the truth conditions of contents in which they appear — and also by the basic principle that it is rational (at least in the absence of any special reason for doubt) to accept instances of certain fundamental rules of inference

for these logical concepts. The nature of the concept ‘yellow’ might be given both by the concept’s semantic value — the property of yellowness that the concept stands for — and also by the basic principle that it is rational to make a judgment applying this concept to some perceptually presented object if one has a visual experience that represents that object in a certain distinctive way (and one has no strong special reason for doubting the reliability of one’s experiences in the circumstances).

According to the version of the idea that “the intentional is normative” that I am outlining here, it is not just *concepts* whose nature is given by the normative principles that apply to them; the same is true of the various *types of attitude* as well. Thus, for example, perhaps the nature of the attitude of *belief* is given both by the principle that a belief is correct if and only if the content of the belief is true, and by some related principle that specifies certain ways of revising beliefs as rational (or specifies certain other ways of revising beliefs as irrational). This approach could also be applied to other types of attitude. Thus, perhaps the nature of the attitude of *admiration* is given both by the principle that it is correct to admire something if and only if the object of one’s admiration really is admirable, and also by some related principle that specifies certain ways of coming to have attitudes of admiration, or ways of responding to such attitudes of admiration, as rational.

This then, roughly, is the core of this normative theory of intentional mental states — an account of the nature of the various concepts, and of the various attitudes that we can have towards contents that are composed of such concepts, in terms of some of the normative principles that apply to mental states involving these concepts or attitudes. But this core will need to be surrounded by a shell that explains what it is for a *thinker* to *possess* those concepts, or to *be capable of* those attitudes. If this is the correct account of the nature of the concepts ‘or’ and ‘yellow’, and of the attitudes of belief or admiration, what has to be true of

a thinker if she is to possess these concepts, or to be capable of having these attitudes?

One plausible answer to this question is that the thinker must have some *disposition* that amounts to an appropriate sort of sensitivity to the normative principles that give the nature of the concept or type of attitude in question.<sup>3</sup> In most cases, it will be more plausible to suppose that the thinker must have a disposition to conform to the principles of *rationality* that feature in the correct account of the nature of the relevant concept or attitude-type than that she must have a disposition to conform to the corresponding principles of *correctness*. As I explained above, whether a mental state is correct is typically determined by the relation between that mental state and the external world; and one's dispositions to have many sorts of mental states do not respond directly to the external world, but only to one's antecedent mental states.

Specifically, then, the sort of disposition that a thinker must have, if she is to possess a given concept, is a disposition to think in ways that the relevant basic principle of rationality specifies as rational. Such principles typically take the form of specifying some set of antecedent mental states such that — according to this principle — it is rational to respond to being in those antecedent mental states by forming a certain further mental state. Then, the disposition that one must have, in order to possess the concept, will be a disposition to respond to one's actually being in those antecedent mental states by forming that further mental state. So, for example, the thinker must be disposed to form a judgment applying the concept 'yellow' to a perceptually presented object whenever she has a visual experience that represents the object in the relevant way (at least if the question of the object's colour arises, and the thinker has no strong special reason for doubting the reliability of her experiences in the circumstances). According to this version of the claim that the "intentional is normative", this disposition is essential to possessing the concept 'yellow'.

Since this version of the claim that “the intentional is normative” crucially involves the notion of a *disposition*, it will be useful to say something about how I am understanding dispositions here. In general, a disposition can be specified by means of a function from stimulus conditions to corresponding response conditions. (For example, the function that specifies the disposition of fragility might be defined by means of a function from any stimulus condition like *being struck at time t* to the corresponding response condition *breaking shortly after t*.) To say that you have a certain disposition is not necessarily to say that if you were *ever* to be in any of the relevant stimulus conditions, you would *invariably* go into the response condition onto which the function in question maps that stimulus condition. It is to say that *ceteris paribus*, or in any *normal* case, when you are in one of these stimulus conditions, you will also go into the corresponding response condition. In effect, each of these “stimulus conditions” and “response conditions” is a property that you might have; and to say that you have the disposition is in effect to say that you fall under a *ceteris paribus* law that connects one domain of properties (the stimulus conditions) with another domain of properties (the response conditions) according to the function that specifies the disposition.<sup>4</sup>

It follows from this conception of dispositions that even if you are in one of the relevant stimulus conditions, you may still fail to manifest the disposition — that is, you may fail to go into the corresponding response condition. This can happen when the case in question fails to be normal, and *cetera* fail to be *paria*. When this happens, there will in principle be some explanation of why the case in question failed to count as normal — for example, an explanation of the interfering factors that blocked or inhibited the manifestation of the disposition in this case. It is only when all such interfering factors are absent, and the case is normal in the relevant sense, that the disposition must be manifested.

It seems that the cases that count as “normal” for one domain of stimulus conditions may not be exactly the same as those that count as “normal” for another domain of stimulus conditions. Indeed, it may even be that certain cases that do not count as normal for one domain of stimulus conditions count as normal for a certain narrower domain of stimulus conditions that is properly included in the first domain. This makes it possible for one to have a disposition with respect to a certain domain of stimulus conditions even if one does *not* have the corresponding disposition with respect to a certain narrower domain of stimulus conditions that is properly included in the first domain. For example, you might have a general disposition to accept *modus ponens* instances as such, even if you do *not* have a disposition to accept a certain special subcategory of *modus ponens* instances. Relative to the whole domain of *modus ponens* instances, none of the cases in which you consciously consider the inferences in this special subcategory counts as normal; and so the fact that you do not accept these special *modus ponens* instances in these cases does not count against the claim that you have a general disposition to accept *modus ponens* instances as such. However, relative to the narrower domain that consists of this special subcategory of *modus ponens* inferences, some of these cases *do* count as normal cases; and so the fact that you do not accept these inferences in these cases does count against the claim that you have a disposition to accept this special subcategory of *modus ponens* inferences.

Another important feature of dispositions, as I understand them, is that since dispositions are essentially like *ceteris paribus* laws, they will also be subject to essentially the same restrictions that apply to *ceteris paribus* laws. In particular, since any such law is in effect a connection between domains of properties, it seems plausible to require that these domains of properties must in a sense be “*natural* property domains”, rather than gruesomely gerrymandered domains of properties.<sup>5</sup>

For example, suppose that you have a general disposition to accept instances of *modus ponens*. To say that you have this disposition is to say that *ceteris paribus*, or in any *normal* case, when you consciously consider an inference that is an instance of *modus ponens*, you accept that inference. There are two relevant domains of properties here: first, the domain of properties that for every *modus ponens* instance  $S$ , includes the property *consciously considering*  $S$ ; and second, the domain of properties that for every such *modus ponens* instance  $S$ , includes the property *accepting*  $S$ . These domains of properties seem to be “natural property domains”. They certainly seem much more natural than a domain that includes the property of consciously considering  $S_1$ , for every member  $S_1$  of a completely random assortment of inferences, or the property of considering  $S_2$  for all inferences  $S_2$  that are instances of some very strange and gerrymandered “form”, which coincides with *modus ponens* in all cases that any human being has actually considered, but diverges wildly from *modus ponens* in some of less easily accessible cases. Those unnatural domains of properties are less apt to enter into *ceteris paribus* laws, and so they are less likely to be the domains of stimulus conditions that your disposition is responding to. So, when you manifest this disposition, it is plausible to say that you accept the inference in question precisely *because* it is an instance of *modus ponens* — not because it is an instance of the strange and gerrymandered “form” that coincides with *modus ponens* in some cases but diverges wildly from *modus ponens* in others.<sup>6</sup>

A final feature of dispositions that I should like to highlight here is that something’s possession of a disposition may in a sense be “realized in” its possession of various other dispositions, or in the possession of various other dispositions by one or more of its parts. There is no requirement that the dispositions that realize a given disposition  $D$  should all be of the same kind, nor that these realizing dispositions should not play a role in realizing other

dispositions besides *D*. For example, most of us are disposed to become annoyed when we are insulted. This disposition is itself realized in the following dispositions: our dispositions to perceive communicative actions of various kinds (such as spoken or written utterances); our disposition to understand these communicative actions correctly, and so to recognize insults for what they are; and our disposition to respond to the belief that one has been insulted with the emotion of annoyance. These realizing dispositions are themselves dispositions of many different kinds — including perceptual, cognitive, and emotional dispositions — and each of them is involved in the realization of many other dispositions besides our disposition to become annoyed with we are insulted. But this does not count against the claim that we have the general disposition to be annoyed when we are insulted.

## **2 An argument for the claim that “the intentional is normative”**

How might one argue for this claim about the nature of intentional mental states? I shall now give a quick sketch of an argument for this claim.<sup>7</sup>

The first step in this argument is to defend a sort of *dispositionalism* about concepts and types of attitude. What makes it the case that a concept that you possess in your conceptual repertoire is the concept ‘yellow’? According to this sort of dispositionalism, it is in virtue of some of your *dispositions for reasoning* with a certain concept in your repertoire that that concept counts as the concept ‘yellow’. If this is right, then it is also plausible that these dispositions are *essential* to possessing the concept, so that two thinkers will count as sharing this concept only if they also share these dispositions for using the concept in reasoning.<sup>8</sup> We could also give a similar answer to the parallel question about attitude-types: What makes it the case that an attitude-type in your repertoire is the attitude of intention?

According to this sort of dispositionalism, it is in virtue of some of your dispositions with respect to an attitude-type in your repertoire that this attitude-type counts as the attitude of intention; and this disposition is essential to being capable of the attitude of intention.

Now there are reasons for thinking that the dispositions that are essential to possessing a concept (or to being capable of an attitude-type) must be *rational dispositions*. One of these reasons is that it seems that any concepts that you have could be shared by a perfectly rational being who had no irrational dispositions at all. (For example, the perfectly rational being would need to possess these concepts in order to ascribe attitudes to you accurately, and to diagnose the various confusions and irrationalities that mar your thinking.)

However, if it is in virtue of some of her dispositions that the perfectly rational being possesses these concepts, and the perfectly rational being has no irrational dispositions at all, then it must be in virtue of some rational disposition that she possesses the concept.<sup>9</sup> And if the dispositions in virtue of which a thinker possesses a concept are essential to possessing the concept, then it must also be in virtue of this rational disposition that *you* possess the concept in question. Moreover, this picture of what it is to possess concepts seems intuitively plausible. For example, take the concept ‘if’. It is plausible that it is at least in part in virtue of your rational disposition to accept instances of *modus ponens* that you possess this concept. It wouldn’t be correct to interpret any concept in your repertoire as the concept ‘if’ unless you had this disposition. On the other hand, it seems that you could still possess the concept if you lost all your irrational dispositions with respect to the concept. So it seems that the dispositions that are essential to possessing a concept must all be rational dispositions.

Rey has a number of objections to this argument, which he labels my “asymmetry argument” (this volume, § 3.3). First, he asks whether it really is impossible for one’s possession of a concept to rest essentially on some irrational disposition. Perhaps there are

some intrinsically irrational or incoherent or problematic concepts? Possible examples of such intrinsically irrational concepts might be the concepts of “free will”, the traditional concept of the “soul”, and the concept of an “angel”. I shall answer this objection in the next section, where I shall argue that such intrinsically irrational concepts are indeed impossible.

Secondly, Rey objects that the sort of intuition that my argument rests on “doesn’t entail any claims about the believer’s adherence to norms *generally*.” Here, however, he seems to be misunderstanding what I am arguing for. I am not arguing that it is necessary that all believers must adhere to “norms *generally*”. On the contrary, all that I am arguing for is that anyone who possesses a concept must have some disposition to reason in accordance with the *basic* principle of rationality that features in the correct account of the nature of the concept. For example, it is quite compatible with my argument that thinkers who possess this concept may have no disposition at all to comply with many of the valid rules of inference for the concept ‘if’, such as *modus tollens* or contraposition. My argument only implies that all such thinkers must have at least some disposition to reason in accordance with the *basic* principle of rationality governing the concept: in the case of ‘if’, presumably, they must have some disposition to accept *modus ponens* inferences. According to my argument, that is all that is required for possession of the concept ‘if’, not a general adherence to *all* norms that apply to the use of this concept.

Finally, Rey presents the following objection:

Lastly, the same “asymmetry” conception of meaning to which Wedgwood is appealing is advocated (in different ways) by both Jerry Fodor ... and Paul Horwich ..., the meaning-constitutive “laws” or “uses” being the ones on which all other uses asymmetrically depend. Why does Wedgwood insist, unlike Fodor and Horwich, that the basic ones are normative?

Rey is quite right that my account resembles the accounts of Fodor and Horwich by appealing to a basic disposition to use a concept on which all other uses of the concept asymmetrically depend. Of course, I have also argued for something that neither Fodor nor Horwich has argued for — that the “basic” or “meaning-constitutive” dispositions with respect to a concept must all be “rational dispositions”. Still, perhaps these rational dispositions can be adequately specified in wholly non-normative terms? If so, then we would not be able to infer that normative terms must be mentioned in any adequate account of what it is to possess the concept; and so we would not be able to use this argument to support the claim that the intentional is normative.

It is precisely in order to answer this objection that I give the supplementary argument that Rey has labelled my “defeasibility argument” (this volume, § 3.4). Rational dispositions are dispositions to engage in rational forms of reasoning. But whenever we specify a form of reasoning in wholly non-normative terms, it will turn out that the form of reasoning in question is *defeasible*, and so circumstances can arise in which it is not in fact rational to engage in that form of reasoning. Even forms of reasoning that involve basing a belief on utterly conclusive grounds — such as forming a belief in a mathematical theorem on the basis of a genuine proof of the theorem — can be defeated, if a sufficiently large number of experts testify, with apparent sincerity and confidence, to the incorrectness of one’s belief.

So the only way to specify a form of reasoning that will always count as rational is for one’s specification to include a proviso that requires the absence of such defeating conditions. For example, the most plausible way of specifying the basic rational form of reasoning with the concept ‘yellow’ might be as the form of reasoning that leads from one’s having a visual experience that presents an object in a certain distinctive way to one’s forming a belief that predicates the concept ‘yellow’ of the object in question — *provided*

*that* one is not in any defeating conditions that give one a strong special reason to doubt the reliability of one's colour experience in the circumstances. But the very notion of a "defeating condition" is a normative notion. So, contrary to Rey's objection to my "asymmetry argument", the relevant rational dispositions cannot be specified except in partly normative terms.

Against this "defeasibility argument", Rey objects as follows:

The question for psychology is whether people could possess a particular content and not be disposed to apply it appropriately, or to respond to defeaters. Of course they could. A person might have all sorts of patently *bad* reasons to withhold a concept — superstitions, silly theories, blind prejudice ...; or he might apply it in the face of genuine defeaters, failing simply to appreciate them, or mistakenly thinking they in turn have been defeated.

Again, Rey seems to be misinterpreting me as claiming that anyone who possesses a concept must have a *general* disposition "to apply it appropriately", and to respond to defeaters to *any* way of "applying" the concept. As I have already explained, I am claiming only that anyone who possesses a concept must have some disposition to reason in accordance with the *basic* principle of rationality that features in the account of the nature of the concept (in the case of the concept 'if', this will presumably be a disposition to accept *modus ponens* inferences — not a disposition to accept *all* valid inferences involving 'if').

What my "defeasibility argument" adds is the following point: since the disposition that is essential to possessing the concept must be a rational disposition, it must be a disposition that tends not to be manifested in the presence of defeaters. This is not to say that this disposition will *never* be manifested in the presence of *any* defeaters — only that there is a range of defeaters in the presence of which the disposition will not be manifested. In saying that possessing the concept requires having a disposition to use the concept in a certain

basically rational way, I need not claim that this disposition must be *perfectly* rational; I need only claim that this disposition must to a greater or lesser degree *approximate* to such perfect rationality.

However, I am still claiming that every thinker who possesses a given concept must have a rational disposition towards a certain basic form of rational reasoning involving the concept. Rey would presumably think that it is still far too Panglossian a view to suppose that all thinkers who possess the concept must have a rational disposition of this kind. Surely this view is incompatible with all the mountains of evidence that we have of how thoroughly and pervasively irrational human beings are? This is the objection that I shall try to answer in the next section.

### **3 A hopelessly Panglossian picture of the mind?**

On my picture, the possession of a concept is always a rational power or ability, and never a defect or liability. But some philosophers have toyed with the idea that some concepts — perhaps especially the concepts that are routinely found problematic by philosophers — might be essentially incoherent concepts; that is, they might be concepts that rest on irrational dispositions of some kind.<sup>10</sup>

In fact, however, it does not seem plausible to postulate incoherent concepts of this sort. The reason for this is similar to the reason against postulating concepts that essentially depend on mistaken beliefs. I am assuming here that the content of thoughts is composed of concepts; so when I deny the existence of witches or unicorns, I am surely using the very same concepts that were used by the medieval thinkers who believed in witches and unicorns. If I were not using the very same concepts, I would not really be *disagreeing* with these

medieval thinkers, in the sense of denying the very thought that they affirmed. But of course I do not make the mistakes that characterized medieval thinking about witches and unicorns. So making these mistakes does not seem to be necessary in order to possess these concepts.

It seems to me that an essentially similar point applies to the concepts that Rey mentions, such as the concepts of an “angel” or of the “soul”. Again, suppose that there were a perfectly rational being, who had no irrational dispositions of any kind. This perfectly rational being would have to possess these concepts in order to ascribe beliefs involving these concepts (for example, by thinking to herself ‘Many human beings believe that they have an immortal soul’), and also in order to reject the mistaken beliefs that many human beings have involving these concepts (for example, by thinking to herself ‘Human beings do not have immortal souls’). It seems possible that there could be a perfectly rational being who could do this. So it does not seem that it is *essential* to possessing these concepts that one should have any irrational dispositions. But it seems that any of the “problematic concepts” that appear to be deeply entwined in confused or irrational thoughts could be shared by a perfectly rational being who was capable of diagnosing and rejecting our mistaken and irrational thoughts. So in fact, these problematic concepts do not seem to be essentially irrational concepts. The concepts themselves are perfectly innocent: it is our *use* of those concepts that is defective.

However, even if I am right that there are no such intrinsically irrational concepts, it is an undoubted fact that our thought is riddled with fallacious thinking of many kinds. How can the claim that the intentional is essentially normative be reconciled with this undoubted fact?

First, my account only implies that any thinker who possesses a concept must have a *disposition* to conform to the basic principle of rationality that features in the correct account

of the nature of the concept. But as I have already explained, the claim that a thinker has a disposition of this kind does not imply that this disposition will be manifested in every possible case. It is only in the relevantly *normal* cases, when *cetera* are *paria*, that the disposition is bound to be manifested.

Thus, this claim — that everyone who is capable of having a certain type of mental state must have a disposition towards a certain basic sort of rational thinking involving mental states of that type — does not imply that these rational dispositions are always manifested. It implies only that when we do not conform to these very basic requirements of rationality, the situation was abnormal (*cetera* were not *paria*), and so there must be some explanation of why the disposition was not manifested in this case — perhaps an explanation that appeals to certain interfering factors that blocked or inhibited the manifestation of the disposition in this case. Indeed, it seems possible, in extreme circumstances, for there to be an agent, or even an entire community of agents, for whom the circumstances are always abnormal, so that these dispositions are in fact *never* manifested.

Secondly, in addition to the ways in which one's dispositions towards these basic kinds of rational thinking can fail to be manifested in abnormal cases of various kinds, one may also have many *other* mental dispositions, which may be dispositions that it is irrational to manifest. This too is quite compatible with its being the case that in general, one *also* has dispositions towards certain basic kinds of rational thinking.

These two points must be taken together with the point (which I emphasized in the previous section, in discussing Rey's objections to my "asymmetry argument") that the claim that I am defending does not imply that all thinkers must have a disposition to adhere to *all* the norms of rationality that apply to them. Instead, it only implies that they must have a disposition to adhere to *those particularly basic* norms of rationality that feature in the

correct accounts of the nature of the various concepts that they possess (and of the various attitude-types of which they are capable).

Taken together, these points help to show how my claim can be reconciled with the well-known evidence about how strikingly bad most humans are at even fairly elementary deductive reasoning.<sup>11</sup> In recent empirical studies of conditional reasoning,<sup>12</sup> for example, only 72% of subjects accepted instances of *modus tollens*. But this is compatible with my claims, since it is plausible that it is *modus ponens*, and not *modus tollens*, that is the basic form of rational inference for the concept ‘if’. In addition to failing to accept certain valid forms of inference, many of the subjects in these experiments endorsed certain fallacious forms of inference: as many as 63% committed the fallacy of affirming the consequent, and 55% committed the fallacy of denying the antecedent. This too is compatible with my claims, since I have only claimed that we must have certain basic rational dispositions, not that we cannot *also* have many irrational dispositions as well. These studies do not undermine my claims about what is involved in possessing concepts, since 97% of the subjects in these experiments accepted instances of *modus ponens*; and in the case of the 3% who did not accept instances of *modus ponens*, we can either appeal to some slippage between language and thought (so that they were not genuinely entertaining the relevant inference in thought), or else to some other interfering factor that inhibited the manifestation of their general disposition to accept *modus ponens* inferences in those cases.<sup>13</sup>

At one point, Rey (2002, 107–08) raises the issue of what Rosalind Hursthouse (1991) has called “arational actions” (like jumping for joy, or rumpling one’s lover’s hair). But my version of the claim that the intentional is normative has no problem with these actions. These arational actions are indeed counterexamples to a certain well-known theory of action, according to which every action is done for a reason, and this reason consists of a

desire for some goal, together with a belief that the action in question is a means towards that goal.<sup>14</sup> But my version of the claim that “the intentional is normative” does not imply that every single action is the manifestation of a rational disposition to act for reasons. At most, my version of this claim implies that if one is capable of making decisions about what to do at all, one must have at least some disposition to take account of reasons for action in deciding what to do, and some disposition to try to carry out one’s decisions. But there certainly can be abnormal cases in which one fails to manifest this disposition; and one may also have many other irrational or non-rational dispositions, which will often influence how one acts. Moreover, as Joseph Raz (1999) has pointed out, these arational actions do not typically involve failing to take account of *all* reasons for action: in performing one of these arational actions, one is usually aware, as part of one’s background beliefs, that there are no strong reasons against the action in question (one would not jump for joy on the edge of a precipice, for example). So these examples pose no problems for the claims that I am defending here.

Recently, Timothy Williamson (2003 and 2006) has emphasized the example of philosophers whose revisionary views lead them to reject some of the laws of logic. For example, Vann McGee (1985) believes that there are counterexamples to *modus ponens*; and a philosopher who agreed with P. F. Strawson (1952) in holding that the natural-language universal quantifier has existential import might (if she also denied the existence of sets) insist that the sentence ‘All sets are sets’ is neither true nor false. As Williamson argues, it is highly implausible to say that these philosophers do not understand the sentences involved, or express different concepts with these sentences from those that we would be expressing by using them. But, Williamson maintains, these revisionary philosophers are not even *disposed* to accept the relevant instances of these logical laws; so how can a disposition to accept these

logical laws be constitutive of possessing the concepts involved?

It is not clear exactly how to interpret these revisionary philosophers. Are these philosophers are being *irrational* (albeit perhaps in a quite blameless way)? Or is it actually *rational* for them, given the unusual philosophical reflections that they have gone through, to take the view that they do? Whichever interpretation is correct, these examples create no problems for my claims. If it is *rational* for these revisionary philosophers to take the view that they do, then this must be because their philosophical reflections constitute a rather unusual *defeater* for the rationality of accepting the relevant instance of the logical law in question. In that case, these philosophers plainly are *not* counterexamples to my claims. The principle of rationality that features in the account of the nature of the concept ‘if’ is, strictly speaking, the principle that it is rational to accept *modus ponens* inferences *unless one has sufficiently strong defeating reasons not to*; and, on this interpretation, Vann McGee *has* a disposition to comply this with this principle.

On the other hand, if it is *irrational* for these philosophers to reject these instances of these logical laws, then we must say that the case in question is abnormal, perhaps because some interfering factor is inhibiting the manifestation of their rational disposition to accept instances of these laws. On this view, Vann McGee has a general disposition to accept *modus ponens* inferences, and the follower of Strawson has a general disposition to accept propositions of the form ‘All *Fs* are *Fs*’; their misleading philosophical reflections, and the false philosophical beliefs that they have acquired as a result, are operating as an interfering factor that inhibits the manifestation of this disposition in these cases.

Williamson (2006, 15–20) considers this second response to his objection, and argues that it is not correct to ascribe these dispositions to these revisionary philosophers. There are, he suggests, two ways of taking the claim that they have these dispositions — as a “*personal*

level account” and as a “*sub-personal* level account”. According to Williamson (2006, 16–17), if the claim that these philosophers have these dispositions is a “personal level account”, it is refuted by the fact these philosophers would persist in their rejections of the relevant instance of *modus ponens*, and of ‘All sets are sets’, even after the most careful and self-conscious reflection.

However, Williamson’s argument is not a conclusive refutation of the “personal level” version of the claim that these revisionary philosophers have these dispositions. As it has recently been persuasively argued, disposition ascriptions are not in general equivalent to counterfactuals.<sup>15</sup> So the fact that the thinker would not manifest the disposition in this specific case is not enough to show that the thinker lacks the disposition altogether.

Moreover, as I explained in § 1.1, even if we concede that Vann McGee does *not* have a *specific* disposition to accept *this particular instance of modus ponens*, it could still be the case that he does have a *general* disposition to accept *modus ponens* inferences as such. Relative to the wider domain of properties that includes the property of consciously considering *S* for every *modus ponens* instance *S*, the cases in which McGee considers these complicated instances of *modus ponens* in the course of philosophical reflections of the sort that he is pursuing do not count as “normal” cases in the relevant sense; and so the fact that McGee does not accept these special instances of *modus ponens* does not count against the claim that he has a disposition to accept *modus ponens* inferences in general. But we could still consistently concede that relative to the much narrower domain of properties that includes only the property of consciously considering this particular *modus ponens* inference, the case in question does count as normal, and so the fact that he does not accept the inference in this case does count against the claim that he has the more specific disposition.

It surely is highly plausible that Vann McGee does have this general disposition to

accept *modus ponens* inferences. Like everyone else, he accepts the vast majority of such inferences that he consciously considers; in the cases in which he does not, there seems to be something distinctly abnormal going on; and there seems to be no more “natural” way of capturing all the cases in which he does accept such inferences in a general *ceteris paribus* law than by classifying them all as involving his consciously considering, and then accepting, an instance of *modus ponens*. Moreover, this disposition is clearly at the “personal level”, in the sense that its manifestations consist in conscious events in the mental life of the whole person, rather than in the unconscious cognitive processes that realize our conscious thinking.

Williamson (2006, 17–20) goes on to develop an objection to the second “way of filling out the dispositional story” — the interpretation of this dispositional story as a “sub-personal level account”. Fundamentally, his objection is that philosophers should not attempt to develop such sub-personal level accounts without consulting the relevant empirical evidence, which may well not tell in favour of this account. This is only a specific instance of a more general objection that is often raised against the claim that “the intentional is normative”. I shall consider this objection in the next section, which will be the last section of this chapter.

#### **4 Psychology, *a priori* and empirical**

Some of the philosophers who have claimed that the intentional is normative have gone on to make some further claims to the effect that empirical psychology will never tell us anything interesting about the mind. Rey (2002) has criticized these further claims effectively; and I have no intention whatsoever of defending any such further claims myself.

Indeed, it seems to me that we have every reason to expect empirical psychology to

reveal many crucial truths about how our minds work. The philosophical account that I have sketched above could at best only reveal what is *essential* to the various types of intentional mental states as such. Empirical psychology could tell us about all the numerous *contingent* truths about the nature of the *human* mind (or the minds of any of the other actually existing species). Thus, empirical psychology can be expected to give us a huge amount of information about the contingent psychological dispositions that we have, which are not essential to the capacity for having the various types of intentional mental states. For example, the contingent facts mentioned above about our dispositions to accept fallacious forms of reasoning are facts that we can know only by doing empirical psychology.

Moreover, the rational dispositions that this philosophical account appeals to are all dispositions at the “personal level”. This philosophical account cannot tell us anything about the “sub-personal level” — that is, about the various unconscious mechanisms and processes that realize our conscious mental processes. Thus, we may also expect empirical psychology to give us much illuminating information about these “sub-personal” processes and mechanisms, including information about the “sub-personal” mechanisms that realize our possession of the dispositions that are essential to our being capable of the various types of intentional mental states. For example, empirical psychology may tell us that we actually have two reasoning systems — in the terms of Stanovich and West (2000), System 1 and System 2; and it may also tell us, as Johnson-Laird and Byrne (1993) suggest, that System 1 does not involve any encoded rules of inference, but operates by means of other mechanisms. So far as I can see, it is quite compatible with my philosophical claim that possession of the concept ‘if’ requires a rational disposition to accept *modus ponens* inferences (at least in the absence of defeaters) that this rational disposition itself is realized in a state of System 1, which does not itself involve any encoding of the rule of *modus ponens* itself.<sup>16</sup> In general,

the claims that I have been defending here are not obviously inimical in any way to any such accounts of the sub-personal mechanisms that realize our reasoning capacities.

In general, my view of the relation between empirical psychology and the philosophy of mind is entirely in line with Rey's analogy (this volume, § 2.2; compare 2002, 101):

our understanding of the minds of people and animals seems like the understanding clever children have of their computers: they know a good deal about their *interaction* with them — they can play games, and get them to do various things — but with only the sketchiest ideas about their internal causal/computational structure.

Philosophical, *a priori* psychology can at best give us a certain basic part of the sort of understanding that a skilled computer *user* has of his computer. Empirical psychology can ideally give us other sorts of understanding as well. Empirical cognitive science could ideally give us the sort of understanding that the computer *programmer* has of the computer's programming code; and empirical neuroscience can ideally enable us to achieve the sort of understanding that an *electrical engineer* has of the computer's internal electrical circuits. Still, it does not follow that the understanding of the mind that we can achieve through *a priori* psychology is worthless. After all, many skilled computer users are paid a decent wage to advise others on how to use their computers to do such things as the following: to import and manipulate sound and image files; to create spreadsheets, web pages, and electronic documents; to root out viruses, and fix other software problems; and to install and configure the computer's programs in ways that will be most useful to the users. In many cases, these computer advisers need no more than the most rudimentary knowledge of the computer's electrical circuits or of its programming code in order to do their job perfectly well.

However, Rey and other philosophers of broadly "naturalist" sympathies may be tempted to make a stronger claim. According to this stronger claim, the various types of

intentional mental states are like “natural kinds”, in that the essential nature of these mental states *cannot* be known *a priori*, but can *only* be known empirically. Now, I am not denying that the essential nature of these mental states can be known empirically. At least, it is certainly possible to observe empirically what mental dispositions various thinkers have, and then further reflection on these empirical observations may lead one to understand which of these dispositions are essential to the various types of mental states and which are not. What I am claiming is that it is *also* possible to know the essential nature of these mental states *a priori*: even without observing the dispositions that thinkers actually have, one can simply try to consider more abstractly which dispositions it is possible to lack, and which it is not possible to lack, given the essential nature of the relevant mental states.

Admittedly, many objections could be raised against the idea of such *a priori* knowledge of the nature of our concepts and attitudes. I cannot discuss all these objections here. But *prima facie*, it would be surprising if the essential nature of the concepts that we possess could not be known *a priori* — that is, by relying not on any empirical observations, but simply on the resources that are already built into the mind itself. When one investigates what it is to possess a given concept, one’s investigation itself involves an exercise of the very capacity that one is investigating. So it should be possible, at least under favourable circumstances, for this investigation to draw directly on the very phenomenon that is being investigated. Traditionally, many philosophers have worried about how one could know anything about *extra-mental* reality by purely *a priori* methods; but there was usually thought to be much less difficulty in knowing the nature of our *concepts* by methods that draw on what is already built into our possession of those very concepts. If the nature of our concepts can be known *a priori*, then it also seems plausible that the nature of the various types of attitude can be known *a priori* as well.

Moreover, the mental facts that can clearly only be known by the methods of empirical psychology — facts about the unconscious sub-personal mechanisms and processes that realize our perceptual, cognitive and behavioural processes — do not seem to be essential to the intentional mental states in question. This is an old point, which Ned Block (1978, 310–11) made in criticizing the view that he called “Psychofunctionalism”. It seems intuitively possible for the very same type of conscious mental process to be realized in very different sub-personal mechanisms. To take Block’s example, suppose that when we get to know the Martians, we develop extensive cultural interaction with them: we study each other’s science and philosophy journals, read each other’s novels, and so on. But it turns out that our sub-personal processing systems are very different. (Suppose, for example, that it is as if our brains had been designed to use as much memory capacity as necessary in order to minimize use of computation capacity, and their brains had been designed to use as much computation capacity as necessary in order to minimize use of memory capacity.) Still, it intuitively possible for these Martians to have many of the same sorts of intentional mental states — beliefs, desires, plans, and so on — as we do. Thus, it seems possible for there to be creatures who share some of our intentional mental states but do not have the specific sorts of unconscious sub-personal mechanisms and processes that we have. Thus, these unconscious sub-personal mechanisms seem not to be essential to these intentional mental states.

So there is no obvious reason to think that knowledge of the sorts of facts that only empirical psychology can tell us is strictly necessary in order to understand the essential nature of these intentional mental states. It seems, *prima facie*, as though it should be possible, at least in principle, to know the essential nature of these types of intentional mental states *a priori*.<sup>17</sup> The fact that my claim that the intentional is normative is based purely on philosophical *a priori* considerations is not in itself an objection to that claim.

## Notes

1. For some very different versions of the claim, see the works of Davidson (1980, Essays 11 and 12; and 2001, Essays 9 and 10), Brandom (1994), and Morris (1992); for a parallel normative theory of linguistic meaning, see Lance and Hawthorne (1997).
2. For more on my conception of normative concepts, see Wedgwood (2001 and 2006a).
3. In effect, this is a principle of interpretive *charity* rather like the principle that was famously advocated by Davidson (2001, Essay 9). A normative theory could invoke such a principle without accepting Davidson's full-blown "interpretivism". As Lewis (1974) suggests, the reference to interpretation could just be taken as a way of dramatizing what is objectively constitutive of the intentional states in question.
4. For an illuminating discussion of *ceteris paribus* laws, see Rey and Pietroski (1995).
5. For this notion of "natural properties", see especially Lewis (1999, Essays 1 and 2).
6. It is this feature of my "dispositionalism" about concepts that I believe will enable me to escape the sort of indeterminacy worries that are inspired by Kripke (1982).
7. For a more detailed statement of this argument, see Wedgwood (2006b).
8. This conception of concepts has been disputed recently, e.g. by Timothy Williamson (2006). I reply to some of Williamson's arguments in § 3, but unfortunately I shall not be able to offer a full defence of this conception of concepts here.
9. I am assuming that dispositions to use a concept in reasoning are all either rational dispositions or irrational dispositions. None of them is a purely non-rational disposition (like the disposition to sneeze when one inhales finely ground pepper). The norms of rationality cover the whole of reasoning, and so there are no processes of reasoning that are neither rational nor irrational in that way.
10. A subtle version of this idea has been defended recently by Matti Eklund (2002).
11. It was careless of me to suggest that the first point alone was enough to answer the objections of Stein (1996); Rey was quite right to criticize me on this score.
12. For these results, see Oaksford (2005, 427).
13. Williamson (2006, 18) points out that "in some cases, when a further premise of the form 'If *r* then *q*' is added to *modus ponens* only a minority endorse the inference (Byrne 1989)." But an inference with an extra premise of this form is simply not an instance of *modus ponens*. So these cases are simply irrelevant to my claim that people who possess the concept 'if' have a suitable disposition to accept instances of *modus ponens*. (There is a difference between having a background belief, and consciously considering an inference that actually involves the content of that belief as a premise of the inference.)

14. This theory of action is often ascribed to Davidson (1980, Essay 1). For the record, I would prefer simply to define an action as the execution of an intention or volition; but I obviously cannot defend this account of action here.

15. See especially Alexander Bird (1998) and Michael Fara (2005).

16. As I explained in § 1, for a sub-personal mechanism to realize this personal-level rational disposition, it is not necessary that it should not realize any other personal-level dispositions as well. All that is necessary is that this sub-personal mechanism should make it *no accident* that normally, *ceteris paribus*, the thinker responds to the mental state of consciously considering an inference that is in fact an instance of *modus ponens* (in the absence of defeaters) with the mental state accepting that inference.

17. I am very grateful to Georges Rey for taking the time to engage in this debate with me, and to Timothy Williamson for some very helpful comments on an earlier draft.

## References

- Bird, Alexander (1998). "Dispositions and Antidotes", *Philosophical Quarterly* 48, 227–34.
- Block, Ned (1978). "Troubles with Functionalism", *Minnesota Studies in the Philosophy of Science*, ed. C. W. Savage, Vol. 9, *Perception and Cognition: Issues in the Foundations of Psychology* (Minneapolis: University of Minnesota Press), 261–325.
- Brandom, Robert (1994). *Making It Explicit* (Cambridge, MA: Harvard University Press).
- Byrne, R. M. J. (1989). "Suppressing valid inferences with conditionals", *Cognition*, 31: 1–21.
- Davidson, David (1980). *Essays on Actions and Events* (Oxford: Clarendon Press).
- (2001). *Inquiries into Truth and Interpretation*, revised edition (Oxford: Clarendon Press).
- Eklund, Matti (2002). "Inconsistent Languages", *Philosophy and Phenomenological Research* 64: 251–75.
- Fara, Michael (2005). "Dispositions and Habituals", *Noûs* 39: 43–82.
- Frege, Gottlob (1892). "Über Sinn und Bedeutung", *Zeitschrift für Philosophie und philosophische Kritik*, NF 100: 25–50.
- Hursthouse, Rosalind (1991). "Arational Actions", *Journal of Philosophy* 88 (No. 2, February), 57–68.

Johnson-Laird, P. N., and Byrne, R. M. J. (1993). "Models and deductive rationality", in K. Manktelow and D. Over, eds., *Rationality: Psychological and Philosophical Perspectives* (London: Routledge).

Kripke, Saul (1982), *Wittgenstein on Rules and Private Language* (Cambridge, Massachusetts: Harvard University Press).

Lance, Mark, and Hawthorne, John (1997). *The Grammar of Meaning* (Cambridge: Cambridge University Press).

Lewis, David (1974). "Radical Interpretation", *Synthese* 23: 331–344.

——— (1999). *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press).

McGee, Vann (1985). "A Counterexample to Modus Ponens", *Journal of Philosophy* 82: 462-71.

Morris, Michael (1992). *The Good and the True* (Oxford: Clarendon Press).

Oaksford, M. (2005). "Reasoning", in N. Braisby and M. Gellatly, eds., *Cognitive Psychology* (Oxford: Oxford University Press).

Raz, Joseph (1999). "Agency, Reason, and the Good", in Raz, *Engaging Reason* (Oxford: Clarendon Press).

Rey, Georges (2002). "Physicalism and Psychology: A Plea for Substantive Philosophy of Mind", in Carl Gillett and Barry Loewer, eds., *Physicalism and Its Discontents* (Cambridge University Press).

Rey, Georges and Pietroski, Paul (1995). "When Other Things Aren't Equal: Saving Ceteris Paribus", *British Journal for the Philosophy of Science* 46: 81-110.

Stanovich, K. E., and West, R. F. (2000). "Individual differences in reasoning: implications for the rationality debate?", *Behavioral and Brain Sciences* 23: 645–665.

Stein, Edward (1996). *Without Good Reason* (Oxford: Oxford University Press).

Strawson, P. F. (1952). *Introduction to Logical Theory* (London: Methuen).

Wedgwood, Ralph (2001). "Conceptual Role Semantics for Moral Terms", *Philosophical Review* 110 (2001), 1–30.

——— (2006a) "The Meaning of 'Ought'", in Russ Shafer-Landau, ed., *Oxford Studies in Metaethics*, Vol. 1 (Oxford: Oxford University Press), 127-160.

————— (2006b). “The Normativity of the Intentional”, in Brian P. McLaughlin and Ansgar Beckermann, eds., *The Oxford Handbook of the Philosophy of Mind* (Oxford: Clarendon Press).

Williamson, Timothy (2003). “Understanding and Inference”, *Proceedings of the Aristotelian Society* Sup. Vol. 77: 249-293.

————— (2006). “Conceptual Truth”, *Proceedings of the Aristotelian Society* Sup. Vol. 80: 1–42.