

# Introduction

**Teresa Marques**

University of Barcelona

University of Lisbon

This is a special issue of *Disputatio* on normativity and rationality. The idea for this volume originated after the fifth European Congress of Analytic Philosophy, ECAP5, which took place in Lisbon in August of 2005. This volume collects the contributions of John Broome, Pascal Engel, Kevin Mulligan, Josep Prades and John Skorupski, who were speakers on that occasion. The common thread in the diverse talks suggested that a volume on the topic would be of general interest. This common thread, which revealed itself naturally, testifies to issues hotly debated in recent years and also to the centrality of normativity in various philosophical areas.

Normativity is a central notion in distinct areas of philosophy, from ethics, legal philosophy or aesthetics, to philosophy of mind, language and action. It is of special importance in accounts of rationality, either theoretical or practical, since rationality can be characterized as a system of requirements or principles. For instance, it seems to be a requirement of rationality that one should not have contradictory beliefs, that one believes what follows by modus ponens from beliefs one holds, or that that one should intend to do what one believes one should do (cf. Broome, p. 162–163 this volume).

What we mean by ‘normativity’, though, can be thought to be ambiguous. And it can be thought to be so in several ways (Broome, Engel and Skorupski, in this volume, make different distinctions between possible senses of ‘normativity’). I will restrict myself here to two ways of understanding what we might mean by ‘normativity.’ In one sense, normativity concerns any principles or rules governing a given practice, activity or conduct. As such, it concerns the evaluation of acts, attitudes or mental states as those which are correct, justified or which one ought to do, have or be in, because they conform to the rules or norms governing them. In other words, an act or state might be correct or justified because, and insofar as, it follows a given rule or norm. For instance, driving on

the right hand side of the road in continental Europe is correct insofar as it conforms to European traffic regulations. Here is a distinct type of example: it might be correct to infer  $q$  from  $p$ , and from if  $p$ , then  $q$ , because doing so conforms to modus ponens. Thus, a given rule or requirement sets a standard according to which an act or state is evaluated as correct or incorrect. It is correct if it follows the rule. Normativity, in this sense, is a standard for the evaluation of *acts* or of *states*.

Some rules or requirements might also be thought to be constitutive of the practices they regulate. A typical example is that of the rules that govern chess as constituting what the game is. An interesting question debated in the current literature on normativity is whether there are constitutive norms for certain activities, such as speech acts like assertion, or intentional states like belief or desire. The constitutive norms for specific acts or intentional states would both individuate those acts or states as distinct from other acts or states, and would also set the relevant standard according to which that act, or state, is correct.

There is another sense of normativity, though. That concerns whether one *ought* to follow, or has reasons to follow, given requirements or principles. To illustrate, consider the traffic regulations example. It is one thing to say that you will follow the traffic regulations in continental Europe only if you drive on the right; it is another thing to say that when you drive in continental Europe you ought, or have reason, to follow the traffic regulations enforced.

Insofar as one is bound to follow given rules, by engaging in a norm-governed activity (either driving in continental Europe, or reasoning), one is held accountable and responsible for complying, or failing to comply, with the rules enforced. Normativity, in this second sense, imposes a standard of evaluation *on agents* as committed to, and responsible for, following given rules or requirements.

The second sense of normativity is of special importance in accounts of rationality, where the two senses can be illustrated with specific examples. So, it is one thing to say that you do not violate the principle of non-contradiction if you do not have contradictory beliefs (first sense of normativity). It is quite another thing to say that it is *irrational* to have contradictory beliefs, that is, that *you ought*, because rationality so requires, not to have contradictory beliefs. The general question is whether these requirements or principles of rationality provide reasons, and whether they impose

obligations, to act in given ways. This relates to more basic issues of practical rationality. Namely, to what it is for  $X$  to be a reason for an agent to do  $F$ , where  $X$  can be a fact that is taken as a reason to  $F$ , or as an end desired or intended by an agent.

The papers in this volume take their position within the different areas of the current debate on normativity and rationality, although naturally they do not exhaust all the different aspects of the ongoing discussion. One of the issues approached by two of the authors in this volume is whether intentional states can be individuated in terms of their correctness conditions.

Kevin Mulligan's paper, 'Intentionality, Knowledge and Formal Objects,' concerns the broad issue of how to characterize the intentionality of factive mental states in contrast with the intentionality of non-factive mental states, or, as Mulligan puts it, the intentionality of states which can, and those which cannot, miss their mark. Mulligan argues for an intrinsic constitutive difference between the two kinds of states. He contrasts two accounts of the intentionality of states that can miss their mark: theories of correctness conditions and theories of satisfaction conditions. Mulligan favours normative accounts of non-factive mental states, i.e., those that appeal to correctness conditions to individuate the intentional states in question.

In contrast with the states that can miss their mark are factive states, like knowing or seeing. Mulligan proposes that the intentionality of factive states is more basic than that of non-factive intentional states, and suggests that the most basic mind-world relations are not normative. He further claims that the type of intentionality of the most basic, factive, intentional states, weighs in favour of normative accounts of non-factive intentional states.

Pascal Engel's paper, 'Belief and Normativity,' also focuses on the possibility of elucidating intentional states, in particular belief, in normative terms, and offers a defence of the claim that the norm of belief is truth. Rather than contrast the norm of truth for belief with other possible constitutive norms, Engel examines versions of the claim that the truth norm is individuating of beliefs and defends this from several criticisms. It is a significant feature of the debate on the normativity of intentional states (i.e., of the debate concerning which norm, if any, governs an intentional state), revealed in Engel's discussion, that objections against a certain norm governing an act or state should appeal to issues pertaining to normativity in the second sense described earlier. That is, criticisms against some-

thing being a constitutive norm often draw on what subjects ought to do, or have reasons to do, and on agents' motivations to act, arguing that the putative norm is impossible to follow, or is not motivating, etc.

In the particular case of belief, which Engel considers, the criticisms against the truth norm claim that it lacks normative force, that it violates the principle that 'ought' implies 'can', and that it is viciously circular. The first objection makes sense once the two ways to understand 'normativity' are distinguished. If there were a norm constitutive of belief, that would not explain in what sense, if any, one *ought* to believe what is true. The second objection elaborates on the normative force of the putative truth norm of belief, and insists that the truth norm for belief is impossible to satisfy, imposing on believers constraints which they cannot meet — people cannot be expected to believe *all* truths, for instance. The concern then becomes how to formulate a constitutive norm for belief which imposes on believers a requirement which they can satisfy. The final objection concerns why speakers care for believing the truth, i.e., how the truth-norm is supposed to be motivating, without being necessarily so. The objection is that if the truth norm is constitutive of belief, then it necessitates motivation. If so, the sense in which it is supposed to be normative is unclear. Engel proposes, in reply, a formulation of the truth norm which, he argues, avoids the objections — for any  $p$ , a subject  $S$  ought to believe that  $p$  only if  $p$ , and proposes that one should distinguish between a norm and its regulation.

As mentioned, it is a revealing fact about debates on normativity that concerns with what agents have reasons or ought to do, and with what they are motivated to do — i.e., concerns that depend on normativity in the second sense introduced earlier — weigh in favour, or against, given particular norms, rules or principles as providing the standards of correctness for specific actions or intentional states.

Josep Prades's paper in this volume, 'Acting without Reasons,' offers a sceptical and critical perspective on the role of reasons in practical rationality in general. The concern is with the patterns of practical reasoning, and with how to explain that an agent can take  $X$  as a reason to do  $F$ . Prades offers a general argument form, which is independent of what provides reasons to act, whether desires or other pro-attitudes. In his paper, he illustrates the argument with examples. The general form of the argument is the following. If a

set  $G$  of pro-attitudes is sufficient to motivate a subject  $S$  to act, then it is also sufficient to motivate a subject  $R$  to act (on the assumption that  $S$  and  $R$  have the same pro-attitudes regarding  $F$ ). However,  $G$  is not sufficient to motivate  $R$  to do  $F$ , because  $R$  does not do  $F$ . Since  $R$  has the same pro-attitudes as  $S$ , but is not motivated to act by them, it follows that those pro-attitudes are not motivating, and therefore, not motivating *even* for  $S$ .

Prades's explanation of this is that when an agent answers a why-question, s/he does not give his/her reasons to form a given intention. Rather s/he identifies the *content* of the intention to do  $F$ . In reply to why-questions, the agent explores conversational commitments to provide relevant information about the intention with which s/he acts. Among the many causal antecedents for an action, the agent chooses to mention those that conversationally implicate information about the relevant kind of reason one is being asked about. These Prades calls *Content Determining Reasons* (CD). The conversational implicature can be cancelled, indicating that the original rationalization was not the correct one. Prades concludes by offering general remarks on the structure of practical reason. There is nothing, he believes, like a logical form of pure instrumental practical reasoning, because there is nothing like pure instrumental practical reasoning. No rational agent chooses a means just in virtue of its instrumental relation with a given goal. Rather, agents make up their minds forming specific intentions, a process that can be ideally rationalized.

A radically different attitude towards the role of reasons in our rational lives is put forward by John Skorupski, in his paper 'What is Normativity?' Skorupski offers a positive attempt to elucidate and formulate what he calls *Reasons Thesis*, the thesis that the fundamental normative concept is that of a reason. Skorupski motivates the Reasons Thesis in different ways; one central motivation for it is that it illuminates, he believes, the 'fact/value' or 'is/ought' distinction. Another motivation for the thesis is, he thinks, that normativity must come down to an agent's reasons for belief, feeling or action, i.e., to the relation between facts, on the one hand, and reason-responsive acts and states, on the other hand.

Skorupski considers two ways to formulate the Reasons Thesis. The first one takes any normative predicate to be definable in terms of a reason predicate. This is the semantic thesis. The conceptual thesis, in contrast, takes the normative component in any normative

concept to be a concept of a reason. Skorupski does not defend the semantic thesis, reducing normative predicates to reason predicates, since it may turn out, as it seems to happen in particular cases, that if a normative predicate is defined in terms of a reason to act, or feel, we may then only be able to define the act or sentiment by appealing to the normative term itself. The project would be jeopardized by the risk of circularity. Thus, Skorupski proposes that any normative concept be elucidated in terms of what is required for its possession. Skorupski dedicates most of the paper to trying to explore how far the semantic thesis can be taken, since this might nonetheless be illuminating for the understanding the Reasons Thesis.

A problem that a proponent of the semantic thesis might have is to identify which terms are normative. Skorupski's strategy is to focus on prominent types of *prima facie* normative terms that are definable in terms of reason predicates. He distinguishes between three types of relational reason predicates: specific reasons of degree, overall reasons of degree and sufficient reasons. Skorupski then defines what are the explicit normative sentences — atomic normative sentences, built from the relational reason predicates, and sentences built from atomic normative sentences by means of the connectives of propositional logic, the quantifiers and the truth operator. It follows, on his account, that the negation of a normative sentence or a conditional sentence with a normative consequent will be normative sentences too. The point of this strategy is to establish that one cannot derive *normative* conclusions from true non-normative premises alone, which, Skorupski believes, captures the 'is/ought' or 'fact/value' distinction. In the remaining of his paper, Skorupski undertakes the task of showing how predicates that we are disposed to classify as normative will also be normative in *his* sense, advancing arguments to so classify both thin and thick normative predicates. (Among the thin predicates he considers 'ought,' 'should,' 'right,' 'good,' or 'bad.' In the case of 'good' and 'bad,' in particular, Skorupski defends a buck-passing account as a plausible reduction to reasons, and replies to possible objections.)

John Broome's paper, 'Is Rationality Normative?' takes a much more sceptical outlook on normative issues, addressing the central question of whether rationality gives us reasons to follow its requirements. Another way to raise the same question is to ask whether one ought to be rational. Broome draws on a distinction between two senses of normativity, along the lines of the one de-

scribed in the beginning of this introduction. In one sense, what is normative is what sets a standard of correctness conditions, correctness according to specific rules or requirements. Another sense, that with which his paper is concerned, is whether those requirements give one reasons to follow them. Broome takes a sceptical perspective about possible answers to the question. Broome thinks that rationality is normative, but he argues that there are no good arguments in support of the claim that it is. Among the different requirements of rationality are, for instance, that you should believe what follows by modus ponens from things you believe, or that you should comply with the requirement of *krasia*: to intend to do what you believe you ought to do. Broome finds no satisfactory argument to show that rationality is normative, that is, that it gives us reasons to follow its requirements.

Broome construes the thesis of the normativity of rationality in three different degrees of strength. Strong: Necessarily, if rationality requires  $S$  to  $F$ , then  $S$  ought to  $F$  because rationality requires it. (Rationality provides a sufficient reason to  $F$ ). Medium: necessarily, if rationality requires  $S$  to  $F$ , there's a reason for  $S$  to  $F$  (either a sufficient or *pro tanto* reason). Weak: necessarily, if rationality requires  $S$  to  $F$ , there's a reason for  $S$  to  $F$  (which is either sufficient or *pro tanto*, but the rational requirement might not be the reason). Weak normativity is entailed by medium normativity, which in turn is entailed by strong. Broome holds that it is difficult to explain why any form of normativity (weak or otherwise) is true, if it is true. In the remaining of the paper, he argues that there are no good arguments to support even weak normativity. He considers the idea that rationality is normative for instrumental reasons, because it helps one to achieve some of the things one ought to achieve, as well as the idea that rationality consists in responding correctly to reasons, and rejects both.

I would not like to conclude this introduction without thanking the authors of the papers collected in this volume for their willingness to participate in the project, their generosity and their patience during the whole process. I would also like to thank the editorial committee of *Disputatio*, the editor João Branquinho, and the man-

aging editor Desidério Murcho, for all the support given to the publication of this volume.<sup>1</sup>

Teresa Marques  
LOGOS research group  
Departament de Lògica, Història i Filosofia de la Ciència  
Facultat de Filosofia, Montalegre, 6-8, 4<sup>a</sup> planta  
08001 Barcelona, Spain

LanCog research group  
Centro de Filosofia, Universidade de Lisboa  
Faculdade de Letras, Alameda da Universidade  
1600-214 Lisboa, Portugal  
teresamarques@fl.ul.pt

<sup>1</sup> Work supported by research fellowship, Fundação para a Ciência e a Tecnologia (FCT), SFRH/BPD/16678/2004, by Spanish Ministry for Education and Science (MEC), DGI SB2006-0036. Research project Content, FCT, POCI/FIL/55562/2004, and Logos, Grupo de Logica Lenguaje y Cognicion, HUMN2006-08236 (C-Consolider), MEC.

# Is rationality normative?

**John Broome**

Corpus Christi College, University of Oxford

## **Abstract**

Rationality requires various things of you. For example, it requires you not to have contradictory beliefs, and to intend what you believe is a necessary means to an end that you intend. Suppose rationality requires you to F. Does this fact constitute a reason for you to F? Does it even follow from this fact that you have a reason to F? I examine these questions and reach a sceptical conclusion about them. I can find no satisfactory argument to show that either has the answer 'yes'. I consider the idea that rationality is normative for instrumental reasons, because it helps you to achieve some of the things you ought to achieve. I also consider the idea that rationality consists in responding correctly to reasons. I reject both.

## 1. The normative question

Rationality requires various things of you. For example, it requires you not to have contradictory beliefs, to believe what follows by modus ponens from things you believe, to intend what you believe to be a necessary means to an end that you intend, and to intend to do what you believe you ought to do. These are only rough formulations, and rough formulations are perhaps enough for this paper.

However, for the sake of accuracy, I shall give more formal statements of the requirement that you believe what follows by modus ponens from things you believe, and the requirement that you intend to do what you believe you ought to do. These ones will come up later in the paper. They are:

*Modus ponens.* Rationality requires of  $N$  that, if  $N$  believes  $p$  and  $N$  believes that if  $p$  then  $q$ , and if it matters to  $N$  whether  $q$ , then  $N$  believes  $q$ .

*Krasia.* Rationality requires of  $N$  that, if  $N$  believes that she herself ought to  $F$ , and if  $N$  believes that she herself will  $F$  if and only if she herself intends to  $F$ , then  $N$  intends to  $F$ .

These are technically requirement-schemata; individual requirements are obtained by making appropriate substitutions for the schematic letters: the name of a person for '*N*', a sentence for '*p*' and a verb phrase for '*F*'. Each formula states a requirement on *N*'s contemporaneous mental states. To make that explicit, I could have added 'at *t*' to each of the attitudes that are mentioned in the formulae; I did not do so simply to make the formulae less complicated. In the formula for Krasia, 'she herself' is a compound reflexive pronoun; it represents in indirect speech the first-person pronoun 'I' in direct speech.<sup>1</sup> The formulae contain some clauses that do not appear in the informal statements; I leave you to work out why they are needed.

I shall use less formal language in this paper. Even so, I shall sometimes use a schematic letter to stand for a verb phrase. For instance, I might say 'rationality requires you to *F*.' This is unattractive, but sadly English does not provide a good alternative. My '*F*' may stand for a complex phrase, such as 'intend what you believe to be a means to an end that you intend.' There is no generic verb in English with a broad enough meaning to stand in for all the substitutions that may be needed. 'Do' is too narrow; it covers little more than actions. However, because a schematic letter is not always tolerable, in places I shall artificially adopt 'achieve' as my generic verb. This is not accurate English, but fortunately the contexts make it possible. I hope you will accept it now that I have warned you.

This paper considers whether rationality is normative. It is automatically normative in one sense. Rationality is a system of requirements or rules. It therefore sets up a notion of correctness: following the rules is correct according to the rules. That by itself makes it normative in one sense, because in one sense 'normative' simply means to do with norms, rules or correctness. Any source of requirements is normative in this sense. For example, Catholicism is. Catholicism requires you to abstain from meat on Fridays. This is a rule, and it is incorrect according to Catholicism to eat meat on Fridays. So Catholicism is normative in this sense.

But I do not use 'normative' in that sense. In my sense, it means to do with ought or reasons. Given a rule or a requirement, we can ask whether you ought to follow it, or whether you have a reason to do so. Is there any reason for you to abstain from meat on Fridays, for

<sup>1</sup> See Castañeda 1968.

instance? I shall ask the corresponding question about the requirements of rationality. ‘Requirement’ is not a normative word in my sense.<sup>2</sup> When I say rationality requires this or that of you, I do not mean anything normative in my sense. Nevertheless, it seems plausible that rationality actually is normative. Is it? I shall call this the ‘normative question’ about rationality. This paper looks at various possible arguments for the answer ‘yes’, and rejects them all.

What does the normative question mean, more exactly? The claim that rationality is normative comes in various strengths. The strongest is:

*Strong normativity.* Necessarily, if rationality requires  $N$  to  $F$ , then  $N$  ought to  $F$  because rationality requires  $N$  to  $F$ .

There is another way to say this. When you ought to  $F$  because  $X$ , we say that  $X$  is a reason for you to  $F$ . When, necessarily, you ought to  $F$  if  $X$ , we say that  $X$  is sufficient for it to be the case that you ought to  $F$ . When the two things are true together, we may say that  $X$  is a *sufficient reason* for you to  $F$ .<sup>3</sup> So strong normativity says that, if rationality requires you to  $F$ , that fact is a sufficient reason for you to  $F$ .

This claim is implicit in the way many philosophers write, including me in the past. It is common to describe requirements of rationality using the normative term ‘ought’. For example, saying ‘You ought not to have contradictory beliefs’ is a common way of expressing the requirement not to have contradictory beliefs. In adopting this style of expression, I was implicitly assuming strong normativity.<sup>4</sup>

On the other hand, some philosophers think rationality has a normative force that is defeasible. They think that, if on some occasion your having contradictory beliefs would be very beneficial — perhaps it would prevent a war — then it is not the case that you ought not to have contradictory beliefs. Nevertheless, these philosophers think that rationality requires you not to have contradictory beliefs, and they think this is a consideration that counts against your having them. So they think the rational requirement not to have contradictory beliefs

<sup>2</sup> I have examined the meaning of ‘requirement’ in Broome 2007b.

<sup>3</sup> Some authors use the term ‘sufficient reason’ differently. By ‘a sufficient reason to  $F$ ’ they mean a reason that is sufficient to make it permissible for you to  $F$ . I mean a reason that is sufficient to make it the case that you ought to  $F$ .

<sup>4</sup> Andrew Reisner persuaded me to change my practice.

is a pro tanto reason for not having them. They claim that rationality is normative in a second, weaker way:

*Medium normativity.* Necessarily, if rationality requires  $N$  to  $F$ , that fact is a reason for  $N$  to  $F$ .

I mean ‘a reason’ in this formula to include both sufficient reasons and pro tanto reasons. So medium normativity is entailed by strong normativity, but it is weaker than strong normativity.

If you believe medium normativity but not strong normativity, you will think that the reasons generated by rationality sometimes conflict with reasons that issue from other sources of normativity, such as morality or prudence. But you will probably believe that conflicts of this sort are rare. Rationality is principally concerned with coherence among your attitudes such as your beliefs and intentions, whereas morality, prudence and other sources of normativity are rarely concerned with those things. Rationality has a domain of application where it is pretty much on its own. Examples of conflict between rationality and other sources of requirements tend to be far-fetched, like my example of a preventing a war. So according to medium normativity, when rationality requires something of you, it will normally be the case that you ought to achieve that thing.

A yet weaker view is:

*Weak normativity.* Necessarily, if rationality requires  $N$  to  $F$ , there is a reason for  $N$  to  $F$ .

According to weak normativity, that rationality requires you to  $F$  entails that you have a reason — either sufficient or pro tanto — to  $F$ , but the fact that rationality requires you to  $F$  may not itself be the reason. Many philosophers think rationality consists in responding correctly to reasons — a view I shall discuss in section 3. They may well accept weak normativity but neither of the stronger views. They do not think the fact that rationality requires something of you is itself a reason to achieve that thing. Still, because they think rationality consists in responding correctly to reasons, they think rationality would not require something of you unless there was a reason for you to achieve that thing.

Weak normativity is not really a version of the claim that rationality is normative. It associates rationality with normativity, but it does

not claim that rationality is a source of normativity. Nevertheless, in this paper I shall concentrate principally on weak normativity. That is because I shall reject even the arguments for that weak claim. Strong and medium normativity both entail weak normativity, so rejecting weak normativity entails rejecting both of them. They are both genuine versions of the claim that rationality is normative.

I do not disbelieve weak normativity. Indeed, I am inclined to believe even medium normativity. An important paper by Niko Kolodny (2005) has a stronger conclusion than mine. It concludes that rationality is actually not normative. But I am not convinced by Kolodny's arguments, for reasons I cannot go into here.<sup>5</sup> I am agnostic about this question.

Compare the corresponding normative question for morality. Medium normativity is surely true of morality; it is surely true that, necessarily, if morality requires you to *F*, that fact is a reason for you to *F*. But it is notoriously difficult to find a convincing explanation of why. Rationality may be in the same position. It may be that rationality is indeed normative, but it is hard to explain why. The fact that I cannot do so may be just because I am not clever enough.

## 2. Rationality and reasons

Many philosophers unhesitatingly assume there is some conceptual connection between rationality and reasons.<sup>6</sup> More specifically, they assume that acting contrary to reasons is irrational.<sup>7</sup> Why make this assumption? There are certainly connections between the words 'rational' and 'reason.' For one thing, they have the same Latin root. But, although etymology can be suggestive, it gives no real ground for thinking there is a connection between the concepts.

More confusingly, the word 'reason' has various senses, and in one of them it refers to the *faculty* of reason. 'Reason' used this way is a mass noun — a noun that has no plural. Our faculty of rationality (which I shall mention again in section 4) is plainly part of, or perhaps

<sup>5</sup> See Broome 2007a.

<sup>6</sup> Just one example is Smith 2004.

<sup>7</sup> Again just one example: Williams 1981.

all of, our faculty of reason. So there is certainly a conceptual connection between rationality and reason in this sense.

But the word ‘reason’ that appears in my formulae for medium and weak normativity does not have that sense. It is a count noun — a noun that has a plural. Its plural is ‘reasons’. Just because rationality is conceptually connected with the faculty of reason, it does not follow that it is conceptually connected with reasons.

When David Hume said

’Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger.<sup>8</sup>

he was using ‘reason’ to refer to the faculty of reason. Since rationality is a part of this faculty, Hume’s remark implies that it would not be irrational to have this preference. (Hume himself does not use the word ‘irrational’.) However, it does not imply that there is no reason not to have it. We should not attribute that extraordinary view to Hume.

At this point, Hume is saying nothing about reasons. In the *Treatise*, he rarely uses the count noun ‘reason’ in a normative sense, and never in the section entitled ‘Of the influencing motives of the will’ that contains this remark about preference. It would obviously be very immoral to prefer the destruction of the whole world to the scratching of your finger, and therefore you ought not to have this preference. Since Hume had no less of a moral sense than the rest of us, we may fairly assume he would have agreed you ought not to have it. The preference is contrary to morality, he might have said. But since Hume did not think morality arises from the faculty of reason, it was perfectly consistent for him to say that the preference is not contrary to reason, this faculty.

Had Hume been willing to speak of normative reasons, I assume he would have agreed that morality gives you a reason, in the normative sense, not to have this preference. He would have agreed the preference is contrary to this reason, that is. Some philosophers have been more shocked by Hume’s remark than they need have been.<sup>9</sup>

<sup>8</sup> Hume 1978, book 2, part 3, section 3.

<sup>9</sup> Johnston (1989: 161) is an example.

### 3. Responding correctly to reasons

The confusion of words could explain why many philosophers associate reasons and rationality. But is there *really* any conceptual connection that gives the answer ‘yes’ to the normative question about rationality? It is very commonly said that rationality consists in responding correctly to reasons. If this were so, it might be the connection I have just asked for. At least, it might support weak normativity.

It would not infallibly do so, because the idea that rationality consists in responding correctly to reasons has two alternative readings. On a narrow-scope reading it means, first, that whenever there is a reason for you to *F* then rationality requires you to respond correctly to this reason and, second, that this is all rationality requires of you. On a wide-scope reading it means, first, that rationality requires of you that whenever there is a reason for you to *F* you respond correctly to this reason and, second, that this is all rationality requires of you. The wide-scope reading does not support weak normativity, but the narrow-scope reading does. Given the narrow-scope reading, if rationality requires you to *F*, that can only be because there is a reason for you to *F*. Weak normativity follows.

Does rationality consist in responding correctly to reasons, under either reading? There is a quick objection to the idea that it does. On some occasion, there might be a reason for you to achieve something but, without any irrationality on your part, you might not believe this reason exists. If you do not believe it exists, then you might well not respond correctly to it, and your failure will not imply any failure of your rationality. Therefore, rationality cannot consist in responding correctly to reasons.

For example, suppose the fish in front of you contains salmonella. This is a reason for you not to eat it. But there may be no obvious evidence that it contains salmonella. So you might not believe it contains salmonella, and you might eat it, and nevertheless you might be rational. So you are rational even though you do not respond correctly to the reason.

Many philosophers find this quick objection convincing. As a result, few support precisely the idea that rationality consists in responding correctly to reasons. Instead, they support a modified idea that connects rationality with the reasons you believe to exist rather than with actual reasons. They say that rationality consists in responding correctly to the reasons that you *believe* to exist, or something of that sort.

I know no quick objection to this idea. But it does not support weak normativity. The narrow-scope reading of this view goes as follows: first, whenever you believe there is a reason for you to *F* then rationality requires you to respond correctly to this believed reason and, second, this is all rationality requires of you. We can derive the conclusion that if rationality requires you to *F*, that can only be because you believe there is a reason for you to *F*. But weak normativity does not follow, because you might believe there is a reason when actually there is not. Mistakes about reasons are possible.

So if the quick objection to the original idea is sound, we can get no support for weak normativity from any conceptual connection that holds between rationality and responding correctly to reasons. Is the quick objection sound? I know of one response that may be made to it. The objection depends on two assumptions. One is that there can be a reason for you to achieve something that, without any irrationality, you might not believe exists. The second is that, if you do not believe a particular reason exists, you may fail to respond correctly to this reason, without any irrationality on your part. On the face of it, both these assumptions seem hard to doubt. But one theory of rationality circumvents the quick objection, and casts doubt on them.

I shall call it 'the theory of attitudinal reasons,' for a reason that will soon appear. It assumes that all the requirements of rationality are in one respect like the examples of requirements I mentioned at the beginning of this paper. In my examples, rationality requires particular relations to hold among your attitudes. It requires you to have or not have a particular belief if you have other particular beliefs; it requires you to have a particular intention if you have other particular intentions and beliefs; and so on. The theory supposes that all rational requirements have this character. According to the theory of attitudinal reasons, then, your rationality consists in the holding of appropriate relations among your attitudes. That is its first feature.

Its next feature is that it understands the requirements of rationality in a particular way. It takes each requirement to imply that some of your attitudes constitute reasons for your having other attitudes. Take as an example the requirement *modus ponens* set out in section 1. One instance of it is that rationality requires of you that, if you believe it is raining and you believe that if it is raining the snow will melt, you believe the snow will melt. (For the sake of setting out the theory of attitudinal reasons, I shall ignore the condition that it matters to you whether the snow will melt.) Suppose you believe it is

raining and you believe that if it is raining the snow will melt. Let us call these beliefs your 'premise attitudes.' Let us call the belief that the snow will melt the 'conclusion attitude.' According to the theory, the premise attitudes together constitute a reason to have the conclusion attitude. Moreover, according to the theory, this is a sufficient reason: it makes it the case that you ought to have the conclusion attitude. Furthermore, according to the theory, to respond correctly to this reason is to have the conclusion attitude.

We are taking it for granted that rationality requires you to have the conclusion attitude if you have the premise attitudes. Therefore, you are irrational if you have the premise attitudes but do not have the conclusion attitude. As the theory of attitudinal reasons understands the requirement, you are irrational if you do not respond correctly to the reason in this case. So the theory implies that rationality consists in responding correctly to reasons. The reasons are of a particular type; they are constituted by the premise attitudes that are specified in requirements of rationality. The theory could allow there to be reasons of other types as well. Those ones would have to be unconnected with rationality. But for simplicity I shall assume that reasons of this type are the only ones it recognizes.

How does the theory of attitudinal reasons circumvent the quick objection? It must deny one or other of the assumptions that this objection relies on. It could deny the first. It could say that, when you have a reason to have some conclusion attitude, you must be irrational if you do not believe the reason exists. This is not entirely implausible. The reason consists of your premise attitudes, and it is not entirely implausible that, if you have attitudes of this sort, you are irrational unless you believe you have them.

However, I think the theory would do better to deny the second assumption. It would do better to say that, even if you do not believe the reason exists, nevertheless you are irrational if you do not respond to it correctly. This is plausible, given that the reason is constituted by your premise attitudes. Plausibly, you can come to have the conclusion attitude through a process of reasoning that does not require you to believe you have the premise attitudes. That indeed is my own view: I think that reasoning does not require you to have

second-order beliefs about your attitudes.<sup>10</sup> You can reason directly from the *contents* of your premise attitudes. In the example, you can reason from the contents of your premise beliefs — the proposition that it is raining and the proposition that if it is raining the snow will melt — to a new belief that the snow will melt. This reasoning process does not require you to have any beliefs about your attitudes; you do not need to believe you believe it is raining, for instance. If this is right, then you can respond correctly to your reason — the premise attitudes — without believing this reason exists. Furthermore, plausibly, you are irrational if you do not do so.

The upshot is that the theory of attitudinal reasons could offer a defence of the idea that rationality consists in responding correctly to reasons. It offers to defend only a very restricted version of the idea: that rationality consists in responding correctly to reasons that are constituted out of premise states. But still, this is some defence.

However, the defence is not yet complete. There is first of all a matter of scope to attend to. Weak normativity says that, when rationality requires of you that you *F*, there is a reason for you to *F*. In the cases we are considering, what rationality requires of you is that, if you have the premise attitudes, you have the conclusion attitude. So we must take *F* to be the conditional that, if you have the premise attitudes, you have the conclusion attitude. But the theory of attitudinal reasons does not claim you have a reason to satisfy this conditional. It claims that, if you have the premise attitudes, you have a reason to have the conclusion attitude. So it does not support weak normativity directly. However, I doubt this is a serious difficulty. I expect it would be easy to argue that a reason to have the conclusion attitude is also a reason to satisfy the conditional *F*. At any rate, I shall ignore this difficulty.

The serious problem lies elsewhere. The theory of attitudinal reasons simply assumes that your premise attitudes constitute a reason to have the conclusion attitude. But this is to beg the normative question. We all agree that rationality requires of you that, if you have the premise attitudes, you have the conclusion attitude. This is not a normative claim, since I do not use ‘requires’ as a normative word. But the theory of attitudinal reasons understands this requirement in a

<sup>10</sup> My objection to what I call the ‘second-order model’ of reasoning is briefly set out in Broome 2006.

normative way. It takes it to imply that the premise attitudes constitute a reason for having the conclusion attitude. It expresses the rational requirement using the normative term 'a reason.' But it has no right to do that until it has established that the rational requirement does indeed make a normative connection between the premise attitudes and the conclusion attitude. Whether this is so is just the normative question.

The theory of attitudinal reasons is not truly a defence of the idea that rationality consists in responding correctly to reasons. It is simply an assertion that it is so, within a limited domain. It offers no argument for the claim that rationality is normative.

#### 4. An instrumental reason to be rational

Section 3 shows that we can get no support for weak normativity from the idea that rationality consists in responding correctly to reasons. I am now going to pursue a different approach to the normative question. I shall pursue it in a generous spirit; I shall do my best to make it work. On its behalf, I shall make assumptions I cannot properly justify, and rely on arguments that are frankly rough. I have already said that my ultimate conclusion will be sceptical: I know no good grounds for thinking rationality is normative. Given that, a generous spirit is appropriate. I want to see how far it can take us in finding grounds.

The approach I am going to pursue starts by recognizing that, if rationality is indeed normative, that seems likely to be because of what we can achieve by being rational. It seems likely to be for instrumental reasons, as I shall put it.

Since some philosophers give a different meaning from mine to the word 'instrumental,' I need to be clear about mine. I am not suggesting the requirements of rationality might be normative because satisfying them is a way of satisfying our desires. I am suggesting they might be normative because satisfying them is a way of achieving some of the things we ought to achieve. I ignore the possibility of total scepticism about normativity; I take it for granted there are some things we ought to do, some things we ought to hope for, some things we ought to believe, some things we ought not to do, not to hope for, not to believe, and so on. There are some *F*s such that we ought to *F*. Rationality seems plausibly a good means of coming to *F* in many instances when

we ought to *F* — of achieving much of what we ought to achieve. Perhaps this explains why we have a reason to satisfy rational requirements. The rest of this paper explores that idea.

It cannot be applied directly to individual requirements. If there is a reason for you to satisfy each individual requirement, the reason cannot always be a directly instrumental one. That is because satisfying a particular requirement will sometimes not contribute to your achieving anything you ought to achieve. Indeed, it will sometimes prevent you from achieving something you ought to achieve. For example, suppose you believe you ought not to sell your car, but your belief is false and actually you ought to sell it. If you satisfy the requirement *krasia* of rationality on this occasion, you will intend not to sell your car. As a result of your intention, you will probably not sell it. But you ought to sell it. So in this case, satisfying a rational requirement will probably prevent you from doing something you ought to do.

Often no doubt, when rationality requires you to achieve something, you do have a direct instrumental reason to achieve it. But that is not necessarily so. There is therefore no direct instrumental argument for weak normativity.

Is there an indirect argument? We can start by noticing it is much more plausible that, for directly instrumental reasons, you ought to have the rational *faculty*. By ‘the rational faculty’ I mean the bundle of dispositions and abilities that causes you to satisfy many of rational requirements. This is only a rough definition because of the vague term ‘many,’ but it will serve for my purposes. The connection between the rational faculty and satisfying individual rational requirements is causal.

Possessing the rational faculty is plausibly part of the best means of achieving much of what you ought to achieve. By ‘best’ I mean better than other means that are psychologically possible for you. In principle, there might be an alternative faculty that could form part of an even better means of achieving much of what you ought to achieve. On some occasions, the rational faculty will steer you wrong; the car examples illustrates how it will sometimes cause you to fail to achieve something you ought to achieve. The alternative faculty would be like the rational faculty, but altered a little to correct some of those glitches. But the necessary alterations would depend on complex circumstances in the outside world, and I assume this alternative would not be psychologically possible for you.

I shall therefore assume that the rational faculty is part of the best means of achieving much of what you ought to achieve. Given that, two claims are plausible. First, you ought to have the rational faculty. Second, if you ought to have the rational faculty, that is because having it is part of the best means of achieving much of what you ought to achieve.

I am sorry to say I cannot offer a proper argument to support the first of these claims. One difficulty standing in the way is that a proper argument would have to depend on an accurate account of the way in which normativity is transmitted from ends to means. If you ought to achieve some end, how does that determine what means you ought to take to that end? In our particular case, I am assuming the rational faculty is part of the best means to the end of achieving things you ought to achieve. How does it follow that you ought to have the rational faculty? An accurate answer to this question would be complicated. It is not clear that you ought always to take the best means to an end that you intend.<sup>11</sup> The best means may be bad in some way that is unconnected with the end. If so, perhaps you ought to take some means other than the best one.

To deal with this difficulty properly would take all the resources of decision theory — decision theory deals with the relation between ends and means, and it can be interpreted in normative terms. But there is no need for that sort of detail in this paper. I am working in a generous spirit, and looking only for rough arguments. Let us accept the plausible claim that you ought to have the rational faculty.

The second claim is that, if the first claim is true, it is true because having the rational faculty is part of the best means of achieving much of what you ought to achieve. I can give this claim some support by describing a world where the rational faculty is not instrumentally effective in this way.

I assume it is only a contingent fact, if it is a fact at all, that the rational faculty is part of the best means of achieving much of what you ought to achieve. I assume there could be quirky worlds where that is not so. In a quirky world, people with the rational faculty generally satisfy the same requirements of rationality as rational people do in our world. They generally intend to do the things they believe they ought to do; they generally do not have contradictory beliefs; they

<sup>11</sup> As Carsten Nielsen helpfully reminded me.

generally believe what follows by *modus ponens* from things they believe; and so on. But because of the way causal processes work in their world, satisfying the requirements of rationality tends to be unsuccessful. These people tend not to end up having the beliefs they ought to have, doing the things they ought to do, and so on. They do not achieve much of what they ought to achieve.

In a quirky world there are also people who do not have the rational faculty. Those people do just what they feel like doing, believe whatever comes into their heads, and so on. The causal processes in their world bring it about that these people achieve much of what they ought to achieve.

Plausibly, it is not the case that people in a quirky world ought to have the rational faculty, since it is not a means of achieving much of what they ought to achieve. This suggests that, if the rational faculty were not instrumentally successful, it would not be the case that we ought to have it. So it supports the claim that, if we ought to have the rational faculty, that is because it is instrumentally successful.

On the basis of these rough arguments, let us now suppose that you ought to have the rational faculty, and that this is so for instrumental reasons. Could it be because of this that you ought to satisfy the individual requirements of rationality? If so, it would give us a good answer to the normative question. Not only would you have a reason to satisfy rational requirements, but the reason would have to be that rationality requires it. Medium normativity would be true. Rationality would be a genuine source of normativity.

By contrast, if there had been a direct instrumental argument for satisfying rational requirements, it would have established only weak normativity. The reason for satisfying each requirement would have been what could be achieved by doing so. If we are to show on instrumental grounds that rationality is normative, we therefore need the indirect argument we are now investigating, rather than a direct one.

But the indirect argument does make it necessary to add one qualification.<sup>12</sup> If rationality is normative on instrumental grounds, it will not be normative in worlds where the rational faculty is not instrumentally successful. It will not be normative in quirky worlds, therefore. This means that medium normativity as it is stated in section 1 will not be true. It will need to be weakened to something like:

<sup>12</sup> Alexander Bird and Hannes Leitgeb pointed this out to me.

necessarily except in quirky worlds, if rationality requires  $N$  to  $F$ , that fact is a reason for  $N$  to  $F$ .

## 5. Still no answer to the normative question

Sadly, this is all pie in the sky. So far as I can tell, the indirect instrumental argument does not go through. We are assuming you ought to have the rational faculty, and that this is so for instrumental reasons. It does not follow by any means of inference I can find that rationality is normative.

If you have the rational faculty, it will cause you to satisfy many individual requirements of rationality. But it does not follow by any general principle of inference that there is a reason for you to satisfy those requirements. ‘You ought to  $F$ ; if you  $F$ , your  $F$ ing will cause you to  $G$ ; so there is a reason for you to  $G$ ’ is not a valid pattern of inference.

This should be uncontroversial. Suppose you ought to take some drug to cure your serious disease. Suppose the drug has the side effect of causing you to feel unsteady on your feet. It does not follow that you have a reason to feel unsteady on your feet. To be sure, if you take the drug and as a result feel unsteady on your feet, we might say there is a reason why you feel unsteady on your feet. That would simply be another way of saying there is an explanation of why you feel unsteady on your feet. ‘The reason why’ in this context has a non-normative meaning equivalent to ‘the explanation of why’ — this is yet another sense of the promiscuous word ‘reason.’ There is no suggestion that there is a reason, in any normative sense, for you to feel unsteady on your feet. In general, normativity is not transmitted from something to a causal effect of that thing.

Still, given the premise that you ought to have the rational faculty, perhaps we might nevertheless be able to derive the conclusion that rationality is normative. The very general inference pattern I set out is invalid, but there might be some more specific way to put through the derivation. For example, perhaps we might be able to develop a valid argument using some extra premise. I can only report that I have found no way to do so.

Would a different, parallel argument be more successful?<sup>13</sup> I defined your rational faculty to be a cause of your satisfying individual

<sup>13</sup> Philip Pettit made this suggestion to me.

requirements of rationality. I could have based the argument on some different property of yours, which has a necessary, non-causal connection with satisfying the individual requirements. For example, let us say you are *fully rational* if and only if you satisfy all the requirements of rationality. Being fully rational and satisfying any particular requirement of rationality are then logically connected. Logically necessarily, if you are fully rational, you satisfy each particular requirement. Could this give us the basis of an argument?

Suppose we could demonstrate that you ought to be fully rational. Then we would be able to call on a different pattern of inference to derive the conclusion that you have a reason to satisfy a particular requirement. We could use the pattern: 'You ought to *F*; logically necessarily, if you *F*, you *G*; so there is a reason for you to *G*'. We could substitute 'be fully rational' for '*F*' and 'satisfy a particular rational requirement' for '*G*'. This would give us a different argument for the normativity of rationality.

But I see two difficulties with this parallel argument. First, the new inference pattern is questionable. Suppose you ought to buy a can of paint and decorate your kitchen. It follows by this pattern of inference that you ought to buy a can of paint. But suppose you are not going to decorate your kitchen; you have no intention of doing so, and you will not do it. Then it seems obvious that there may be no reason for you to buy a can of paint. If you are not going to decorate your kitchen, it may be entirely pointless to buy one. This example suggests the inference pattern is invalid.

However, that pattern is nevertheless defensible. Standard deontic logic even validates the stronger pattern: 'You ought to *F*; logically necessarily, if you *F*, you *G*; so you ought to *G*'. So for safety I shall not put much weight on this difficulty.

The other difficulty is more severe. I can see no satisfactory way of arguing for the premise that you ought to be fully rational. The argument is supposed to be an instrumental one, so it would depend on showing that being fully rational is instrumentally effective. How could that be shown?

To make the corresponding argument that you ought to have the rational faculty, I assumed that having this faculty is part of the best means of achieving much of what you ought to achieve. By 'best' I meant best among those means that are psychologically possible for you. Granted that assumption, it is plausible that you ought to have the rational faculty.

But being fully rational is not even psychologically possible for you; no one could be as rational as that. And once we are looking into psychologically impossible properties, being fully rational would not be the most instrumentally effective among them. To be fully rational is to satisfy each of the individual requirements of rationality. But the example of selling your car shows it is more instrumentally effective for you to fail to satisfy a few of these requirements. By doing so, you will achieve more of what you ought to achieve. Given all this, I do not see how one could argue that you ought to have the psychologically impossible property of being fully rational.

In sum, I have found no successful argument on instrumental grounds for the conclusion that rationality is normative. Neither the rational faculty nor the property of being fully rational seems to offer a sound basis for an argument. It still seems intuitively plausible to me that rationality is normative. Moreover, it seems plausible to me that this is so for a broadly instrumental reason. But I am sorry to say that I simply cannot find an argument.

## 6. Conclusion

After all this, I have been unable to show that the requirements of rationality are normative. I have failed to show even that weak normativity is true: that, necessarily, when rationality requires you to *F*, there is a reason for you to *F*. Often when rationality requires you to *F*, you have a reason to *F*, but I cannot guarantee this is necessarily so.

When we accuse someone of irrationality, we are surely criticizing her. How could we be entitled to do so if there is no reason for her to satisfy the requirements of rationality in the first place? Well, it may indeed be the cases that, necessarily, she has a reason to satisfy the requirements; I merely have not been able to show that is so. But even if she does not have a reason, there is another possible explanation of why a charge of irrationality is a criticism.

I said that, plausibly, you ought to have the rational faculty. This is for instrumental reasons. If you do not satisfy some particular rational requirement, that is evidence that you do not have this faculty, or at least that you do not have it to the highest degree. So it is evidence that you are failing to achieve something that, plausibly, you ought to achieve. That makes it grounds for justified criticism. If you fail to satisfy a particular rational requirement, that may show you have

gone wrong. But where you have gone wrong may not be in failing to satisfy the requirement. You may have gone wrong in failing to have the rational faculty, which your failure is evidence of.

In any case, my conclusion is sceptical. I can find no grounds for thinking that rationality is normative.<sup>14</sup>

John Broome  
Corpus Christi College  
Oxford OX1 4JF, UK  
john.broome@philosophy.ox.ac.uk

### References

- Broome, John. 2005. Does rationality give us reasons? *Philosophical Issues* 15: 321–37.
- Broome, John. 2006. ‘Reasoning with preferences?’ In *Preference Formation and Well-Being*, ed. by Serena Olsaretti. Cambridge: Cambridge University Press.
- Broome, John. 2007a. Wide or narrow scope? *Mind* 116: 359–70.
- Broome, John. 2007b. ‘Requirements’. In *Homage à Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz*, ed. by Toni Rønnow-Rasmussen, Björn Petersson, Jonas Josefsson and Dan Egonsson, at [www.fil.lu.se/HommageaWlodek/site/papper/BroomeJohn.pdf](http://www.fil.lu.se/HommageaWlodek/site/papper/BroomeJohn.pdf).
- Castañeda, Hector-Neri. 1968. On the logic of attributions of self-knowledge to others. *Journal of Philosophy* 65: 439–56.
- Hume, David. 1978. *A Treatise of Human Nature*. Edition by L. A. Selby-Bigge and P. H. Nidditch. Oxford: Oxford University Press.
- Johnston, Mark. 1989. Dispositional theories of value. *Proceedings of the Aristotelian Society* supp. vol. 63: 139–74.
- Kolodny, Niko. 2005. Why be rational?. *Mind* 114: 509–63.
- Smith, Michael. 2004. Humean rationality. In *The Oxford Handbook of Rationality*, ed. by Alfred R. Mele and Piers Rawling. Oxford: Oxford University Press.
- Williams, Bernard. 1981. Internal and external reasons. In *Moral Luck*. Cambridge: Cambridge University Press.

<sup>14</sup> This paper has developed out of Broome 2005. It has benefited from the discussion at the European Congress of Analytical Philosophy, Lisbon 2005 — particularly from Philip Pettit’s criticism — and from discussions at the many universities where I have subsequently presented it.

# Belief and normativity

**Pascal Engel**

University of Geneva

## **Abstract**

The thesis that mental content is normative is ambiguous and has many forms. This article deals only with the thesis that normativity is connected to our mental attitudes rather than with the content of the attitudes, and more specifically with the view that it is connected to belief. A number of writers have proposed various versions of a 'norm of truth' attached to belief. I examine various versions of this claim, and defend it against recent criticisms according to which this norm lacks normative force, that it violates the principle that 'ought implies can,' and that it is viciously circular. I defend the view that we should distinguish the statement of the objective norm and the way it is regulated, and that this distinction can answer most of the criticisms of the norm of truth for belief.

## 1. Introduction: issues about normativity

It seems to be a platitude that belief is governed by a norm of truth. Is not the point of belief to believe truths? Is not it a requirement of believing that we should not believe falsehoods? A number of philosophers, however, deny this. Although they recognise that there is an intimate connection between belief and truth, they reject the idea that this connection is normative. Indeed, they tell us, beliefs 'aim' at truth in the sense that to believe that P is to believe that P is true, or in the sense that it is the 'direction of fit' for beliefs that they should adapt to the world, but there is no more in this than a trivial fact about belief as a propositional attitude. To elevate this trivial fact to the status of a 'norm' is to transform an innocent platitude into a pompous falsehood. For there is nothing normative about believing: neither we believe with an eye fixed on the horizon of an ideal of truth nor we obey any prescription to believe the truth. On the contrary, we believe all sorts of things, some rational, some irrational, some justified, some unjustified, some true, some false, and there is no particular norm that we follow, no particular prescription that we obey and no particular sanction that we incur when we go off

track. Believing is just a natural mental state, which has certain causes and certain effects, and which answers no particular rational or normative essence. These reactions to the thesis that belief is governed by a norm of truth seem to bear the seal of common sense.

Nevertheless, the thesis that belief carries or involves a normative dimension, which is intrinsically connected to truth, seems to me both true and important. It is important for the philosophy of mind, since it is part of what makes belief specific among other attitudes and is connected to the impossibility of believing at will; for epistemology since it helps us to understand the role of belief in an analysis of knowledge; and for the philosophy of normativity since it helps us to understand the vexed question of whether there is a normative dimension, and which one, in mental content. But this thesis is not true without qualifications, and it has several versions, depending on how one understands the normative involvement.

Before trying to assess the issue of the normativity of belief, it is important to draw a rough — and necessarily incomplete — map of the general questions which arise about the notion of normativity. When one talks about norms, there are several strands. In the first place, ‘norm’ and ‘normative’ belong to a family of notions which are often not distinguished easily. In particular, are norms the same thing as rules? If norms carry a dimension of evaluation, how does one distinguish them from values? Should they be expressed in terms of deontic concepts, such as those of obligation or permission? Do all norms trade into *oughts* and *shoulds*? What is their domain of application? No one contests that there are moral norms, social norms, and aesthetic norms. But are there norms for beliefs and for mental contents? Are there epistemic norms in addition to practical and aesthetic ones? All of these issues are moot, and the concept of norm is, in many respects, a vague one. I shall not try to settle them here. Three kinds of questions, however, are prominent:

- a) *Semantical*: How should we formulate the norm for belief? It is generally agreed that the normative dimension in belief is its dimension of correctness, and that the norm for belief, if there is such a norm, is that a belief is correct if and only if it is true. But what is the relationship between this correctness condition and its application to particular beliefs? In particular, does it entail special prescriptions in the form of statements about one *ought* to believe?

- b) *Epistemological*: Given that there are specific norms for belief, how do we come to know them? How are they used when we attempt to conform to the norm? It seems to be a requirement on any norm that someone who is subject to it has to know how to conform to it. It is also an apparent requirement that if one is subject to a normative requirement one is at least able to conform to it, and to see how one can do so (*ought* implies *can*). In other words, how does the norm *regulate* the behaviour of the agents or subjects which are supposed to be subject to it? I group all such questions under the epistemological heading
- c) *Ontological*: are the norms of belief real properties of belief? If so, are they essential or derivative? In general, there are two positions relative to the ontology of moral norms in meta-ethics: one can be a cognitivist about them, and take them as objective, or one can be a non-cognitivist or an expressivist, and take them as mere expression of our psychological attitudes. Is there a parallel opposition about epistemic norms and norms of thought? There is no reason to think that there is not.

Here I shall deal mostly with the semantical and epistemological issues, and shall leave aside the ontological ones.

Several kinds of epistemic norms are said to govern belief: truth (a belief is correct if it is true), evidence (a belief is correct if it rests upon sufficient evidence), knowledge (a belief is correct if and only if it aims at knowledge), rational norms (a belief is correct if and only if it is rational). A full account of the norms for belief would need to consider all these, and would have to analyse their relations. It would also have to determine whether there is a hierarchy among these norms, and whether one of them is candidate for being more fundamental than the others are. Similar issues arise about assertion, which are, in many respects, close to those about belief, and it is interesting to consider these similarities and differences. But here I shall abstract from all these issues, and consider only the proposal that there is a basic norm for belief, namely truth. My main question is not whether truth, or another norm, is the fundamental norm for belief. It is rather this: in so far as we admit that truth is the fundamental norm for belief, in what sense is it normative? My objective here is to try to assess various versions of the view that truth is the norm for belief, to clarify them and to explain which version is, in my view, the most credible. In doing so, I shall try to answer some criticisms of the normativity of belief thesis, which have been voiced recently, in particular, by Kathrin

Glüer and Åsa Wikforss (forthcoming), Asbjörn Steglich-Petersen (2006), and Bykvist and Hattiangadi (2007).

## 2. How to formulate the truth norm?

The idea that there are conditions under which a belief is correct seems to be the most general way for characterising the normative dimension of belief. For instance, Alan Gibbard says:

‘For belief, correctness is truth. Correct belief is true belief. My belief that snow is white is correct just in case the belief is true, just in case snow is white. Correctness, now, seems normative ... The correct belief, if all this is right, seems to be the one [a subject] ought, in this sense, to have.’ (Gibbard 2005: 338–39)

From this we can derive a formulation of the norm of truth for belief:

(NT1) For any P, a belief that P is correct iff P is true.

and if we express the notion of correctness in prescriptive terms:

(NT 2) For any P, one ought to believe that P iff P.

which is the one favoured by various writers would have defended what I shall call the normative account (Wedgwood 2002, Boghossian 2003, Engel 2002, Shah 2003).

A different formulation is James’ famous declaration:

‘There are two ways of looking at our duty in the matter of opinions — ways entirely different, and yet ways about whose difference the theory of knowledge seems hitherto to have shown very little concern. We must know the truth; and we must avoid error — these are our first and great commandments as would-be-knowers; but they are not two ways of stating an identical commandment, they are separable laws.’ (James 1896)

James’ statement occurs within the context of his famous analysis of the will to believe and of the ethics of belief. James expresses himself in deontic terms — he talks about duties, commandments, and ‘musts,’ but elsewhere in his article he makes it clear that he is not simply talking of duties, but also of epistemic aims and interests, and

of epistemic values in general.<sup>1</sup> If we express the relation between truth and belief in terms of interests, it is natural to express the relevant notion of correctness in terms of our *desire* about our true belief. The ‘norm’ for belief simply becomes

(DES) We desire that we believe that P if and only if P is true (Piller 2006).

Of course, the desire in question is not simply a contingent and transient desire, but a second-order belief to have certain desires. From this formulation it is easy to move to a formulation in terms of value (given the dispositional theory of value which derives values from desires about desires):

(VAL) We value that we believe that P if and only if P.

Before trying to assess these formulations, we need to understand what the status of a norm of truth for belief is. Is it a requirement on belief as a *mental state*? Is it a property of the *contents* of our beliefs, i.e. of their propositional content? Or is it a property of our beliefs within the general context of inquiry? The question becomes particularly pressing if we consider the norm for belief in relation to other norms which are said to govern belief. Beliefs are not subject simply to a truth norm, but also to *rationality* norms. In general,

(NR1) For any P, a belief that P is correct iff it is rational.

Which we can express as a prescriptive requirement on believing what our beliefs entail:

(NR2) For any P, one ought to believe that Q if one believes that Q is entailed by P.

For instance, Frank Jackson says:

‘Someone who believes that P, and that if P then Q, *ought* to believe that Q. It is not simply that, by and large, they do believe that Q. It is that if they don’t, there is something *wrong*.’ (Jackson 2000: 101)

<sup>1</sup> For an analysis of the ambiguities in James’s article between values and duties, and between epistemic and moral obligations, see in particular Haack 1997.

Now, what is the relationship between the truth norm (NT) and the rationality norm (NR)? On the one hand, it is certainly a requirement that our beliefs are rational, but being a rational believer who has false beliefs is not a very desirable situation. On the other hand having true beliefs but being unable to see their rational connexions is not very desirable either. So the two norms seem to function together and to be on a par. But what exactly are their relations? Nick Zangwill (2005) has an interesting way of characterising the difference. He calls norms like (NT) ‘vertical requirements,’ about links between beliefs and the world, and norms like (NR) ‘horizontal requirements,’ about links between beliefs and beliefs or between beliefs and other mental states.

Now this distinction is related to another. Glüer and Wikforss (to appear) are concerned to discuss ‘normativism,’ the view according to which norms are in some sense essential or constitutive of *contents*. They distinguish two senses in which content can be said to be ‘normative:’ a) the sense in which the norms of thinking in general determine the normative character of the content or our beliefs (and of other mental states), which they call ‘content determining normativism’ CD, and b) the sense in which the norms associated to the concepts which feature in the content which engender the norms ‘content engendered normativism,’ CE). In the first sense the norms come, so to say, from outside contents because they belong to the attitudes (here belief), whereas in the second sense, the norms so to say, come from within the contents, together with the concepts which figure in them (if ‘meaning is normative,’ presumably this is true for every word or concept). CE normativism is the view most commonly attached to Kripke and to the writers who claim that the normative dimension of contents come from the meaning or concepts and from the inferential role associated to them. CD normativism is the view that the norms are associated not to concepts but to the attitudes and mental states.<sup>2</sup>

My objective here is not to deal with the problem of the ‘normativity of content’ as such.<sup>3</sup> I am concerned with the normativity of *belief as*

<sup>2</sup> A related distinction is Bilgrami’s (1992), between ‘high profile’ norms of rationality and ‘low profile’ norms attached to particular concepts and meaning.

<sup>3</sup> See among many others, Gibbard, Engel 2000, Boghossian 2003, Glüer 2000, Wikforss 2001.

*an attitude*, and therefore with the claims of CD normativism, but my purpose is not to claim that mental *content* is normative, or in what sense it is, although this issue is obviously orthogonal to the present one, which is to investigate in what sense a norm of belief can be said to govern this mental attitude. Although Glüer and Wikforss' distinction between CD normativism and CE normativism is useful, I am not sure that it is always relevant to characterise normativist theses. It is relevant if we construe meanings and concepts as inferential roles *independently* from the truth conditions, as in views which like Brandom's (1994) characterise inferential role in terms of assertion conditions and rational relations. But it is irrelevant for those normativist theories of concepts which, like Peacocke's (1992, 2004) do not divorce inferential role from truth conditions, and insist that truth-links are as important as inferential links. In this sense it is not clear that rational requirements are independent from truth requirements.<sup>4</sup>

Similar questions arise about the relationship between the truth norm (NT) and *evidential* norms. It is often said that belief is subject to a norm of evidence, as well as to a norm of truth:

(NE) A belief is correct iff it is based on appropriate evidence.

There are, however, several concepts of evidence. If we associate evidence to subjective probability, and adopt the Bayesian concept of evidence, we shall have a fairly different concept from the one that we have if we characterise beliefs as governed by a norm of truth. Presumably (NT) goes with a categorical notion of belief as full belief, whereas the Bayesian notion goes with a notion of degree of belief determined by subjective probability. The relations between the two are notoriously problematic, and this problems transfers to the relations between NT and NE.

I shall also make three questionable assumptions. First, I shall abstract from the problem of what we may call the location of norma-

<sup>4</sup> Another reason why the question whether it is belief as an attitude or the concepts within belief contents which carry the normative load may not be two different questions is that when we attribute belief to ourselves and others, we use the concept of belief. Is 'S believes that P' normative because the concept of belief figures in this attribution (CD normativism) or because the attitude of belief is normative (CE normativism)? Both, presumably. That ascriptions of content are, according to normativism, normative, is used by Steglich-Petersen (to appear) as an argument against 'normative essentialism.'

tive content — is it a property of belief or a property of the concepts which figure in beliefs? — and I shall assume that the norm for belief is attached to *belief*. Second I shall suppose that there is only *one* main norm for belief — the truth norm — and that the other norms are in some sense derivative from it. And third, I shall not try to assess the relationships between the norm of truth and these other derived or associated norms.<sup>5</sup> Each of these assumptions may be questioned, but they are independent from the kind of question that I want to raise: supposing that truth is the fundamental norm for belief in what sense can we say that it is normative at all? This question has been the focus of many objections, and it these that I want to address.

So, I shall suppose that the main norm for belief is (NT). The problem I want to address here is this: is this formulation the right one? What are the conditions for its being right? Should we revise it in the face of the objections addressed to it? The main objection which is addressed against the truth norm is the following: in what sense is NT supposed to be genuinely *normative*, *i.e.* to *regulate* and to give us any *guidance* for our beliefs? If NT does not regulate belief at all, it is reduced to an abstract and empty requirement. In other words in so far as NT is supposed to cash out the intuition that ‘beliefs aim at truth’ there is just no such aim or norm for believers, for beliefs do not have any such target (many beliefs are not formed through a concern for truth), and it is completely idealistic to claim that believers could consciously entertain NT when they believe something.

### 3. The objection from normative force

The first objection which is addressed to NT is that it lacks normative force. Certainly, the objection goes, NT expresses a general requirement on belief. It is a basic condition on rationality that one’s beliefs are true, but this condition merely tells us what our beliefs *are* and it gives us no directive about what we should do with our beliefs. Indeed, our beliefs aim at truth, and are supposed to be true if we are believers at all, but we have no choice. For there to be a norm, however, there has to be a must, a normative force, but also a normative freedom: the norm can be violated. But understood as a require-

<sup>5</sup> Elsewhere (Engel 2005) I have argued that the norm of truth is actually closely associated to the norm of knowledge, and derivative from it.

ment on belief NT cannot be violated. As Kevin Mulligan (1999) has reminded us, a norm, to be a norm but be such that it can be broken and such that the person who breaks it can be criticised or sanctioned. And as Peter Railton (1999) reminds us, the normative force or the authority of a norm or normative principle (which is supposed to constrain us) goes hand in hand with normative freedom (our freedom to break the rules). Now, the objection goes, if a norm is merely a general rational principle, such as (NT) or (NR), it only says what a belief is (perhaps for an ideal rational agent), but it has no normative force. The point is well expressed by Glüer and Wikforss:

‘The point can be put in terms of the notion of *internal relations*. The idea is that beliefs stand in basic internal relations to one another, such that being a believer in the first place requires that certain general patterns of very basic rationality are instantiated between those beliefs one has. Otherwise, it becomes unclear what the very content of those beliefs are, i.e. which beliefs it is that one has. This, also, makes it perfectly clear how beliefs differ from other cognitive attitudes, such as imaginings: Beliefs stand in various internal relations that imaginings do not. If I believe that  $p$  and that *if  $p$  then  $q$* , I have a decisive reason to believe that  $q$ , whereas imagining that  $p$  gives me no such reason.

However, to say that beliefs stand in various internal connections to one another is not to say that these connections are *normative*. On the contrary, precisely because the connections are internal or analytically necessary, they are not normative, not optional. If the connection were merely normative, it would be *possible* to violate the norm in question. That is, it would be possible to be in the one state without being in the other. This is precisely what is *impossible* if a relation between the states is internal. If the relation is internal, there is, so to speak, not enough room for any norm to enter between the two states. Of course, even if I (fully) believe that  $p$  and that *if  $p$ , then  $q$* , I can fail to draw the conclusion.’ (Glüer and Wikforss, to appear)

A good example of a theory of rational norms which is not normative is Davidson’s. Davidson talks a lot about the ‘norms of rationality’ which an interpreter of language and mind is bound to use, and takes these norms to be intrinsic to what meanings and mental contents are. But, to use his own metaphor, these norms are principles for ‘measuring the mind’ analogous to principles for measuring weight or temperatures. There are merely descriptive and offer us no guidance

at all. Timothy Schroeder (2003; see also Glüer 2001, Engel *to appear*) distinguishes in this sense two notions of ‘norm:’

- a) As categorisation or classification schemes, in the sense of general idealised principles of description;
- b) As force makers, that is as prescriptions or governance principles giving us aims to follow.

According to Schroeder, a theory of mind or a theory of content is fully normative only if it has norms in *both* senses a) and b). Otherwise, it is not normative, or is ‘normative’ only by courtesy. The normative force of a norm is this feature of it, which is such that it is susceptible to *motivate* us in doing what the norm prescribes, or, as Schroeder says, to have some sort of normative ‘oomph.’ As Schroeder rightly points out, Davidson’s theory of mind is ‘normative’ only in the first sense, and not in the second sense. Hence, it is not ‘normative:’

‘His interest in rationality is thus an interest in it only insofar as it picks out a certain set of propositional attitude clusters (those which it would be fairly rational to hold) and distinguishes them from a different set of propositional — attitude clusters (those which it would be wildly irrational to hold). The fact that the patterns exhibited by the propositional attitudes of a rational organism are normatively commanded — that there exists a force-maker for the patterns — is of no significance in Davidson’s theory.’ (Schroeder 2003)

Schroeder’s diagnosis seems to me perfectly right, and his confirmed by Davidson’s answer to those who, like me, hold that there is a norm of truth for beliefs:

‘When we say we want our beliefs to be true, we could as well say we want to be certain that they are, that the evidence for them is overwhelming, that all subsequent (observed) events will bear them out, that everyone will come to agree with us. It makes no sense to ask for more. But I do not think it adds anything to say that truth is a goal, of science or anything else. We do not aim at truth but at honest justification. Truth is not, in my opinion, a norm.’ (Davidson 1998, in reply to Engel 1998)

The objection from normative force therefore says that if *all there is to the norm of truth for belief* that a belief is correct if and only if it is true, this ‘norm’ is no norm at all, and his perfectly trivial or shallow.

What can we answer to this objection? There is something correct in it, which is that a mere categorisation scheme cannot be a norm, unless it is susceptible to have a normative force (I shall below have to qualify this). And for the norm to have force, we must be able to see in some way how it can guide our conduct, or, to use Shah's and Velleman's phrase (2005) to *regulate*, our conduct or our mental states.

Where, however, the objection goes wrong, is that from the fact that a norm is a categorisation scheme, it concludes that it cannot have normative force. But there is no reason why we should not distinguish two levels:

- (a) The statement of the norm (the kind of analytic or constitutive or essential truth about belief it expresses);
- (b) How the norm is regulated (its regulation).

It is one thing to say what the norm is, that is what kind of truth (analytic, or essential) is expressed by it, and it is another thing to say how the norm is regulated, and realised in the psychology of the believers. In this sense, (NT) expresses a basic truth, perhaps conceptual, perhaps essential (depending upon the kind of ontological status one grants to normative judgments or principles).<sup>6</sup> But the question of how the norm is regulated is another matter. In particular, we cannot simply *read off* the regulation from the basic truth. And the fact that we simply state the rational or normative principle (NT) does not imply that the agent is necessarily motivated by the norm (a point familiar from Lewis Carroll's story of Achilles and the Tortoise).<sup>7</sup> The distinction between the statement of the norm and the conditions of its regulation is reminiscent of the distinction between the formulation of a general norm on the one hand, and its conditions of application, or between the law and its decrees of application.<sup>8</sup>

<sup>6</sup> I said above that I would not deal with these ontological issues. One can be a conceptualist about the norms (it is a feature of our concepts), an expressivist (it is a feature of our psychological attitudes) or a realist-cognitivist (it expresses a real essence). See Wedgwood 2006, Zangwill 2005. Although I do not need to enter into these ontological issues, I believe, like Wedgwood, that the proper defence of the normative account needs a form of cognitivism.

<sup>7</sup> See Engel 2005, 2007 on Carroll's paradox.

<sup>8</sup> Several people have pointed out to me that the distinction is reminiscent of the distinction in moral theory of the general principle of utilitarianism and the particu-

So in a sense, I grant the objection from normative force. Simply stating a rational principle like NT does not tell us how it is implemented in a believer's psychology. Still, it would be wrong to say that there is simply no relation between the principle and the regulation. There must be *some* relation between the principle and the regulation. Although the normative truth is necessarily independent from the way it is regulated, there has to be a connection between the two. What kind of connexion? In the first place it must be the case that the agent who violates the norm can be criticised for doing so. 'Criticised' does not necessarily mean: sanctioned or castigated. If I violate the norm that my beliefs be true, by entertaining false beliefs, I am not going to be beaten with sticks. But someone can certainly criticise me for having held false beliefs, and if I myself realise this, I ought, normally to change my beliefs. I shall say more on this in section 5 below, but in this respect it is not right to say that NT is such that it cannot be violated, since it applies to rational agents and that they cannot fail to conform to the norm. In the second place, the norm as an analytic or essential truth must be such that it can be obeyed. In other words, it must be such that the *ought* that it contains must imply *can*. I can be under no obligation to conform to a norm to which no human being can conform. And this condition on norms is the one which prompts the other set of objections against (NT).

#### 4. The objection from '*ought*' implies '*can*'

According to this objection, the norm for belief is impossible to satisfy, because it imposes to believers constraints which are impossible to satisfy. It is not clear, however, that such an objection is always correct, for the fact that an agent cannot perform the action which a norm prescribes does not imply that he is not under the obligation to perform the action. In other words it is not clear that the '*ought* implies *can*' principle is always correct.<sup>9</sup> Let us come back to our initial formulation:

(NT2) For any P, one ought to believe that P iff P.

lar rules by which it is implemented (thanks to Tom Stoneham and Klemens Kappel for this).

<sup>9</sup> For objections, see for instance Stocker 1990, Ogien 2003.

It is more complex that this statement seems to say. In the first place, John Broome has attracted our attention to differences of scope for what he calls 'normative requirements' (Broome 1999). (NT) can be read with a narrow or with a wide scope.

On the *narrow* reading, it says:

(NT2a) For any  $S, P$ :  $S$  ought to (believe that  $P$ ) if and only if  $P$  is true.

On the *wide* reading, it says:

(NT 2b) For any  $S, P$ :  $S$  ought to (believe that  $P$  if and only if  $P$  is true).

The difference might not be apparent at this stage, but at first sight, the narrow scope reading seems to be the most natural one: the left hand side of the biconditional tells us what condition we must respect if we believe that  $P$ : to believe it iff it is true. The wide reading on the other hand tells us that we have to obey the whole biconditional (believe that  $P$  iff it is true). So let us, for the moment examine the narrow scope reading.

The narrow scope reading can itself be broken into two conditionals depending upon whether one reads it from right to left or left to right:

(NT2a\* ) For any  $P$ , if  $P$  is true then  $S$  ought to believe that  $P$ .

(NT2a\*\*) For any  $P$ ,  $S$  ought to believe that  $P$  only if  $P$ .

Suppose we interpret (NT2) through the first reading (NT2a\*). A common objection to it (Haack 1997, Engel 2002, Boghossian 2003, Sosa 2008) is that on such a reading the norm is unsatisfiable or useless. Unsatisfiable: there are infinitely many truths, and by logic infinitely many truths equivalent to a given truth, which not only no one cares to believe, but also that no one could possibly believe.

(NT2a\*) is also unsuitable as a norm for belief because there are plenty of trivial or uninteresting beliefs that are true, but that no one would, at least in usual circumstances, care to believe. For instance, that there are presently 36547 blades of grass on this corner of my garden is not something which I care to believe, although I could do so, if I cared to gather this truth. So it is not true, says the objection, that we have to believe any truth whatsoever. So (NT2a\*) violates the *ought* implies *can* constraint or it is useless.

For this reason a number of writers prefer the (N2a\*\*) formulation (Boghossian 2003):

(NT 2a\*\*) For any P, S ought to believe that P only if P.

or

(NT 2a\*\*\*) For any P, if S ought to believe that P, then P is true.

Now, Bykvist and Hattiangadi 2007 have argued that it is not clear it works either.

(NT2a\*\*), according to them, it ‘does not capture the thought that the truth is what you ought to believe, since (2a\*\*) is not normative in any interesting sense — it does not imply that a subject is under any obligation under any circumstances whatsoever.’ Bykvist and Hattiangadi write:

‘Obviously, if  $p$  is true, nothing whatsoever follows from (NT 2a\*\*) about what  $S$  ought to believe. Less obviously perhaps, if  $p$  is false, nothing whatsoever follows about what  $S$  ought to believe. For, if  $p$  is false, it only follows that it is not the case that  $S$  ought to believe that  $p$ . It does not follow, from the falsity of  $p$ , that  $S$  ought not to believe that  $p$ . There is an important difference between ‘it is not the case that  $S$  ought to believe that  $p$ ’ and ‘ $S$  ought not to believe that  $p$ ’ — the former states that  $S$  lacks an obligation to believe that  $p$  and the latter states that  $S$  has an obligation not to believe that  $p$ . The former is compatible with it being permissible for  $S$  to believe that  $p$ , while the latter is incompatible with its being permissible for  $S$  to believe that  $p$ . Hence, whether  $p$  is true or false, (NT2a\*\*) does not tell  $S$  what to believe.’ (Bykvist and Hattiangadi 2007)

Now, this objection has force only if the proposition  $p$  is true or false independently of what the thinker takes it to be. In other words (NT2a\*\*) makes sense when the think considers the proposition  $p$ , and asks himself whether it is true. I agree with Bykvist and Hattiangadi that nothing follows from (NT2a\*\*) when the agent is not aware in any sense of the proposition. And actually they report the suggestion by Wedgwood that (NT2A\*\*) that we should replace it by:

(NT 3) For any  $S$ ,  $P$ : if  $S$  considers whether  $P$ , then  $S$  ought to (believe that  $P$ ) if and only if  $P$  is true.

To this Bykvist and Hattiangadi point out that there are some sentences which act as what Sorensen (1988) calls ‘blindspots,’ i.e. as truths such as when we consider them we cannot satisfy the requirement of believing them, such as:

It is raining and nobody believes that it is raining;  
There are no believers.

They remark that we could reformulate the condition (NT2a\*) such that it is restricted only to believable truths:

(NT3a) For any  $S, P$ : if  $S$  considers whether  $P$ , and  $P$  is truly believable, then  $S$  ought to (believe that  $P$ ) if and only if  $P$  is true.

But then this seems to trivialise the requirement, and to say only: if  $P$  is true and believable you ought to believe that  $P$ .

I do not find this objection to (NT3) very convincing. For certainly a requirement on  $P$  in (NT3) is that  $S$  actually *understands*  $P$ , and it is not clear that the blindspot sentences in question can be understood. Moreover, the question of their truth can arise. And in so far that it can arise, the norm is in place. I shall come back below to the sense of (NT3) for the regulation of truth.

Now what about the wide scope reading of (NT2)? Remember that it says:

(NT2b) For any  $S, P$ :  $S$  ought to (believe that  $P$  if and only if  $P$  is true).

As Bykvist and Hattiangadi comment, (NT2b) tells you that there are two combinations that will satisfy the requirement: either you believe that  $p$  and  $p$  is true, or it is not the case that you believe that  $p$  and  $p$  is false. At the same time, it tells you that there are two combinations that you ought to avoid: either you believe that  $p$  and  $p$  is false, or it is not the case that you believe that  $p$  and  $p$  is true.

The advantage of (NT2b) is that it is not clearly objectionable as (NT2a) is. For, (NT2b) cannot be broken down into the conditionals (NT2a\*) and (NT2a\*\*), for in those conditionals, the ‘ought’ took narrow scope. But now the problem, raised by John Broome about wide scope rationality requirements or norms is that we can’t detach.

(NT2b) does not capture the intuition that the truth is what one ought to believe, or that a false belief is faulty or defective. Broome

remarks that when we have a wide scope formulation of a *modus ponens* kind of argument of the form:

You ought (if you believe that  $p$  and believe that  $p$  implies  $q$ , believe that  $q$ );  
 And that you believe the antecedent of what is in the scope of the ‘ought’;  
 You believe that  $p$  and believe that  $p$  implies  $q$ .

The inference to

You ought to believe that  $q$ .

does not go through.

Transposing now to the (NT2b) case, the same non detachment phenomenon appears. As Bykvist and Hattiangadi say, ‘The reason is that what (NT2b) enjoins are *combinations*: the combination of your believing that  $p$  with its being true that  $p$  and the combination of its being false that  $p$  and your not believing that  $p$ . Because the ‘ought’ takes wide scope, one cannot detach from (NT2b) that you ought to believe that  $p$ , even when  $p$  is true.’

I agree with them that this makes (NT2b) unsuitable for being the norm for belief.<sup>10</sup> But the narrow scope reading and NT2a\*\* stands.

## 5. Truth and epistemic interests

A third objection raised against NT is that it does not capture our *interest* for truth Piller (2006) argues that what he calls the ‘standard view’ (NT) is wrong if we formulate it in desire terms:

(DES) We desire that we believe that  $P$  iff  $P$  is true.

or:

DES ( $BP \leftrightarrow P$ ).

This, like (NT2), can be decomposed into two conditionals:

<sup>10</sup> Some writers, in particular Kolodny 2005, have accepted wide scope requirements on rationality.

- (i) DES ( $P \rightarrow BP$ )
- (ii) DES ( $BP \rightarrow P$ )

Now Piller claims that (ii), which is the counterpart of (NT2A\*\*) in desire terms, is implausible, because we can derive from it the implausible consequence that if someone believes that P, he desires that P, through the plausible ‘transition principle’ that if someone desires that if A then B, and that A is the case, then she is rationally required to desire that B.

(Des ( $A \rightarrow B$ ) &  $A \rightarrow$  Des B).

To take one of Piller’s examples: I want that if Jim does not get the post, then John should, and I hear that the appointment committee has already eliminated Jim, it follows that I hope that John will get it. Applying this to (ii) we get:

- (1) Des ( $B P \rightarrow P$ ) [ii]
- (2) Des ( $A \rightarrow B$ ) &  $A \rightarrow$  Des B [transition principle]
- (3) Bel  $P \rightarrow$  Des P

and (3) is certainly absurd: wanting that if A then B and noticing B certain does not commit me to want B.

As Kappel (to appear) has remarked, however, it is not clear that the desire formulation leads us to such paradoxical claims. If we contrapose we get the following from (1):

(4) Des ( $\text{not-}P \rightarrow \text{not } B p$ ).

And from this we may plausibly infer (with the help of (2)):

(5)  $\text{not-}P \rightarrow$  Des ( $\text{not Bel } P$ ).

which makes sense of something similar to what (NT2A\*) expresses: we desire to avoid error, i.e., not to believe that P if P is false.

Even if we can agree with Piller that the desire formulation of (NT) is problematic, all it shows is that the proper formulation of (NT) may not be one in terms of desires like (DES), but the normative one. I quite agree with Clemens Kappel that:

‘The general lesson to be learned from this is that it is a mistake to try to capture our epistemic interests and commitments in terms of desires. There are senses in which if P, you ought to believe that p, and senses in which, if you believe P, then P should be true, but neither are captured in terms of ordinary desires’ (Kappel, *to appear*).

Piller claims that NT in its standard formulation implies that we want ‘the truth and nothing but the truth.’ He points out in his account of our interest in truth that this interest is not pure, and can coexist, or can be overridden, by our interests. This is similar to a common objection against taking truth as a goal of inquiry: we transform truth into a goddess. But there is no need to defend this sort of view to have norm of truth like (NT). The fact that our beliefs have side effects, or that we might want to believe certain things does not in any way abolish the distinction between our reasons for belief (our epistemic reasons) and our reasons for wanting to believe (which have nothing to do with an interest for truth). This is what the norm of truth is about. The norm of truth is not a *truth goal*, reflecting our interests and our desires. It is wrong to interpret the claim that one ought to have true beliefs and avoid having false beliefs as saying that we have a concern for truth for truth’s sake. On the contrary, this claim is a claim about the regulation of our beliefs, and about their minimal epistemic regulation. This is what the last section is about.

## 6. Truth and the regulation of belief

The specificity of the regulation problem has been well isolated by Railton (1994) Velleman (2000) and Shah (2003): if a norm of truth for correct belief is in place, how can it actually guide our believings, without being either idle or the expression of a requirement too strong to be followed by any human agent?

As it has been suggested above about (NT3), NT makes most sense when a subject is considering her beliefs and asks herself the question ‘do I believe that P?’ in the context of a deliberation about her beliefs. There are, however, two ways of understanding this.

The first one is the *intentional* or *teleological* account, which takes seriously the metaphor the ‘belief aims at truth:’ to believe that P is to have the conscious aim of regarding P as true if and only if it is true. On this view, the regulation of NT is done through a conscious, intentional mental act of the believer. Velleman (2000) who proposes

this account, allows that the teleological aiming at truth can be accounted, for those of our beliefs which are not conscious or explicit, by a teleological mechanism embedded in the believer's cognitive system. But even in this hypothesis believing is a matter of having a certain *goal*.

The main objections for the teleological account are these (Shah 2003, Engel 2005a). In the first place the teleological account fits only those beliefs which are consciously entertained and reflexive, and does not account for those which are not directed at truth, but at other aims, such as comforting the believer (e.g. cognitive dissonance, wishful thinking and all such 'irrational' believings. Even if we consider the non conscious beliefs, there is no reason to suppose that they are governed by a truth aim. In the second place, the teleological account represents believing as directed — consciously or not — towards a goal, truth. But we have seen that this idea, which goes along with the analysis of the norm of truth in terms of desire, misrepresents the regulation of belief. It is not at all clear that belief has an aim in the sense in which stamp collecting or any other intentional activity has one (Owens 2003). As a result of these tensions, the teleological account is caught into what Shah (2003) calls the 'teleological dilemma:'

'one horn, the teleologist must allow the disposition that constitutes aiming at truth to be so weak as to allow paradigm cases in which beliefs are caused by such non-evidential processes as wishful thinking, in which case he cannot capture the exclusive role of evidence in one particular type of belief-forming process, reasoning. On the other horn, in order to account for the exclusive role of evidence in reasoning about what to believe, the teleologist must strengthen the disposition that constitutes aiming at truth so that it excludes the influence of non-truth-regarding considerations from such reasoning. However, by strengthening the truth-aimed disposition, the teleologist cannot accommodate the cases of wishful thinking, in which non-evidential factors clearly exercise influence over belief.' (Shah 2003: 461)

Instead of the teleological account, Shah and Velleman have proposed what they call the transparency account, which analyses the process of 'doxastic deliberation' not in terms of an intentional mental act, but in terms of a simple recognition of the truth of the belief.

Transparency (Evans 1982, Moran 2001) is a phenomenon occurring in such processes, namely, the fact that whenever one asks one-

self whether to believe that  $p$ , one must immediately recognize that this question is settled by, and only by, answering the seemingly different question whether  $p$  is true. When our beliefs are in this sense transparent, i.e., to paraphrase Gareth Evans, when we direct our minds not to her beliefs, but to the world itself, no intentional aim is present. We recognise directly that we have the beliefs by considering their truth. The step is immediate and not inferential.

The transparency account allows us to understand how a normative truth about belief, to the effect that believing  $p$  is correct if and only if  $p$  is true, can explain transparency in doxastic deliberation. For in asking oneself *whether to believe that  $p$* , one applies the concept of belief. If NT is a conceptual truth about belief, then it is a constitutive feature of the concept of belief that the correctness of believing  $p$  is settled by settling the question whether  $p$  is true. So applying the concept of belief in forming a belief thus involves applying the correctness norm to one's own belief-formation.

The transparency account also explains the difference between reasons to believe and reasons for wanting to believe. One can want to believe that  $P$  without considering (indeed trying to bracket) whether  $P$  is true, but one cannot believe that  $P$  in the deliberative sense of considering whether  $P$  without asking oneself whether  $P$  is true.

The transparency account, however, seems to imply that

‘the motivation stemming from the thought that true beliefs are correct has to be so strong, if it is to do the desired explanatory work, that it is implausible to regard it as motivation stemming from acceptance of a norm at all’ (Steglich Petersen 2006).

The point is that the relation between the norm and its regulation becomes now so intrinsic that it cannot be normative: a norm which necessarily motivates does not motivate at all. This objection is very similar to the one from normative force above. As Steglich Petersen says:

‘If transparency is produced by the norm of belief, this norm motivates one necessarily and inescapably to act in accordance with it. The transparency is immediate, and does not involve an intermediary question about whether to conform to the norm for belief; the norm is thus unlike norms such as the one governing promising. It is thus doubtful whether a consideration which necessitates motivation should be considered a normative consideration at all.’

I do not see, however, why the internal relationship between the norm and its regulation which the transparency account introduces implies that the norm *necessarily* motivates us. It does not motivate in the many cases where we are not self conscious in this way, and many cases where we simply disregard the norm. Just as cases of *akrasia* or *acedia* can arise where the agent considers the norm but does not follow it, cases where the norm of truth is considered by the agent but is not followed can arise. One could analyse self deception along these lines. In this respect we can break the norm, or fail to conform to it. Does it motivate us in the conscious case? Steglich-Petersen objects that in this case too the transparency account implies that we are necessarily motivated. But I do not see why the norm, as I understand it here, necessarily motivates us. As I understand it, the norm is a constitutive principle implicit in our understanding of what a belief is. This principle may not be always present to our mind, and even when it is present, we need not follow it ( I grant that there can be 'epistemic akrasia' and other such internal inconsistencies within a subject <sup>11</sup>). So it is not clear that there is any necessary connexion between the expression of the norm as a constitutive feature of belief, and its regulation.

The transparency account of NT that it applies both to the cases of conscious deliberations about beliefs. and to non conscious cases. We can associate the normative account of belief to a set of rational dispositions of the believer, which can, in a number of cases, fail to be triggered (Wedgwood 2007).

The transparency account of truth regulation allows us also to answer an objection formulated by Glüer and Wikforss Seeing correctly the difference between the objective norm for belief (NT) and the way it regulates our believing through subjective norms, they write:

'The question is how such norms [subjective] *guide* our actions. Being guided by a norm such as 'buy low and sell high' requires having some beliefs about the market. For instance, if I believe that the market is at a low (and I intend to follow this rule plus have the required background beliefs and desires) I will buy. However, applying this to N1 [ NT2a\*] we get: If you hold *p* true, then you ought to hold *p* true. It is rather obvious that no guidance can be had from this. The trouble with N1, there-

<sup>11</sup> I have defended this view in 'Akrasia Pratique et Akrasia Épistémique,' *Le Philosophique* 19, 'L'Action,' 2007.

fore, is not that it is an objective norm, but that it cannot guide our belief formation and hence is not a norm for belief.'

The point is that the prescription 'Believe that P only P is true' is void, since in order to obey it, one must already believe P. This objection would make sense if the correctness condition for belief were a prescription such that, in order to obey it, a precondition (to believe that P) had to be satisfied. But if the regulative condition is conceived in the transparency sense, there is no circularity here: to ask oneself whether to believe that P and asking oneself whether P is correct are one and the same thing as asking oneself whether P is true.<sup>12</sup>

The transparency account of truth regulation allows us also to answer an objection formulated by Glüer and Wikforss seeing correctly the difference between the objective norm for belief (NT) and the way it regulates our believing through subjective norms, they write:

'The question is how such norms [subjective] *guide* our actions. Being guided by a norm such as 'buy low and sell high' requires having some beliefs about the market. For instance, if I believe that the market is at a low (and I intend to follow this rule plus have the required background beliefs and desires) I will buy. However, applying this to N1 [NT2a\*] we get: If you hold *p* true, then you ought to hold *p* true. It is rather obvious that no guidance can be had from this. The trouble with N1, therefore, is not that it is an objective norm, but that it cannot guide our belief formation and hence is not a norm for belief.'

The point is that the prescription 'Believe that P only if P is true' is void, since in order to obey it, one must already believe P. This objection would make sense if the correctness condition for belief were a prescription such that, in order to obey it, a precondition (to believe that P) had to be satisfied. But if the regulative condition is conceived in the transparency sense, there is no circularity here: to ask oneself

<sup>12</sup> Steglich-Petersen also objects to the transparency account of NT that it applies only to the cases of conscious deliberations about beliefs. But I do not see why it does not apply to other cases as well. We can associate the normative account of belief to a set of rational dispositions of the believer, which can, in a number of cases, fail to be triggered (Wedgwood 2007).

whether to believe that P and asking oneself whether P is correct are one and the same thing as asking oneself whether P is true.<sup>13</sup>

## 7. Conclusion

I conclude, therefore, that, properly understood, through distinguishing the truth expressed by the norm for belief and its regulation, and by having a proper account of the regulation of belief, the normative account of the correctness condition for belief stands and that the objections from normative force and from the unsatisfiability of the norm can be answered. Many issues are still unsettled, such as the consequences that this conclusion has for the normativity of content in general, and for the ontology of norms. But I am confident that we can raise these issues, which are left open, by presupposing that the normative account is correct.<sup>14</sup>

Pascal Engel  
 Département de Philosophie, Faculté des Lettres  
 Université de Genève  
 2, Rue de la Candolle (4e étage)  
 CH 1211 Genève, Switzerland  
 Pascal.Engel@lettres.unige.ch

## References

- Bilgrami, A. 1992. *Belief and Meaning*. Oxford: Blackwell.  
 Boghossian, P. 2003. The normativity of content. *Philosophical Issues* 13: 31–45.

<sup>13</sup> Glüer and Wikforss write about Shah's transparency condition. Instead, Shah argues, we have to say that there is an *internal relation* between discovering that *p* is true and believing that *p*, and this requires accepting that N1 is a condition of possessing the concept of belief. Shah is surely right that there is an internal relation here, one that does not depend on the aim of believing what is true. However, precisely because of this there is no room for a norm. Shah, it might be said, falls prey to the very same objection that he levels against Velleman: His account leaves open the possibility that one may discover that *p* is true, without the question of whether to believe that *p* being settled. But this objection can only be made if one understands the norm of truth as external to the activity of believing.

<sup>14</sup> I would like to thank, for their comments on this text and its previous versions: John Skorupski, Kathrin Glüer, Asa Wikforss, Asbjorn Steglich-Pedersen, Ralph Wedgwood, Kevin Mulligan, Julien Dutant, Anne Meylan, David Owens and Klemens Kappel.

- Brandom, R. 1994. *Making it Explicit*. Harvard: Harvard University Press.
- Broome, J. 1999. Normative Requirements. *Ratio*, repr. in Dancy 2000.
- Bykvist, K. & Hattiangadi, A. 2007. Does *Thought* Imply *Ought*? *Analysis* 67 (296): 277–285.
- Dancy, J., ed. 2000. *Normativity*. Oxford: Blackwell.
- Davidson, D. 1998. 'Reply to Pascal Engel,' in L. Hahn, ed. *The Philosophy of Donald Davidson*, 460–1.
- Engel, P. 2000. Wherein lies the normative dimension in mental content? *Philosophical Studies*. 2000, 100, 3, 305–321
- Engel, P. 2002. *Truth Bucks*: Acumen.
- Engel, P. 2005. Truth and the aim of Belief, in D. Gillies, ed. *Laws and Models in Science*, London: King's College Publications, 79–99.
- Engel, P. 2006. Epistemic norms and rationality,' in W. Strawinski, M. Grygianca, & A. Brodek, eds., *Mysli o Języku, nauce I wartosciach, Ksiega ofiarowana Jackowi Juliużowi Jadakiemu*, Warszawa: Semper, 355–370.
- Engel, P. (to appear) Davidson on epistemic norms, in C. Amoretti & N. Vassalo, *Davidson: Language, Mind and Interpretation*. Munich: Ontos Verlag.
- Evans, G. 1982. *The Varieties of Reference*. Oxford: Oxford University Press.
- Gibbard, A. 2003. Thoughts and Norms. *Philosophical Issues* 13: 83–98.
- Glüer K, 2001. Dreams and Nightmares. Conventions, Norms, and Meaning in Davidson's Philosophy of Language, in Kotatko, P. Pagin, G. Segal, eds., 2001: 53–74.
- Glüer, K & Wikforss, A. (to appear) Against Content Normativity.
- Haack S. 1997. The Ethics of Belief Reconsidered, in L. Hahn, *The Philosophy of Roderick Chisholm*. Open Court, La sale, Ill., repr. in Zagzebski and Fairweather 2000.
- Heal, J. 1986. The disinterested Search for Truth. *Proceedings of the Aristotelian Society* 97–108.
- Jackson, F. 2000. Cognitivism, Normativity, Belief, in Dancy 2000.
- James, W. 1896. *The Will to Believe*, reed. New York Dover Books, 1956.
- Kappel, K. (to appear) Comments on Piller.
- Kolodny, N. 2005. Why be Rational? *Mind* 114:509–562.
- Kotatko, P. Pagin, P. & Segal, G., eds. 2001. *Interpreting Davidson*, Stanford: CSLI Publications.
- Moran, R. 2001. *Authority and Estrangement*. Princeton: Princeton University Press.
- Mulligan, K. 1999. Justification, Rule Breaking and the Mind. *Proceedings of the Aristotelian Society* 99 (2):123–139.
- Ogien, R. 2003. *Le rasoir de Kant*. L'Eclat, Paris, Tel Haviv.
- Owens, D. 2003 Does Belief Have an Aim? *Philosophical Studies* 115, No. 3.
- Peacocke, C. 1992. *A Study of Concepts*. Cambridge, Mass: MIT Press.

- Piller, C. (to appear) Desiring the Truth and Nothing but the Truth, in A.Haddock, A. Millar and D. Pritchard, eds., *Epistemic Value*, Oxford: Oxford University Press.
- Railton, P. 1994. Truth, Reason, and the Regulation of Belief, in Engel, P. 2000 ed. *Believing and Accepting*, Dordrecht, Kluwer repr in Railton 2003.
- Railton, P. 1999 Normative Force and Normative Freedom, in Dancy 2000 and in his *Rules, Values and Norms*, Cambridge, Cambridge University Press, 2003: 179–208.
- Schroeder, T. Davidson's Theory of Mind is Non Normative, *Philosophers' Imprint* 3 (1): 1–14.
- Shah, N. 2003. How Truth Regulates Belief. *Philosophical Review*: 113.
- Shah, N. and Velleman, D. 2005. Doxastic deliberation. *Philosophical Review* 114, 2: 497-534.
- Sorensen, R. 1988. *Blindspots*. Oxford: Oxford University Press.
- Sosa, E. 2007. *A Virtue Epistemology*, Oxford: Oxford University Press.
- Steglich-Petersen, A. 2006. The Aim of Belief: no Norm Needed. *The Philosophical Quarterly* 56, 225: 500-516.
- Stocker, M. 1990. *Plural and Conflicting Values*. Oxford: Oxford University Press.
- Velleman, D. 2000. On the Aim of Belief, in his *The Possibility of Practical Reason*, Oxford: Oxford University Press.
- Wedgwood, R. 2002. The Aim of Belief. *Philosophical Perspectives* 16: 276–297.
- Wedgwood, R. 2007 Normativism Defended, in Cohen, J. and McLaughlin, J. eds. *Contemporary Debates in Philosophy of Mind*, Oxford: Blackwell.
- Wedgwood, R. 2007. *The Nature of Normativity*. Oxford: Oxford University Press.
- Wikforss, A. 2001. Semantic Normativity. *Philosophical Studies* 102: 203–226.
- Zangwill, N. 2005. The Normativity of the Mental. *Philosophical Explorations* 8, 1, March: 1–19.

# Intentionality, knowledge and formal objects

**Kevin Mulligan**  
Université de Genève

## **Abstract**

What is the relation between the intentionality of states and attitudes which can miss their mark, such as belief and desire, and the intentionality of acts, states and attitudes which cannot miss their mark, such as the different types of knowledge and simple seeing? Two theories of the first type of intentionality, the theory of correctness conditions and the theory of satisfaction conditions, are compared. It is argued that knowledge always involves knowledge of formal objects such as facts and values, that emotions are reactions to (apparently) known values and that beliefs are reactions to known or apparently known facts or to the objects of relational states.

## 1. Intentionality<sup>1</sup>

Any philosophy of intentionality, of the property peculiar to mental acts, states and activities of being 'directed' towards or about something, should contain many chapters. It should provide an account of the different mental acts, states and activities. It should provide an analysis of the relations and other ties hiding behind the metaphor of directedness. And it should provide an account of the sorts of things mental acts, states and activities are directed towards. A philosophy of intentionality should, further, tell us about the intentionality of all the main types of mental states, acts and activities. It should tell us, at the very least, about the intentionality of

acquaintance, admiration, attention, belief that, belief in, certainty, choice, deliberation, desire, doubt, expectation, hate, hope, imagination, judge-

<sup>1</sup> An earlier version of this paper appeared in an electronic Festschrift: *Homage à Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowicz*. Ed. T. Rønnow-Rasmussen, B. Petersson, J. Josefsson & D. Egonsson, 2007 <<http://www.fil.lu.se/hommageawlodek>>.

ment, knowledge, love, meaning that p, memory, perception, preference, regret, shame, sympathy, striving, supposition, time-consciousness, trust, uncertainty, understanding, vision, willing and wishes

and not limit itself to, say, the intentionality of belief and desire. A philosophy of intentionality should provide an account of the difference between collective or shared intentionality, for example that of shared shame or shared certainties, and solitary intentionality, such as that of judgement. It should also tell us how the intentionality of different acts and states hang together, how, for example, the intentionality of emotions is related to the intentionality of perception and belief, how the intentionality of visual imagination is related to that of vision, a desideratum which cannot be met by philosophies of intentionality which consider only a handful of types of mental states and acts.

Some types of acts, activities and states are such that tokens of the type may miss their mark, go wrong or be unsuccessful. Memory misleads, suspicions are unfounded, beliefs turn out to be incorrect. Mental states which can go wrong contrast strikingly with types of states and acts such as knowledge, seeing and perception which cannot miss their mark. In what follows, I explore the relations between states and acts which can go wrong, on the one hand, and knowledge, seeing and perception, on the other hand. I consider two accounts of states and acts which can miss their mark, the theory of satisfaction conditions and the theory of correctness conditions (§2). I then consider two objections to the theory of correctness conditions: correctness conditions are not truth-evaluable and one central type of correctness condition, for judgement and belief, is superfluous. I then argue that one plausible account of the intentionality of knowledge gives us some reason to reject the objections to the very idea of correctness conditions. The preferred account comes in two parts, an account of knowledge and its relation to objects and facts (§3) and an account of knowledge of value (§4). I then argue that the intentionality of knowledge, understood in the preferred way, is more fundamental than the intentionality of acts and states which can go wrong (§5).

## 2. Correctness conditions vs. satisfaction conditions

Consider those states and acts which may miss their mark. One account of what it is for states and acts to miss their target, is the theory of satisfaction conditions. This theory is part of an account of

what it is for such states and acts to enjoy the property of intentionality. A simplified version of Searle's account of the satisfaction conditions for belief is that a belief that  $p$  is satisfied only if  $p$ . Similarly, a simplified version of the satisfaction condition for  $x$ 's desire to  $F$  is:

Cause ( $x$ 's desire that  $Fx$ ,  $Fx$ )<sup>2</sup>

Another account of what it is for states and acts to miss their mark is the theory of correctness conditions, a theory associated with Husserl. In the following table, the sentences on the right express putative correctness conditions for the psychological reports on the left:

$x$ desires to $F$	$x$ ought to $F$ ('Tunsollen')
$x$ wishes that $p$	It ought to be the case that $p$ ('Seinsollen')
$x$ values $y$	$y$ is valuable
$x$ admires $y$	$y$ is admirable
$x$ fears $y$	$y$ is dangerous
$x$ 'values' that $p$	That $p$ is valuable, is a 'Wertverhalt'
$x$ regrets that $p$	It is regrettable that $p$
$x$ is ashamed that $p$	It is shameful that $p$
$x$ prefers $y$ to $z$	$y$ is better than $z$
$x$ judges (believes) that $p$	The state of affairs that $p$ obtains The proposition that $p$ is true
$x$ conjectures that $p$	It is probable that $p$
$x$ has an interrogative attitude toward $p$	It is questionable whether $p$
$x$ doubts whether $p$	It is doubtful whether $p$
$x$ is certain that $p$	It is certain that $p$

<sup>2</sup> Cf. Searle 1983: 13 (belief), ch. 3 (desire).

Kenny distinguishes between the material and formal objects of mental attitudes and argues that emotional and other attitudes have formal objects.<sup>3</sup> He attributes to the medieval schoolmen the view that the formal object of fear is a future evil, of envy another's good. Similarly, one might say that propositions, states of affairs, truth and obtaining, values, norms and probabilities are the formal objects of different attitudes, states and acts. Then fear will have, for example, a dog as its material or proper object and a future evil or disvalue as its formal object. A conjecture that the dog will attack has as its material object the dog and the probability that it will attack as its formal object. One unfortunate but perhaps harmless feature of this terminology is that it runs together a narrow and a wide sense of 'object'. In the narrow sense, propositions and states of affairs are the formal objects of judgements and beliefs. Truth, obtaining, oughtness and value, on the other hand, are not formal objects, in the narrow sense of the word, of judgements, beliefs, desires and emotions.

It is a peculiarity of judgement (and of belief and convictions) that it seems to have two correctness conditions: the truth of propositions and the obtaining of states of affairs or the existence of facts. We shall return to this feature of judgement and belief.

How should the theory of correctness conditions be formulated? Presumably, as follows:

- x desires to F  $\rightarrow$  (x correctly (rightly) desires to F iff x ought to F).
- x wishes that p  $\rightarrow$  (x correctly (rightly) wishes that p iff it ought to be the case that p).
- x conjectures that p  $\rightarrow$  (x correctly (rightly) conjectures that p iff it is probable that p).

And so on. But the theory of correctness conditions contains an extra type of claim, an explanatory claim:

- If x correctly judges that p, then (x correctly judges that p *because* the state of affairs that p obtains).
- If x correctly judges that p, then (x correctly judges that p *because* the proposition that p is true).
- If x correctly conjectures that p, then (x correctly conjectures that p *because* it is probable that p).

<sup>3</sup> Kenny 1963, ch. 9. On the formal objects of emotions, cf. Teroni 2007.

And so on.

What are the relations between correctness and satisfaction conditions?

Firstly, 'correct' ('right', 'richtig'), unlike 'satisfied', is obviously a normative predicate. Correctness (right) and incorrectness (wrong) constitute one of the main families of normative or non-theoretical predicates along with, for example, the family of deontic predicates and the family of value predicates, thin ('valuable') and thick ('cool', 'unjust').

Secondly, the fit of satisfaction is either mind-to-world fit (belief) or world-to-mind fit (desire). Whatever this amounts to it is presumably compatible with the claim that if a state or attitude is satisfied, then it is satisfied because its satisfaction condition holds. If there is a fit of correctness, it is always a mind-to-world fit in the following sense: attitudes, states and acts are correct, if they are correct, *because* the world is the way the correctness conditions say it is.

Thirdly, correctness conditions, unlike satisfaction conditions, refer to formal objects (propositions, states of affairs) or are dominated by formal predicates or functors (truth, obtaining, value, ought, probability).<sup>4</sup>

Fourthly, mental states and acts or their contents, it is claimed, represent (conceptually) their satisfaction conditions. Do mental states and acts represent their correctness conditions? This is a question which needs to be posed and answered for each type of mental act and state which is supposed to have correctness conditions. I shall briefly consider three cases in order to make plausible the view that mental states and attitudes do not represent their correctness conditions.

Consider emotions. Many philosophers have thought that

If x favours y, then x believes that y is valuable,  
If x disfavors y, then x believes that y is disvaluable.

Thus Kenny says:

<sup>4</sup> Searle sometimes refers to states of affairs in giving the satisfaction conditions for belief (e.g. Searle 1983: 14). The simplified theory of satisfaction conditions referred to here is assumed to have no use for formal objects.

It is not, of course, correct to say e.g. that the formal object of envy is another's good *tout court*: one must say that it is something *believed to be good*... (Kenny 1963: 193).

But is it not possible to have a pro-attitude towards an object, to admire a gesture or an ankle, for example, without believing it to be valuable, for example, graceful? May emotions not colour non-conceptual content?

Might a creature not undergo certain emotions and lack any value concepts? Might a creature not have emotions based on simple seeing and lack beliefs entirely? In §4 I shall put forward a view of emotions according to which to undergo an emotion is indeed to stand in an intentional relation to value. But this relation, as we shall see, is not belief nor does it involve any representation (thought) of value.

Consider desire. Is it a condition on desire that whoever desires employs a deontic concept? Considerations very like those adduced against the claim that emotions involve axiological beliefs suggest that desires do not constitutively involve deontic beliefs.

Judgements or beliefs, it has often claimed (by Husserl, Pfänder and Bernard Williams), aim at truth or make a claim to be true.<sup>5</sup> (In the same spirit one might think that desire aims at oughtness and emotions at value). Does this mean that if one judges that *p*, then one judges that the proposition that *p* is true or that it is true that *p*? But the ensuing regress would not be harmless. Suppose that 'It is raining' and 'That it is raining is true' express the very same thought or proposition or are synonymous. Then to judge that it is raining is just to judge that that it is raining is true. But since

That it is raining is true because it is raining

our two sentences cannot express the very same thought or be synonymous.

Are the theory of satisfaction conditions and the theory of correctness conditions rival (albeit partial) accounts of intentionality? It is obvious that the thesis that beliefs have satisfaction conditions and the thesis that they have correctness conditions are not incompatible. And the same is true of desires. Nevertheless it seems that emotions and

<sup>5</sup> See Engel 2005, especially the rejection of the 'intentional interpretation of truth directedness' at §5.

preferences have correctness conditions but no satisfaction conditions. It is true that, if emotions and preferences were definable in terms of beliefs and desires,<sup>6</sup> then it might be possible to show that emotions and preferences do indeed have satisfaction conditions, combinations of the satisfactions conditions of the belief-desire combinations which constitute them. But emotions and preferences are a *sui generis* category of mental states. Beliefs and desires are propositional states but some emotions are not propositional (scorn, admiration, hate); preference is sometimes propositional sometimes it is the preference for one person over another. Desires are future-directed but some emotions are past-directed (regret). It is therefore very difficult to see what satisfaction conditions for emotions might look like. If this is right, then the theory of intentionality in terms of correctness conditions enjoys the advantage of greater generality over the theory of intentionality in terms of satisfaction conditions.

There is one final striking difference between correctness and satisfaction conditions. The former but not the latter are widely held to be problematic. Correctness conditions refer to entities the existence of which has been roundly rejected by naturalists and by nominalists — propositions and states of affairs. Correctness conditions employ predicates to ascribe properties which have often been considered suspect — value, oughtness. Indeed much twentieth century philosophy has been marked by scepticism about formal objects and properties. Thus philosophers have argued not only that there are no propositions or facts (obtaining states of affairs) but also that ‘It is raining’ and ‘The proposition that it is raining is true’ say the very same thing. Similarly, it has been argued that value-ascriptions and norm-ascriptions have no truth-values and that probability can be dispensed with in favour of frequency.

There is also an objection to one particular type of correctness condition, the conditions for judgement (belief, conviction, certainty). Mention of propositions or states of affairs in the correctness conditions for judgements, the objection goes, is superfluous. The only correctness condition we need is:

$$x \text{ judges that } p \rightarrow (x \text{ judges correctly that } p \text{ iff } p)$$

<sup>6</sup> Cf. Searle 1983: 31-36.

In other words, the satisfaction condition for judgement is just its correctness condition.

How should a friend of correctness conditions react to the many different objections I lumped together under the claim that correctness conditions are problematic? To the objections that there are no propositions or states of affairs, no values and no norms? To the objection that some or all correctness conditions have no truth values? To the claim that correctness conditions for judgment and belief can be given without mentioning states of affairs or propositions?

A philosopher who intends to provide a philosophy of intentionality and thinks that an account of the intentionality of attitudes, acts and states which can miss their mark can be given in terms of correctness conditions must in any case provide a complementary account of the intentionality of knowledge. Suppose that a plausible account of the intentionality of knowledge could be shown to entail that there are facts, values, norms, probabilities etc. Were that the case our philosopher would be able to kill two birds with one stone. He would have an account of the two main types of intentionality and his account of the intentionality of knowledge would give him the very best of reasons for holding that correctness conditions are unproblematic.

What would such an account of the intentionality of knowledge look like? Is such an account plausible? In §3 I argue that knowledge that *p* involves a relation to facts. In §4 I identify the most plausible version of the view that there is knowledge of values.

### 3. Knowledge & facts

Is knowledge knowledge of facts? Is *knowledge that p* knowledge of facts, understood in some suitably non-anaemic way? Russell, Vendler (1967), Angelika Kratzer (2000) and Keith Hossack (2007) give affirmative answers to this question. The perhaps more popular, negative, answer is given by Ramsey (1931) and Timothy Williamson (2000). The conception of facts which is shared by friends and enemies of the view that knowledge that *p* is knowledge of facts is not the view that facts are true truth-bearers, for example, propositions. The shared conception is one of two more robust accounts of facts. According to the first robust account a fact is just an obtaining state of affairs. According to the second robust account, a fact is just a *sui generis* type of entity in which objects exemplify properties or stand in

relations. Each of the two robust accounts claims that facts contain objects and properties whereas propositions contain only concepts. In what follows, I shall assume that the first robust account of facts is the right account.<sup>7</sup>

As far as I can tell, reflection on the concept of knowledge that *p* has not come up with any decisive argument in favour of the view that knowledge that *p* is knowledge of facts. There is nevertheless the possibility that types of knowledge other than knowledge that *p* constitutively involve an intentional relation to facts and the possibility that reflection on the relations between knowledge that *p* and other types of knowledge might lead us to the conclusion that knowledge that *p* does after all involve an intentional relation to facts.

Is knowledge that *p* the only type of knowledge? No. We can distinguish at least four distinct kinds of knowledge. Knowledge is propositional or non-propositional, episodic or non-episodic. In distinguishing between propositional and non-propositional knowledge I have in mind only the distinction between what makes true knowledge ascriptions of the form ‘*x* knows that *p*’ and what makes true ascriptions of the form ‘*x* knows *y*’. Knowledge that *p* is propositional and non-episodic; it is either a relational state<sup>8</sup> or a disposition or constitutively involves an intentional relation. Knowledge that *p* is the type of knowledge which dominates contemporary epistemology. But there are at least three other types of knowledge.

There is coming to know that *p* or apprehending that *p* (*erkennen, dass p*), which is propositional and episodic. There is no established or happy English translation of ‘*erkennen*’ unless, like some anglophone epistemologists long ago, we talk of an ‘act of knowing that’. There is acquaintance, which is non-propositional and non-episodic (‘I have known her for years’). And there is coming to be acquainted with (*kennenlernen, Kenntnisnahme*) someone or something, which is non-propositional and episodic. This is what might be called making the non-social acquaintance of something or someone.

<sup>7</sup> On these three accounts of facts, see Correia & Mulligan 2007.

<sup>8</sup> For the view that such a relational state is a trope cf. Mulligan & Smith 1986, Smith 1984. For the view that it is a non-trope, cf. Williamson 2000. I now believe that knowledge that *p* is what I call a functorial and so formal state and no relation. But this complication plays no role in what follows.

How do the four types of knowledge hang together? One very plausible view is that apprehension typically marks the beginning of the state or disposition which is knowledge that *p*. And that coming to be acquainted with someone or something typically marks the beginning of the state or disposition of being acquainted with that person or thing. On this view, epistemic episodes are more fundamental than epistemic states or dispositions. This view is supported by the observation that we can always ask with respect to any claim to know that *p* or any claim to know someone ‘*How* do you know that *p*?’, ‘*How* do you know her?’ Questions of this type<sup>9</sup> make little sense with respect to belief claims:

\*How does she believe that *p*?

Answers to the question ‘How (cf. ‘*unde*’, ‘*woher*’) does *x* know that *p*?’ specify the putative *source* of *x*’s knowledge that *p*. And this source is what is apprehended in the episode of apprehending that *p*. Answers to the ‘How does *x* know *y*?’ question specify the putative source of *x*’s knowledge of *y*. And this source is what is apprehended in the episode of coming to be acquainted with *y*.

Let us look first at the most basic case, making the acquaintance of something or someone, episodic knowledge by acquaintance. Is seeing (hearing, touching) someone or something enough to constitute episodic acquaintance? Answers to this question will depend on the account of seeing employed. Suppose, with Dretske, that

*x* simply sees *y* iff *y* is visually differentiated for *x*

and that if *x* simply sees *y* and *y* = *z*, then *x* simply sees *z*. Does simple seeing so conceived suffice for making the acquaintance of someone or something? Suppose a young child glimpses out of the corner of her eye a fat man, who is in fact the President. She has then seen the President. Has she become acquainted (in a non-social way) with the President? Does she enjoy epistemic contact with the President? Most of us, I suspect, would give a negative answer to this question. What further condition, then, must be satisfied by simple seeing if it is to count as coming to be acquainted with?

The relevant condition, I suggest, is identification:

<sup>9</sup> Cf. Wittgenstein 1969 §550; Austin 1961: 46; Reiner 1934: 39.

If  $x$  comes to be visually acquainted with  $y$ , then  $x$  sees  $y$  at  $t_1$  and then at  $t_2$  and sees  $y$  at  $t_2$  as the same object.

Seeing things and people as the same, identification, is a phenomenon investigated in psychology under the name of 'object constancy'. Object constancy typically occurs in visual (tactile, auditive) perception along with different types of property constancy, colour constancy, shape constancy etc.

Identification is a mental act which has correctness conditions:  $x$  correctly identifies  $y$  and  $z$  only if  $y = z$ . Such identification may but need not take the form of a judgement. Simple seeing has no correctness conditions; it is an intentional relation we stand in to things and processes. Coming to be visually acquainted with something has no correctness conditions either. But it involves identification, which does have correctness conditions. What we simply see are substances, states, processes and events. Episodic visual acquaintance is acquaintance with objects and properties. It is based on a relation to substances, states, processes and events.

This account of episodic acquaintance provides in all essentials the model for the account to be given of apprehending that  $p$ . Simple seeing of things and persons stands to episodic visual acquaintance as seeing that stands to episodic visual apprehension. Suppose Sam sees that Maria is sad. 'See that' is factive, just as 'see' is veridical. Is seeing that  $p$  coming to know that  $p$ ? Just as the fact that Sam sees Maria does not make it true that he thereby makes her acquaintance, so seeing that Maria is sad is not knowledge that she is sad. What do we need to add to seeing that  $p$  to obtain visual apprehension that  $p$ ? The missing ingredient, as before, is identification:

Sam sees that Maria is sad at  $t_1$  and then at  $t_2$  and identifies what he sees at  $t_1$  and at  $t_2$

But how should this be understood? One bad answer is

\*Sam identifies what he sees at  $t_1$ , that Maria is sad, and what he sees at  $t_2$ , that Maria is sad

This is a bad answer because

\*That Maria is sad = that Maria is sad,

like all instances of

\*That  $p =$  that  $q$ ,

is unacceptable. All instances of

The fact that  $p =$  the fact that  $q$ ,

on the other hand, are well-formed. This suggests that identification should be understood as follows:

Sam identifies the fact that Maria is sad, which he perceives at  $t_1$ , and the fact that Maria is sad, which he perceives at  $t_2$

Identification is a mental act (a 'synthetic' mental act), unlike the identity predicate or concept. Some identifications are identity judgements. But this is not always the case. When it is not the case, identification typically provides us with good reason to form identity judgements. Similarly, one can identify facts without identifying them *as* facts. Whether it is a judgement or not identification has, as already noted, correctness conditions.

Suppose Sam is asked whether Maria is sad. Motivated by the desire to reply to the question and the desire to know whether she is sad, he observes her. As before, the identification theory will not claim that

\*Sam identifies that Maria is sad, what Sam sees, and that Maria is sad, what Sam refers to.

But rather that

Sam identifies the fact that Maria is sad, which he perceives, and the fact that Maria is sad, which he refers to.

The theory of apprehending that  $p$  as identification might be called the theory of 'fact constancy' by analogy with the theory of object constancy in the area of visual acquaintance with things and persons. The identification theory of apprehension has implications which not all philosophers will find equally acceptable. For example, that to apprehend that  $p$  by inferring validly from known premises to  $p$  involves going through the inference at least twice. And, another example, that coming to know that  $p$  through testimony requires a double-take.

Visual apprehension, like episodic visual acquaintance, is based on a relation. Episodic visual acquaintance is based on the relation of simple seeing. Visual apprehension is based on perceptions of facts not merely on seeing that. Episodic visual acquaintance and visual apprehension are types of knowledge. Are they relations? No. They are complex mental acts and are partly relational and partly non-relational. The identifications which are constitutive of each type of knowledge can go wrong. They are therefore not relations. But each type of knowledge involves a relation. Thus even those cases of visual acquaintance and visual apprehension which involve judgements differ fundamentally from the case where one judges truly and for good reason on the basis of visual information that *p*. For such judgements are not relations and do not involve relations.

What is the relation between visual apprehension that *p* and knowledge that *p*? The answer will depend in part on what we say about the relation between belief and knowledge that *p*. We have seen that to apprehend that *p* is not to judge truly that *p* and for the judgement to be justified. A number of impressive arguments have recently been marshalled against the parallel claim that knowledge that *p* is true, justified belief (Williamson 2000). An argument to the same conclusion in the spirit of the foregoing runs as follows. If knowledge that *p* were a type of belief, however qualified, it would be possible to ask, with respect to any knowledge claim: 'Why do you know that *p*'? But, as we have seen, this is not possible. This argument is not, however, conclusive since it might be argued that it is the qualifications of belief which are supposed to constitute knowledge which make it inappropriate to ask the 'why' question.

Two better objections to the view that knowledge is a type of belief have to do with the nature of belief. Firstly, belief (like conviction) but not knowledge comes in degrees. Secondly, belief (like conviction) is either positive or negative. I shall concentrate in what follows on the second objection.

Negative belief that *p* is not the same thing as positive belief that not-*p*. Negative belief is disbelief. Belief and disbelief are polar opposites. But knowledge that *p* does not have this property. Knowledge has a contradictory opposite, ignorance, and contrary opposites, error

and illusion. But, unlike belief, knowledge does not come in two kinds, positive and negative.<sup>10</sup>

One argument against the view that coming to know that *p* is to judge truly and for good reason that *p* relies on a putative similarity between judgement and belief. There is a view of judging according to which it comes in two polarly opposed kinds: positive judging or acceptance and denial or rejection<sup>11</sup> just as belief comes in two polarly opposed kinds. If this view of judgment is right, then coming to know that *p* is never a type of judging. But if Bolzano, Husserl and (perhaps) Frege are right, judging does not come in two polarly kinds: to reject *p* is just to judge that not-*p*. I shall assume here that they are right.

What, then, is the relation between knowledge that *p*, apprehension that *p*, judgment and belief? Consider

If *x* apprehends that *p*, then *x* believes that *p*.

This is clearly false. Belief is a *reaction* to, for example, what is (apparently) apprehended when we (apparently) apprehend that *p*.<sup>12</sup> That it is a reaction follows from the fact that it comes in two varieties, positive belief and negative belief (disbelief). Reactions and responses (which manifest intentionality) to phenomena of different sorts may be more

<sup>10</sup> The claim that belief comes in two kinds, positive and negative, is compatible with different views about the ontology of belief. Suppose belief is a disposition and, unlike its categorical basis, not any kind of state. Then to believe that *p* is to be disposed to accept *p* and to reject not-*p*. Suppose that belief is a state but no disposition. Then it is either a state of positive belief or a state of negative belief. Presumably, anyone in such a state, for example, the state of disbelief that *p*, will be disposed to accept not-*p* and to reject *p* because of his doxastic state. A third view has it that belief is both a disposition and a state. Finally, there is the view that some beliefs are dispositions but not states and some are states but not dispositions. Perhaps the most plausible candidates for belief states which are not dispositions are beliefs which are (a) important for us, in particular evaluative beliefs, and (b) are continually being tried, tested and reinforced. Sam, for example, is one of those people who continually refer to the European Union as ‘the Belgian Empire.’ His attitude towards what he takes to be pervasive EU rhetoric is one of strong disbelief. One objection to the view that Sam’s disbelief could be a state is that psychological phenomena are episodic. But perhaps Sam’s disbelief is *one* state albeit a scattered state. After all, some substances are spatially scattered.

<sup>11</sup> Cf. Rumfitt 2000.

<sup>12</sup> To believe, Augustine says, is ‘cum assensione cogitare’ (*de praed. sanct.* 2 5).

or less well entrenched, likely or typical. But it is always a contingent matter whether or not a reaction of a certain kind to phenomena of certain other kinds occurs. Often, but not always, we respond doxastically to seeing something or seeing that *p*. Often, but not always, we respond doxastically to (apparently) apprehending that *p*.

*That* we do so react is essential to the state of knowing that *p*. That is why

If *x* knows that *p*, then *x* believes that *p*.

What should we say about

If *x* apprehends that *p*, then *x* judges?

This, I argued above, is false since identification of a fact one perceives in one way with a fact perceived in another way need not involve judgement. Nevertheless identification may and often does take the form of a judgement. Apprehension that *p* which involves an identity judgement consists of perceptions of a fact and an identity judgement which is correct. Since perception of facts, like simple seeing and simple seeing that, is veridical, apprehension that *p* is based on a relation to a fact.

If the foregoing is correct, then knowledge that *p* has its source in apprehension. Knowledge that *p* is not knowledge of facts. Apprehension is not apprehension of facts. But perception of facts is constitutive of apprehension. And apprehension that *p* brings knowledge that *p* into being. But that this is so will only be apparent to us if we trace knowledge that *p* back to its roots, if we ask *how* we know that *p*. Thus the theory of apprehending that *p* in terms of identification seems to give the friend of correctness conditions three things he needs: the beginnings of an account of the intentionality of knowledge; reason to think that the reference to obtaining states of affairs or facts in the correctness conditions for belief is neither superfluous nor problematic; an account of the way the intentionality of belief, judgement and knowledge that *p* hang together.

#### 4. Knowledge & values

Perhaps the most problematic family of correctness conditions in §2 is the group of correctness conditions for emotions, wishes, preferences

and desires. For these conditions are dominated by axiological and deontic predicates and functors. And sentences dominated by such predicates and functors are widely held to have no truth-values. Even philosophers who are prepared to allow that such sentences have truth values often reject one claim made by the friend of correctness conditions. Thus consider again

- (1)  $x$  regrets that  $p \rightarrow$  ( $x$  correctly regrets that  $p$  iff it is regrettable that  $p$ );
- (2) If  $x$  correctly regrets that  $p$ , then ( $x$  correctly regrets that  $p$  *because* it is regrettable that  $p$ ).

A friend of correctness conditions who thinks that these provide a partial account of the intentionality of one type of state or attitude, regret, endorses both (1) and (2). But (1) might be combined with the denial of (2) and endorsement of

- (3) If it is regrettable that  $p$ , then (it is regrettable that  $p$  *because* ( $x$  regrets that  $p \rightarrow x$  correctly regrets that  $p$ )).

(1) and (3) yield one version of what are sometimes called ‘neo-sentimentalist’ or ‘buck-passing’ accounts of what it is to be valuable. According to this version, if something is valuable this is because a pro-attitude towards it would be correct; if something is sublime, this is because a mixture of awe and something like fear would be correct; if it is shameful that  $p$ , this is because being ashamed that  $p$  would be correct. It is not the most popular version of such theories. The more popular versions do not typically appeal to the correctness of emotions but rather to appropriate emotions, to justified emotions, to permissible emotions or to emotions we have undefeated reasons to feel.<sup>13</sup>

A friend of (1) and (2) (like a friend of (1) and (3)) owes us an account of the intentionality of knowledge of values and norms. Is there any plausible such account which will enable him to claim that the correctness conditions for emotions, desires, wishes and preferences have truth-values? And to claim against the neo-sentimentalist that (2) is to be preferred to (3)? And to specify the nature of the relation between knowledge of values, on the one hand, and desires, emotions, wishes and preferences, on the other hand?

<sup>13</sup> Cf. Rabinowicz & Rønnow-Rasmussen 2004, Skorupski this volume, Scanlon 1998, Mulligan 1998.

As far as I can see, the main theories of what it is to have knowledge of values are the following. There is a type of intuitionism which claims that we have intuitive knowledge of values but which says nothing about the nature of this type of intuition. Moore sometimes endorses such a position. Then there is the view which combines the following three claims: (a) we are acquainted with values and know that certain objects are valuable; (b) such acquaintance and knowledge are merely special cases of perceptual acquaintance or intellectual knowledge differing from the more familiar cases only in having unusual objects; (c) acquaintance and knowledge of these kinds are the *only* form of epistemic contact we have with values.<sup>14</sup> A third type of theory has it that we have affective knowledge of values. A fourth that it is desires rather than affective phenomena which may constitute knowledge of values. Clearly, a philosopher who thinks that we have affective or ‘conative’ knowledge of values may also hold that we have intellectual axiological knowledge that.

One version of the view that there is affective knowledge of values is the view that affects may ‘disclose’ values. Such a view has been put forward by Mark Johnston (2001). Another view has it that emotions may disclose values. Such a view has been defended by Christine Tappolet (2000). The view that desires may constitute knowledge of values has been defended by Graham Oddie (2005).

I shall first formulate what I take to be the general form of these or related claims. I shall then formulate some objections to such claims and put forward an alternative. Consider first the view that emotions or affects — favouring — can amount to knowledge of values:

- (4)  $x$  is affectively acquainted with the value of  $y$  iff  $\text{Val}(y)$  &  $x$  appropriately favours  $y$ ;
- (5)  $x$  affectively knows that it is valuable that  $p$  iff  $\text{Val}(p)$  &  $x$  appropriately favours that  $p$ .

Similarly, we may formulate the view that desires can give us knowledge of values as follows:

- (6)  $x$  ‘desideratively’ knows that it is valuable that  $p$  iff  $\text{Val}(p)$  &  $x$  appropriately desires that  $p$ .

<sup>14</sup> Cf. Thomas 2006: 215, 51.

The first thing to notice about (4)-(6) is their similarity to

(7)  $x$  knows that  $p$  iff  $p$  &  $x$  justifiably believes that  $p$ .

Thus appropriate favouring or desire plays the same role in (4)-(6) as justifiable belief in (7). And the first clause on the right hand side of (7) plays the same role as the first clause on the right hand side of (4), (5) and (6). Perhaps, then, there are objections to (4)-(6) which resemble the objections to (7) above.

One striking feature of affects and emotions is that they often have positive or negative valence. Another is that if an emotion or an affect has a valence, there is often an emotion or affect which is its polar opposite. Being pleased (respect, liking, happiness) has positive valence and a polar opposite, being displeased (scorn, disliking, unhappiness). But surprise has no valence and so can have no polar opposite and enjoyment has a positive valence but no polar opposite. Now knowledge in all its manifestations has no valence and no polar opposite. That is one reason for thinking that emotions and affects cannot constitute knowledge.

There is a second reason for rejecting the view that emotions and affects can yield knowledge. Emotions and affects are *reactions*, affective reactions. Indeed having a valence suffices to make a state a reaction. States of opposed valence or 'sign' constitute opposed reactions. That is why so many psychological theories consider an action-tendency to be essential to many types of emotions and affects. But knowledge is no reaction. So emotions and affects can never yield knowledge.

Very similar reasons can be advanced against the view that desires may constitute knowledge. Desires (like wishing, wanting, willing and striving) come in two kinds, positive and negative. There is the positive desire that  $p$  and the striving to realise  $p$  but there is also negative desire, negative willing, aversion, shunning (and *Widerstreben*). Knowledge is not like that.

I suggested in §3 that belief is a reaction,<sup>15</sup> that reactions which are intentional states are reactions to something, and that belief is a reaction to, for example, what we apprehend or apparently apprehend. If emotions, desires and affects are reactions and states or attitudes which

<sup>15</sup> Are beliefs reactions in the same sense in which emotions are reactions? Cf.: 'The concepts of believing, expecting, hoping are less distantly related (*artfremd*) to one another than they are to the concept of thinking' (Wittgenstein 1968 §574).

enjoy intentionality, then we might expect that they, too, are reactions to what is known or apprehended or apparently known or apprehended. But what kind of knowledge could be such that emotions, desires and affects are reactions to its (apparent) deliverances?

An ideal candidate for the role of affective knowledge of value, it is now clear, should satisfy five desiderata. It should not be any sort of reaction and should have no valence and so should not be any sort of emotion, affect or desire. But it should be an affective state or episode. Finally, it should make true a psychological ascription which is veridical or factive. Is there any such thing?

Suppose that Maria is walking down the street and observes a scene in which bread is being distributed unequally to equally needy children. She is struck, as we say, by the injustice of the situation. She has felt the injustice of the situation. Perhaps she reacts with indignation. Perhaps she is suffering from indignation fatigue and feels no emotion whatsoever. We are often struck by the elegance of a gesture or the grace of someone's gait, by the rudeness of a remark, by the beauty of a building. Typically experience of value prompts affective reactions, admiration, annoyance, pleasure. That is one reason why it is a mistake to think that experience of value, feeling value, is an affective reaction. Another reason is that although feeling value is an affective phenomenon it has no valence, it is neither positive nor negative, and hence no polar opposite.

Experience of value seems to be very common. Of course, if axiological nihilism is correct, if nothing is a value and nothing has value, then there is no experience of value, only experience *as of* value. But if ordinary language and experience are taken at face value we are all the time experiencing (dis)value and comparative value. Non-cognitivism — axiological and deontic sentences have no truth-values — is difficult to reconcile with many entrenched assumptions. But the claim that what seem to be affective experiences of value are always merely experiences as of value is even more difficult to swallow. The theory of value has suffered from over-concentration on the arguments for and against cognitivism and on the arguments for and against unnatural values and value-properties. The different semantic and metaphysical options can come to look very different if we bear in mind experience of and the experience as of value. Consider, for example, the possibility that only affective value-experience can explain how value-predicates can have any meaning. Neglect of value-experience also has consequences for substantive ethical and political

questions. Consider the justification of tolerance. It is one thing to appeal to the fact that different people hold different and often incompatible axiological beliefs. It is quite another to appeal in addition to the fact that we are all in different ways value-blind or, more exactly, insensitive to different types of value; and to the fact that sensitivity to one type of value often makes one insensitive to other types of value.

'Feel' in the sentence 'Maria felt the injustice of the situation' is veridical. If Maria felt the injustice of the situation, then the situation was unjust. If she is struck by the beauty of the building, it is beautiful. Maria's indignation is a reaction either to a felt disvalue, the injustice of the situation or to a merely apparently felt value. In the latter case she is the victim of an illusion. Her admiration of the elegance of Giorgio's gait is a reaction to a felt, positive value or it is a reaction to an apparently felt value.<sup>16</sup> Above I objected to the claim that

If x favours y, then x believes that y is valuable.

We now have a more plausible alternative:

If x favours y, then x feels the value of y or x merely seems to feel the value of y or x believes y to be valuable.

Is feeling value an exception to the principle that all knowledge involves identification? No. Values are felt more or less clearly, more or less fully and transitions along these two dimensions involve identification. Aesthetic experience is perhaps the clearest example of the phenomenon of continuously feeling the same value as the same under different modes of presentation. Just as we distinguished between simple seeing and episodic visual acquaintance, so too, we should distinguish between (a) feeling value which is no form of knowledge but rather the analogue of simple seeing and perception and (b) the case where feeling value does constitute knowledge because it involves identification.

The claim that affective knowledge consists at bottom of feeling values and disvalues is, I have argued, superior to theories according to which emotions, affects or desires can yield knowledge. If axio-

<sup>16</sup> Some emotions are factives. If Sam regrets that p, then p. But it is not true that if Sam regrets that p, then it is regrettable that p.

logical nihilism is false, this claim is, I suggest, the best available approach in the epistemology of values. For it is not only preferable to other theories of affective knowledge and to the idea that desires yield knowledge, it is also preferable to any epistemology the neo-sentimentalist can come up with.

Suppose with the neo-sentimentalist that being valuable is understood in terms of appropriate emotions or good, undefeated reasons to feel emotions. What, then, would knowledge of the value of an object amount to? The neo-sentimentalist account of what it is to be valuable does not allow one to infer what a neo-sentimentalist account of the knowledge of value would look like. And I have not come across any developed neo-sentimentalist account of the knowledge of value (later than that given by Brentano). But presumably a neo-sentimentalist must hold that we sometimes have knowledge of the appropriateness of an emotion. And in the most basic cases this could only be knowledge *that* an emotion is appropriate, *that* there are undefeated reasons to feel an emotion. But if we have knowledge of values it is extremely implausible to think that such knowledge consists *only* of knowledge that, a knowledge by description which has no anchorage in any knowledge by acquaintance. On one common and plausible view, knowledge that *p* cannot motivate; even axiological knowledge, knowledge that it is valuable that *p*, is an intellectual state and, like all such states, cannot motivate. Feeling (dis)values, however, is no intellectual state and can motivate.<sup>17</sup> Finally, it seems that neo-sentimentalism cannot do justice to the fact that the very best reason one could have for feeling, for scorn, admiration etc. is knowledge of the value of the material or proper object of the emotion.<sup>18</sup>

## 6. Knowledge vs reactions to what is (apparently) known

How, then, do the intentionality of knowledge and the intentionality of states and attitudes which can go wrong hang together?

Our answer runs as follows. We have distinguished three types of intentional states and acts: (1) acts and states such beliefs, judgments

<sup>17</sup> Closely related to the distinction between emotions and feeling value is the distinction in neuro- and affective science between emotions and motivational saliency; cf. Berridge & Robinson 2003.

<sup>18</sup> Cf. Mulligan 2008.

and emotions which have correctness conditions; (2) acts which are genuinely relational — seeing things, seeing that *p*, perception of facts — but which are not types of knowledge; (3) five types of knowledge — episodic acquaintance, enduring acquaintance, apprehension that *p*, knowledge that *p* and feeling value.

Knowledge ascriptions, like ascriptions of relational acts, are factive. Intentional relational acts are essential parts of knowledge. Beliefs and emotions are reactions. Part of what it means to say that belief and emotions are intentional states and attitudes is given by specifying their correctness conditions. Another part of what it means to say that they are intentional is given by an account of their material or proper objects, an account about which I have said nothing. Correctness conditions for judgements and beliefs mention states of affairs and predicate the formal property of obtaining of states of affairs. Correctness conditions for emotions mention the material objects of these emotions and predicate formal value properties of these.

The intentionality of beliefs and emotions is triply dependent on the intentionality of knowledge or on the intentionality of relational acts and states. First, beliefs are reactions to what is known or to what seems to be known or to what is the case or to what seems to be the case. Sam judges or believes that Maria is sad on the basis of his acquaintance with the fact that she is sad or on the basis of what seems to be acquaintance with this fact or on the basis of seeing that she is sad or on the basis of seeming to see that that she is sad. Emotions, on the other hand, are reactions to (apparent) non-intellectual, affective knowledge or (apparent) intellectual knowledge of the exemplification of value. Secondly, the correctness conditions of judgements, beliefs and emotions tell us what would be known if these states and acts were reactions to knowledge of the right kind. Finally, if we had no relational contact with facts and values, we would have no right to mention states of affairs or predicate value in correctness conditions or, indeed, anywhere else.

One consequence of the theory of intentionality sketched here is that although the intentionality of many states is explained in part in normative terms ('correctness') and part of the intentionality of knowledge is explained in normative terms (the correctness conditions for identifications), there is nothing normative about the most

basic types of intentional relations: simple seeing of things and persons, perception of facts and feeling value.<sup>19</sup>

Kevin Mulligan  
 Département de Philosophie, Faculté des Lettres  
 Université de Genève  
 2, rue de la Candolle  
 CH 1211 Genève, Switzerland  
 kevin.mulligan@lettres.unige.ch

### References

- Austin, J. L. 1961. Other Minds. In *Philosophical Papers*. Oxford: Clarendon Press, 44-84.
- Berridge, K. C. & Robinson, T. E. 2003. Parsing Reward. *Trends in Neuroscience* 26, 9: 507-513.
- Correia, F. & Mulligan, K. 2007. Facts. *Stanford Encyclopedia of Philosophy*, <<http://plato.stanford.edu/entries/facts/>>.
- Engel, P. 2005. Truth and the Aim of Belief. *Laws and Models in Science*, ed. D. Gillies, 77-97.
- Hossack, K. 2007. *The Metaphysics of Knowledge*. Oxford: Oxford University Press.
- Johnston, M. 2001. The Authority of Affect. *Philosophy and Phenomenological Research* 63(1): 181-214.
- Kenny, A. 1963. *Action, Emotion, and Will*, London: Routledge.
- Kratzer, A. 2001. Facts: Particulars or Information Units. *Linguistics and Philosophy* 25, 5-6: 655-670.

<sup>19</sup> *History*. The view that the intentionality of many types of states and attitudes can be partially specified in terms of correctness conditions goes back to Husserl. The view that knowledge arises through identification is also Husserlian. The idea that we can feel positive and negative values goes back to Reinach and indeed to Hutcheson and to Kant in his celtic (sometimes misleadingly called his pre-critical) period. This idea together with the claim that emotions are reactions is defended by Scheler and von Hildebrand. The view that belief (but not judgement) comes in two polarly opposed kinds is defended by Reinach. The view that propositional judgement and belief comes in two polarly opposed kinds is defended by Meinong. For details, see Mulligan 2004, 2004a, 2006. Thanks to Wlodek Rabinowicz, who did not know what he was reacting to, to Toni Rønnow-Rasmussen, who did, and to Philipp Keller who saw through what he was reacting to; to Graham Oddie, Barry Smith, Fabrice Teroni, John Skorupski, Christian Beyer, Fabian Dorsch, Olivier Massin and Teresa Marques and also to Philip Gerrans for putting me on to Berridge & Robinson 2003. My paper was written under the generous auspices of the Swiss NCCR in affective sciences and the FNS project on properties and relations.

- Mulligan, K. & Smith, B. 1986. A Relational Theory of the Act. *Topoi* 5/2: 115-130.
- Mulligan, K. 1998. From Appropriate Emotions to Values. *Secondary Qualities Generalized*, ed. P. Menzies. *The Monist* 84, 1: 161-188.
- Mulligan, K. 2004. Essence and Modality. The Quintessence of Husserl's Theory. In M. Siebel & M. Textor, eds., *Semantik und Ontologie. Beiträge zur philosophischen Forschung*. Frankfurt: Ontos Verlag, 387-418.
- Mulligan, K. 2004a. Husserl on the 'Logics' of Valuing, Values and Norms. *Fenomenologia della Ragion Pratica. L'Etica di Edmund Husserl*, eds. B. Centi & G. Gigliotti, 177-225, Naples: Bibliopolis.
- Mulligan, K. 2006. Facts, Formal Objects and Ontology. *Modes of Existence. Papers in Ontology and Philosophical Logic*, eds. Andrea Bottani & Richard Davies. Frankfurt: ontos verlag, 31-46.
- Mulligan, K. 2008. On Being Struck by Value. *Wohin mit den Gefühlen? Emotionen im Kontext*, ed. Barbara Merkel, mentis-Verlag, forthcoming.
- Oddie, G. 2005. *Value, Desire and Reality*. Oxford: Oxford University Press.
- Rabinowicz, W. & Rønnow-Rasmussen, T. 2004. The Strike of the Demon: On Fitting Pro-Attitudes and Value. *Ethics* 114: 391-423.
- Ramsey, F. 1931. Facts and Propositions. In *The Foundations of Mathematics*. London: Kegan Paul.
- Reiner, H. 1934. *Das Phänomen des Glaubens*. Halle: Niemeyer.
- Rumfitt, I. 2000. Yes and No. *Mind* 109: 781-823.
- Scanlon, T. M. 1998. *What we Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Searle, J. 1983. *Intentionality. An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Smith, B. 1984. Acta cum fundamentis in re. *Dialectica* 38: 157-178.
- Tappolet, C. 2000. *Emotions et Valeurs*. Paris: Presses Universitaires de France.
- Teroni, F. 2007. Emotions and Formal Objects. *Dialectica* Vol. 61, Fasc. 3: 395-416.
- Thomas, A. 2006. *Value and Context. The Nature of Moral and Political Knowledge*. Oxford: Clarendon Press.
- Vendler, Z. 1967. *Linguistics in Philosophy*. Ithaca: Cornell University Press.
- Williamson, T. 2000. *Knowledge and its Limits*, Oxford: Oxford University Press.
- Wittgenstein, L. 1968. *Philosophical Investigations*. Oxford: Blackwell.
- Wittgenstein, L. 1969. *On Certainty*. New York: Harper.

# Acting without reasons

Josep L. Prades  
University of Girona

## Abstract

In this paper, I want to challenge some common assumptions in contemporary theories of practical rationality and intentional action. If I am right, the fact that our intentions can be rationalised is widely misunderstood. Normally, it is taken for granted that the role of rationalisations is to show the reasons that the agent had to make up her mind. I will argue against this. I do not object to the idea that acting intentionally is, at least normally, acting for reasons, but I will propose a teleological reading of the expression 'for reasons.' On this reading, it is quite possible to act for reasons without having reasons to act. In a similar way, paradigmatic cases of cogent practical reasoning do not require the transference of justification from the premises to the practical conclusion.

Let us consider a putative paradigmatic case of cogent practical reasoning. I am in Barcelona with some friends and I start thinking about the best way of going back to my home in Girona before dinner. I have been drinking too much to drive. So, I decide to leave the car in Barcelona and, after checking the time table, I realise that taking the train is an adequate way of getting home. Consequently, I form the intention of taking the train. Here, we have a process of thought in which I have reasoned in order to reach the practical conclusion. This conclusion does not crucially depend, let us assume, on false or unjustified beliefs or on invalid inferential transitions. And, by knowing the relevant thoughts that rationalise my intention to take a train, it seems that you can know the reason for which I act: getting home by dinner time.

When trying to describe what is going on in those kinds of cases, there is a well-known difficulty. On the one hand, it is commonly assumed that the previous process of reasoning determines a reason *for which* I take a train — being in time to have dinner in Girona with my wife, let us say. On the other, it would seem that, for everything that has been said, it does not follow that *I have* a reason to take the train. For I could engage in the previous, cogent pattern of practical

reasoning even if I knew that I have no reason to go to Girona for dinner time, and therefore even if I knew that I could not have any reason to take the train that goes to Girona. For instance, my having dinner with my wife is not a reason to go to Girona tonight when my wife does not expect me and the life of some friend depends on my staying in Barcelona. It seems then that two apparently incompatible descriptions of the phenomenon are required: someone can take the train for a reason without having any reason to take the train.

My solution to the seeming paradox crucially differs from most contemporary accounts. If I am right, the relevant sense of ‘acting for a reason’ in which, in the previous example, I act for a reason, fixes a very peculiar relation between the agent’s intentional action and the existence of this kind of reason. It is not, as it is commonly assumed, that I could have reasons of a special kind M (‘motivating’ or ‘explanatory’, let us say), while lacking reasons of some other kind N (‘normative’ or ‘justifying’, let us say). It is, on the contrary, that my *acting for* reasons when I take the train to go to Girona to have dinner there with my wife does not require the *independent* relation of having (or believing I have) reasons *of any kind*. Those reasons for which I act are just the intention with which I act. To act for them does not require the conjunction of two independent facts: having a reason and being motivated by it. The fact that I *act for* reasons does not require the independent fact that I *have* reasons to act. In the end, my account points towards a teleological interpretation of the process of practical reasoning. If ‘practical reasoning’ denotes a kind of process, it is just a process of adopting specific goals.

In the first section of this paper, I will try to describe the inconsistency of standard Humean accounts when trying to solve the apparent paradox I have just mentioned. In the second one, I will generalise the previous conclusions to show that the same inconsistency also affects most habitual non-Humean accounts. They share a common illusion that is described in the third section, where I will justify my own proposal. Finally, in the last section, I will make some very general remarks about the structure of practical reasoning that are entailed by my account.

## I

Imagine someone — let us call him Roger Bad — who decides to kill his neighbour in order to steal an old copy of Hume's *Treatise*. Being a clever philosopher, Roger Bad makes a careful plan, and considers some different alternatives. In the end, he makes up his mind: poison is the best way of fulfilling his goal. So, the next day he goes to the shop in which, as he knows, they sell poison because he believes that he can buy there the stuff with which he can kill his neighbour. Many philosophers think that our Roger Bad has a reason to enter the shop. Just compare Roger Bad with Roger Irrational. Roger Irrational decides also to buy poison *in order to* kill his neighbour, *in order to* get a copy of Hume's *Treatise*. But there is a problem with him: he is unable to overcome his unjustified fear of being discovered. So he decides, against all evidence, that poison is not an effective way of killing his neighbour and chooses to start a sequence of secret magical rituals with the intention of producing his neighbour's death. There is a difference between Roger Bad and Roger Irrational. Roger Bad satisfies certain normative requirements that Roger Irrational does not satisfy. For Roger Irrational chooses, against all evidence, instrumentally inadequate means for his goal. One of the main conclusions of this paper is that we cannot express the difference by saying that Roger Bad, as opposed to Roger Irrational, has reasons for his intention. It is true that Roger Bad cannot be fairly accused of the specific sort of irrationality that Roger Irrational's intentions exemplify. We should resist the conclusion that, just because Roger Bad can escape a certain charge of irrationality, he must have reasons of a certain kind to buy the poison.

Humean accounts on motivation accept that the mere fact that there is a desire/intention of Roger Bad that is subserved by his buying the right kind of poison, gives him a special kind of reason to enter the right kind of shop. We can describe this Humean strategy by saying that, according to it, the mere fact that Roger Bad's practical inference tracks an efficient means-ends relation guarantees that his original desires provide reasons for his practical conclusion.<sup>1</sup> Standard anti-Humean accounts do not accept this connection. But they fall in a trap

<sup>1</sup> In fact, certain Humean accounts would consider this condition as too restrictive. They would hold that Roger Bad still would have certain kinds of reasons to enter the wrong shop if he had the adequate set of justified, but false beliefs.

when they try to look for reasons somewhere else. A typical anti-Humean approach is to accept that the mere generic desire presupposes some reason to desire. According to this point of view, the source of the reason Roger Bad has is not the fact that he desires to get a copy of Hume's *Treatise*. It is some valuable aspect that he thinks there is in having the book, for it is impossible to have this desire without finding desirable some aspect of the desired object. Another, slightly different, position is to oppose Humeanism by distinguishing two kinds of desire: motivated, or reason sensitive, desires, desires for which the agent has reasons that are not desires, and unmotivated desires. Motivated desires are those that can be rationalised. In these motivated desires at least, the reasons the agent has for the conclusion of practical inference is transmitted from the reasons the agent has for the original desire. I will argue that there are some common mistakes in both approaches.<sup>2</sup> Both share the crucial assumption that Roger Bad must have some kind of reason for his intention.

There is a strong intuition in favour of the anti-Humean camp: it is difficult to understand the idea that to desire/will/intend to  $\phi$  is, by itself, a reason to  $\phi$ . It seems that we cannot get reasons in such an easy way. This cheap, easy way of getting reasons for our actions has been described by M. Bratman as 'bootstrapping'.<sup>3</sup> It does not seem difficult to agree with Bratman: bootstrapping is illegitimate. For, without reasons for the end, we do not have reasons for the means. And, if bootstrapping has to be avoided, we should not say that willing an end is always a reason to take certain steps that subserve this end. On the contrary, the standard Humean intuition is that, after all, there must be some special sense of 'reason' according to which some reasons can escape this diagnosis. In other words, a sense of 'reason' in which we can get reasons from the goals that we accept, in which the mere fact that an end is wanted constitutes a reason for taking the subserving steps:

<sup>2</sup> For the argument that motivated desires presuppose reasons, see Nagel 1970. For recent versions of the idea that desiring requires believing that there is some aspect in the desired object that is worth desiring, see, for instance, Quinn 1993, Scanlon 1998, Moran 2001.

<sup>3</sup> Bratman 1987.

'R at t constitutes a motivating reason for an agent A to  $\phi$  iff there is some  $\psi$  such that R at t consists of a desire of A to  $\psi$  and a belief that were he to  $\phi$  he would  $\psi$ .' (Smith, 1987: 36)

Yet, it is not easy to understand which kind of reason can be identified by this stipulation. Friends of motivational reasons tend to tell them apart from the standard, justificatory sense of 'reason'. I doubt that following the spirit of this stipulation could give us a coherent sense of the term. To see this, consider now Roger Good. Roger Good is a decent human being and a competent philosopher longing for an old copy of Hume's *Treatise*. He also knows that his neighbour has one of those rare editions. And he also knows, as Roger Bad does, that poisoning his owner is a very easy way of getting his copy. There is nothing wrong with Roger Good's motivational set. There is nothing wrong in longing for this kind of book. Also, he has perfect knowledge about certain instrumental connections. Look again at the previous stipulation about the conditions under which someone has motivating reasons. Roger Good satisfies all of them. Even if he has not the slightest desire of poisoning his friend, he has a perfect motivating reason to poison him. What could this exactly mean? Well, it can certainly mean that there are certain possible worlds in which a nasty counterpart of Roger Good, for instance some Roger Bad, forms the desire, or the intention, or even acts with the intention, of getting the book by poisoning his friend and that, if this were the case, then this process could be rationalised by mentioning the relevant desires and beliefs, namely those desires and beliefs that Roger Good — a perfectly decent human being — already has in the actual world. There is, then, a possible world in which the motivational set that Roger Good actually has can rationalise the intention-of-poisoning-a-neighbour. The desires and beliefs of a decent man longing for a book have the power of rationalising a possible intention of killing someone. This could be a sense — a strange sense — of 'motivating reason.' In this sense, *we all actually have* millions of motivating reasons to do the nastiest things we can imagine. I actually have thousands of motivating reasons to kill my wife, to end the human race, to torture an innocent child, etc., even if I have not the slightest desire to do those nasty things.

It is not an accident that, even those philosophers that use the notion of a motivating reason, do not seem to think that the previous stipulation conveys what they really have in mind. If we were to accept the

consequence that Roger Good has a reason of some special kind to poison his friend, even if he has no desire, and no intention, of poisoning him, then it would be difficult to explain — even for the most fervent Humean — what this special kind of reason could be. In fact, when mentioning paradigmatic cases of actually having a motivating reason to  $\Phi$  *without* having normative reasons to  $\Phi$ , those Humean philosophers who are fond of the notion tend to describe a situation in which the agent decides or wants to  $\Phi$ , moved by his desire to  $\Psi$ , when there is something wrong in trying to get  $\Psi$  by  $\Phi$ ing: just (i) because  $\Psi$  itself is a wrong goal, (ii) because, even if  $\Psi$  is not wrong, it is wrong to  $\Phi$ , or (iii) because the relevant beliefs are false or unjustified. In the paper where Smith introduces the quoted stipulation about necessary and sufficient conditions for having a motivating reason, every given example is of one of those types. There is no single example similar to the case of Roger Good: a case in which an agent has a motivating reason even if he is not minimally moved by it. Given that this would be a case that, according to Smith's own stipulation, should be considered an instance of having this kind of reason, it is quite legitimate to wonder why the selection of examples that are supposed to illuminate the notion of motivating reason is so biased. And anyone who is familiar with contemporary philosophy of action must accept that this is not an eccentricity of Smith, it is, on the contrary, the common strategy: the clear cases of *having* motivating reasons are supposed to be just cases of being motivated without having justificatory reasons, not clear cases of having some kind of non justifying reason that does not move the agent that has the putative reason. The clear cases of having a motivating reason are just cases of being moved by the putative reason. Obviously, something wrong is going on here. Whatever they say, it seems that the phenomenon that philosophers have in mind, when they introduce the notion of motivating reasons, is the phenomenon of being in fact motivated: the kind of reason that Smith identifies is constitutively linked to the fact that it moves the agent in a minimal way.

Be that as it may, if, contrary to the literal meaning of Smith's stipulation and according to the spirit of his own examples, we said that Roger Good has no motivating reason to kill someone, then we would have still not identified a proper relation of having a reason, independent from the fact of being motivated by it. Roger Good is certainly in a motivational state that is *potentially* explanatory. It is potentially explanatory and it *would* rationalise the decision of poisoning someone made by his counterpart in a counterfactual world. How can anyone

have ever thought that those truisms justify the conclusion that his actual state (a state in which he actually is) includes a reason he has to poison someone? This conclusion crucially depends on considering the counterfactual situation in which he, or his counterpart, decides to poison someone to get the book. In this counterfactual situation, we can clearly talk about the reasons for which he acts. So, the clear cases of motivating reasons are cases in which the putative reasons are in fact motivating. Their being motivating states is fixed by the fact that they are in fact motivating. We can, if we want, say that those states have explanatory, motivating *powers*, even in the case in which they are not motivating at all. But, still, it looks very strange to say that being in one of states, when they are not motivating at all, is having a reason of a special kind. To say, for instance, that I do have a reason of a special kind to kill my best friend in order to steal a book, when I do not have the slightest desire of acting in this way.

Notice that, under the previous assumption, I am supposed to have thousands of motivating reasons to kill my best friend, or my wife, or my neighbour *without* having any desire of killing them. I can have a motivating reason for  $\phi$ ing without having any desire of  $\phi$ ing. Is this acceptable for a Humean? It does not seem so. It seems that no Humean should be tempted by this picture. The standard argument for motivational Humeanism relies on the idea that desires are, by themselves, a special kind of reasons to act, just because, as opposed to beliefs, they have the right direction of fit. This is why Smith, for instance, has argued that desires are motivating reasons (Smith 1994:116). Putting aside terminological stipulations, the Humean story is that those special reasons cannot be dissociated from desires, since to desire is to be in a state in which the world must fit.

There are many ways of showing that the Humean story is incoherent. I am just pointing to one that is crucial for the purposes of my argument. The crucial premise of Humeanism requires accepting that nobody can have one of these reasons without desiring. If it is right, there is no way of understanding the explanation of the transference of reasons in practical reasoning that underlies conventional wisdom about motivating reasons. Such transference would entail that I have many reasons to kill my best friends, without having the slightest desire to do such a thing. So, the Humean account of the putative reasons that Roger Bad gets for his practical conclusion is incoherent. It presupposes that to have those reasons requires desiring, while, on

the other hand, it has to accept that Roger Good has reasons to do things that he has no desire to do.

## II

Let us try to generalise now the previous diagnosis. For the incoherence of the idea that Roger Bad has certain reasons for his practical conclusion can be detached from the particular account that Humeanism provides of this putative phenomenon. As I have already mentioned, non-Humean explanations tend to assume that the relevant reasons are not provided by desires: according to them, the relevant role has to be played by the content of certain beliefs and certain pro-attitudes (evaluative judgements, for instance). Desires only play a role in fixing the reasons that the agent has to act insofar as they presuppose a certain evaluation about certain desirable aspect of the desired object. Be that as it may, the problem is still to try to understand the sense in which standard rationalisations that mention pro-attitudes and beliefs make the idea that the agent has (or believes she has) certain reasons for her practical conclusion intelligible. If I am right, the inconsistency of standard non Humean accounts is very similar to the one I have diagnosed in Humean analysis.

Consider, for instance, the two following pieces of practical reasoning:

### A

- (a) I wish I had his old copy of Hume's *Treatise*.
  - (b) If I poison him I can easily get his old copy of Hume's *Treatise*.
- 
- (c) I shall poison him as a way of getting the copy

### B

- (a\*) His old copy of Hume's *Treatise* is a valuable object.
  - (b) If I poison him I can easily get his old copy of Hume's *Treatise*.
- 
- (c) I shall poison him as a way of getting the copy.

A and B are different. If you are a Humean, you would tend to think that the specific cogency they have as pieces of practical reasoning

depends on the fact that pattern B is grounded on something similar to pattern A. If you are not a Humean, you would tend to think just the opposite. You would tend to think, then, that this particular cogency can only be accounted for if there are some valuable aspects under which the desired object is desired: something like pattern B must underlie, then, pattern A. Either way, it is commonly accepted, under the assumptions that have just been commented, that the special cogency of those pieces of practical reasoning, as pieces of practical reasoning, crucially depends on the fact that the agent gets some reasons for (c), from (a–a\*) and (b). I think this is wrong. The basic problem with this model of practical reasoning is quite independent of the option (Humean or not Humean) of considering that desires are reasons-providers or not. Even if we say that pro-attitudes require certain kind of evaluative judgement (as in B), we still have to face the basic problem: generic pro-attitudes are used to provide reasons for *more specific* intentional contents.

The common assumption that, by following one of those patterns (A or B), the agent gets proper reasons for the practical conclusion, involves a no less common understanding of the relation of having reasons: having those reasons is different from forming the intention of acting for them in a specific way. Roger Bad does in fact form the intention expressed in (c), and he acts for those reasons he has. On the contrary, Roger Good does not want to act for those reasons he has. He acts for some other reasons he has. Nevertheless, all the reasons that are supposed to be relevant for the rationalisation of Roger Bad's intention are reasons that Roger Good has. Just because both of them accept the premises that are introduced in the model A-B. We could say, of course, that Roger Good acts for other reasons he has. But then we cannot insist that the relevant reasons, that are enough to show that Roger Bad's piece of practical reasoning is minimally cogent, are just the premises (a–a\*) and (b). This assumption is incoherent, because the choice of a specific way of acting requires some difference in the relevant pro-attitudes (desires, if you are Humean, or value judgements, if you are not Humean) that both (i) must rationalise, in the relevant sense, the final intention and (ii) cannot be rationalised, at all, by the generic pro-attitude. They cannot be so rationalised, because, in that case, it would also be cogent for Roger Good to form the intention (c) in spite of the fact that he has no specific desire of getting-the-book-by-poisoning-a-friend. If the generic pro-attitudes *can* be relevant for the cogency of the piece

of practical reasoning, those more specific attitudes *must* also be relevant, in the very same sense. And so it cannot be that Roger Bad's intention is rationalised in the relevant way just by patterns A-B.

The point is not that the attitudes that correspond to the premises in A and B generate only an incomplete explanation of the practical conclusion. Let us assume that different agents accept the premises and form different intentions, depending on many other factors in their motivational set. In the case of Roger Good, he has those reasons, the reasons that are mentioned in the premises of A-B, but they are not the reasons for which he acts. I am also prepared to concede, for the sake of argument, that those putative reasons could be interpreted as *pro tanto* reasons, reasons that could be overridden by more demanding reasons that the agent has. But still: the present assumption is that Roger Bad's intention can be rationalised by pattern A-B. Roger Good's very different intention of buying the book can be rationalised by different premises. The problem is that Roger Good *does accept* the premises that are mentioned in pattern A-B. The only option is to point out that the difference in rationalising force depends on the fact that he also accepts different premises. But this would be, under the present assumptions, self-refuting. For now we need to assume that those premises only manage to rationalise Roger Bad's intention because of many other desires/values that he puts in the particular way of getting the book: because he also accepts many other premises that are not mentioned in pattern A-B. The mere fact that the intention that Roger Bad in fact forms is rationalised by an instrumental belief (an instrumental belief that Roger Good shares even if it is not the belief that rationalises his intention) shows something important. Roger Bad needs some more specific pro-attitudes towards some specific ways of getting the book. By hypothesis, Roger Good does not have any desire to poison anyone. If you are Humean, you will say that Roger Bad must have this more specific desire about a particular means. If you are anti-Humean, *mutatis mutandis*: you will accept, at least, that Roger Bad has to make some different, specific value judgements. In any case, the more specific reason for his intention is not described by the model of practical reasoning we are now considering — the pattern A-B. And this fact crucially breaks the kind of asymmetry between the two premises required by the model. The function of the second premise cannot just be to fix certain instrumental connections; under the present assumptions, the 'epistemic' premise becomes relevant because it conveys information

about the presence of some other conative attitudes that the agent has to form the practical conclusion. Those conative attitudes would generate or presuppose, under the present assumptions, reasons that would have the same right to be introduced in the model as the generic pro-attitude (a-a\*) has. There is no way in which, under the present assumptions, Roger Bad's piece of practical reasoning could be cogent if he had not some pro-attitude about the specific way of subserving the generic goal of getting the book. And, then, it is impossible to maintain that Roger Bad's intention is cogently rationalised by the *mere fact* that it is formed according to our pattern A-B.

In the next section, I will try to show the last source of these difficulties. I will defend that we must give up the very idea that (a-a\*) provide reasons for (c). The fact that our intentions can be rationalised does not entail that we have reasons of any special kind to form our intentions. The reasons that we discover when we engage in a successful process of rationalisation are just more refined descriptions of the content of the intention that the agent has formed. They are not, in any relevant sense, reasons he had to form this intention.

### III

There is then a crucial link between what I have described as standard interpretations of the practical syllogism and certain issues about intentional action. Which is the sense of 'reason' in which Roger Bad's decision is supposed to have reasons? Why can many different, incompatible, intentions be rationalised in the same way by the same conative premises in practical inferences? My diagnosis is that certain basic confusions about the transference of justification are connected to a more basic confusion about the general phenomenon of rationalising explanations. The grain of truth behind the idea that intentional action is (normally) action for reasons is that intentional action accepts certain paradigmatic why-questions, questions that ask the agent to specify the content of the intention-with-which she acts. I will defend that the typical answer to those questions, the answer that the rationalising explanation is expected to provide, does not mention the reasons that the agent has to form the intention. Those putative reasons that are mentioned are just determinations of the content of the intention. Of course, they can make intelligible what the agent is doing, just by showing that certain particular actions of hers have a

purpose. That the agent sees them as steps in a wider plan she is in fact following.

To see what I mean, let us consider some basic cases. It is not difficult to accept, I guess, that, at least in certain cases, a typical why-question is correctly answered by providing information about the intention-with-which-the agent acts. For instance, to adapt an example by Davidson, let us imagine that Ada presses the key with the number 7 in her computer because she wants to calculate the cube root of 728 in her computer, now. Asked the relevant why-question, she tells us that the reason why she acts in this way is that she wants to take the cube root of 728 in her computer, now. There is no way of understanding what she says, except that she acts with the intention of taking a certain cubic root. In such basic cases, there is no room for strange deviant causal chains. Also there is no gap between the reason she has and the reason for which she acts. She cannot press that key because she wants to take the cube root of 728 in her computer now, and act with some different intention. If that were the case, we could say that she has stopped wanting to take this cube root.<sup>4</sup> So, in certain basic cases, the want that we attribute to a subject has all the marks of intention-in-action. In those basic cases, it seems obvious that the fact that an action can be so rationalised does not mean that the agent had any reason to form the intention. The function of the rationalising explanation is to tell us what the intention is. In the ordinary sense of 'having a reason', it might be true that Ada presses that key in her computer because she wants to take the cube root of 728 in her computer now, while she has no reason to do so. Of course, we can say that the reason for which she acts is that she wants to take the cube root of 728 in her computer now. But, if the previous lines are right, this is a very specific sense of 'reason'. In this sense of 'reason,' her reason is just a part of the intention with which she acts. Obviously, it is not the reason why she made up her mind. I am not deny-

<sup>4</sup> Davidson 1982: 263. This entails that the kind of Davidsonian challenge that has dominated contemporary discussions on action is misguided. There are certain Davidsonian reasons (certain wants) that have the following feature: having them is not independent from acting for them. So it is not true that all Davidsonian reasons are such that an agent could have had them without they being the reasons for which the agent acted. And all the interest of the Davidsonian challenge depends on its generality. If there are some reasons to which it does not apply, it loses all its putative force.

ing that our agent might have had such a reason. I am just saying that the existence of this kind of reason is not guaranteed by the mere fact that her intention can be rationalised.

Rationalisations, then, at least in certain basic cases, do not inform us about the reasons that the agent had to make up her mind. The reasons that proper rationalisations convey can be reasons for which an agent acts, without being reasons the agent had to make her mind up, before forming the relevant intention. Let us call them ‘content-determining’ (‘CD’) reasons. CD reasons are not proper reasons, in the sense that an agent can act for them, or form the corresponding intention for them, without having reasons to act or to make up her mind. We can say that Ada’s CD reasons, then, are not proper reasons, in the sense that they are not reasons, not even reasons for Ada, to act in a certain way, or to form an intention. Just because those reasons do not exist before Ada is motivated to form her intention. It is true that Ada might have had proper reasons for her action. It might be, for instance, that the happiness of her children might depend on her acting as she did, it might be that some valuable goal might be attained by her action. On most views about what reasons are, those facts would count as reasons. They would be reasons she would have had, quite independently of how she decided to act. This is possible, but it is not guaranteed by the mere fact that Ada’s action can be rationalised in the way we have just done. The fact that her intention can be rationalised does not guarantee the presence of proper reasons.

Let us consider now the following answers to a typical why-question that an honest and knowledgeable agent might give, when asked by a friend to explain his presence in Barcelona’s station:

- (1) I have the intention of going to Paris tomorrow.
- (2) I am buying a train ticket because I have the intention of going to Paris tomorrow.
- (3) I am buying a train ticket because I want to go to Paris tomorrow.
- (4) I am buying a ticket for tomorrow's train to Paris because I have just remembered that Maria will be in Paris tomorrow.
- (5) I am buying a ticket for tomorrow's train to Paris because I have always had the desire to meet Maria and she will be in Paris tomorrow.

It seems that, in the sense in which (1)–(3) specify a reason for action, the specified reason is just the content of the intention with which the agent acts. Somebody could say that this is not necessarily what those

statements say. For instance, (3) could be literally true, even if I was not buying a ticket for Paris. The fact that I was buying a ticket for *London* might have been, via a complicated causal story, a causal consequence of the fact that I wanted to go to Paris. In this case, (3) would not give us information about the intention with which the agent acts. My diagnosis of this possibility will be crucial for my argument in this section. In this case, the person who is asking for the rationalisation would feel cheated. That is, (3) would still be an explanation, but not the kind of rationalisation that my interlocutor has the conversational right to expect.

To see this more clearly, look now at cases (4)–(5). They might seem quite different. In them, a reference to desires and beliefs seems to work as a way of specifying the reasons why the agent acts. There are, of course, certain differences, but, in my opinion, they are systematically misunderstood in contemporary literature. It is true that (5) can be a conversationally appropriate answer to a paradigmatic why-question ('Why are you going to the station?'). The Humean is obviously impressed by (i) the fact that a similar answer — an answer that mentions desires, for instance — is always possible and (ii) the fact that desires seem obviously appropriate to explain behaviour — the 'direction of fit' intuition. Conventional forms of anti-Humeanism insist that the explanatory role of desires is not the relevant issue when we think of the reasons why we intentionally act. Even if my desire causes my intention, the fact that I desire is not, at least not normally, my reason to act. I agree, but this is irrelevant for the typical anti-Humean conclusion, the conclusion that in those cases the true reasons for action must be some reason-providing features in what is desired.

Of course, it might seem that (4) and (5) are appropriate answers to the typical why-question that asks for the reasons for which we act. The diagnosis is quite simple: they are a perfectly appropriate way of conveying the relevant information about those reasons (the CD reasons). If they manage to fulfil this role, it is not because desires or beliefs are reasons, or because they presuppose the existence of reasons that the agent had. It is because the agent is exploiting her conversational commitment to provide relevant information about the intention with which she acts. She is just talking about the special causal antecedents of the action that are internally linked to the purpose that guides her. There are hundreds of causal antecedents of the action that might be mentioned and that are, nevertheless, quite

irrelevant for the information that the agent is expected to provide. By *choosing* this particular causal antecedent the honest agent conveys information about the relevant kind of reasons she is being asked for: the CD reasons that fix the content of her intention.

On the one hand, as Davidson remarked, the fact that I have had certain desires or beliefs can hardly count as a reason that is necessarily effective. I could have had them without being moved in the corresponding way. On the other hand, we have the intuition that mentioning desires and beliefs is a standard way of conveying the intention, the goal or the purpose with which the agent acts. Both things are true. We can say that there is a mechanism of conversational implicature by which the agent conveys her intention just by selecting as relevant a certain causal antecedent of the action. Our intuition that (4) and (5) are ways of conveying CD Reasons goes hand in hand with our intuition that the agent tries to exploit this conversational mechanism. And there is a crucial test for my thesis: if the agent cancels the corresponding conversational implicature, then we are forced to stop assuming that (4) and (5) convey information about the relevant CD reasons. For instance, the agent might add to (4):

Even so, I do not go to Paris with the intention of meeting Maria. It is just that my thoughts about Maria and Paris made me remember that I promised Joanna to visit her in Paris before the end of the month.

By doing this, she is not contradicting herself. Simply, she has cancelled any right to assume that the intention was to meet Maria. My main point is that we all share the intuition that, when the conversational implicature about the relevant CD reasons is cancelled, what is shown is that the original answer was not the right rationalisation. Certainly, it was not the kind of explanation that the person who asked the why-question in a normal case was looking for. She had the right to feel cheated if someone used this mechanism of conversational implicature to give causal information that was literally true without conveying the relevant information about the CD reasons. Rationalisations do look, then, for CD reasons. But those reasons, as I have been arguing, are not reasons that the agent had for making up her mind. They are just the content of the intention with which the agent acts.

## IV

There is, I have been suggesting, an interesting connection between the role of rationalisations of intentional action and the structure of paradigmatic cases of practical reasoning. In the sense in which intentional action can always be rationalised, this rationalisation only gives us the content of the intention with which the agent acts. In what has usually been accepted as a standard case of practical syllogism, the relevant connection between the premises cannot be properly described as instrumental. The agent does not choose a specific way of satisfying a goal because of its instrumental virtues. The agent, simply, makes his mind up, forming a specific intention. This process can be ideally rationalised by something similar to the mention of a generic goal and a particular way of subserving it. But this idealization is just a way of specifying, in a given dialectical context, the content of the intention. It does not follow that his having a goal is a (believed) reason to choose a particular means, nor that the instrumental virtues of the specific means are — given the goal- a reason to choose it. If I have no reason to choose a goal, I won't have reasons to choose particular ways of subserving it. And, even considering only the cases in which I have reasons to choose a goal, the mere instrumental virtues of the selected means cannot be a reason to choose this particular means.

If this is right, there is nothing like a logical form of pure instrumental practical reasoning, just because there is nothing like pure instrumental practical reasoning. No rational agent chooses a means just in virtue of the instrumental relation with a given goal.<sup>5</sup> As I have just said, the problem is not only the well known difficulty that this instrumental relation could not provide reasons when the agent had no reason to pursue the goal: nothing is gained if we try to bypass this difficulty by reducing the scope of pure instrumental rationality to a certain kind of relativised rationality, the specific rationality that is conditional or relative to a given goal. If the instrumentally adequate means were not sensitive to some other kind of value, then we would have to say that chopping my head off is a rational means of stopping my mild headache. If, on the other hand, we insist that the instrumental connection has to be sensitive to some other value of the possible

<sup>5</sup> For an interpretation of Aristotle's *Nicomachean Ethics* that would be congenial with my account, see Wiggins 1987.

means, then the relativisation to a given goal is a completely empty move. For, then, there is no interesting distinction between (i) the case in which the most rational thing to do is to choose a not-very-efficient way of satisfying the goal, and (ii) the case in which the most rational thing to do is just to give up the goal, because no efficient way of satisfying it would be acceptable. Whatever is meant by 'the most rational thing to do' in (i) is not different from what it is meant by 'the most rational thing to do' in (ii). And this entails that we have still not described the kind of instrumental practical rationality we are looking for.<sup>6</sup>

This conclusion could seem trivial. It is difficult to understand the idea that there is a kind of practical reasoning whose point is to look for the way of satisfying our previous goals at any cost. In fact, we pursue certain generic goals just because we are confident that among the possible ways of satisfying them there are some that can be accepted by us. So, we cannot explain our acceptance of them in terms of a putative instrumental relation. If I have been right, the reluctance to recognise this fact that most part of contemporary literature on action shows can only be explained by a misreading of the surface grammar of paradigmatic cases of rationalisation of intentional action. Once we assume that intentional action requires having reasons to act, we are tempted to look to the instrumental relation between the chosen means and the

<sup>6</sup> Broome 2002 has argued that a cogent piece of practical reasoning does not transfer reasons from the premises to the practical conclusion. Under certain terminological stipulations, I agree: for, once we assume that the successful process of rationalisation of a particular intention entails a corresponding ideal piece of practical reasoning, I endorse the idea that the agent can form the rationalised intention without having reasons for it. Nevertheless, it is important to notice the differences between our arguments. His way of describing the special cogency of instrumental practical reasoning depends on the idea that the practical conclusion must satisfy certain normative requirements: if I have the intention of doing X and I know that doing Z is a necessary means for doing X, then I should not give up Z while keeping the original intention. Even if this is true, it cannot cast light on the general phenomenon of practical reasoning, just because, as Broome himself recognises, in most common cases the agent does not choose a necessary means for the satisfaction of his goal. In fact, I would not say that there is anything practical in the satisfaction of the normative requirements that Broome mentions: the kind of cogency that he analyses in terms of the satisfaction of normative requirements is, in the end, a consequence of (i) the mere relation of instrumental necessity, and (ii) a crucial aspect of the concept of intention: I cannot have the intention of doing X, while at the same time knowing that my present way of acting is incompatible with my doing X.

pursued end as a special source of reasons. This move loses all its appeal when we realise that standard rationalisations of intentional action do not try to specify the reasons that the agent has to act. Intentional action does not require having reasons to act.<sup>7</sup>

Josep L. Prades  
 Departament de Filosofia i Filologia  
 Facultat de Lletres, Sant Domènec  
 Campus de Barri Vell  
 Universitat de Girona  
 17071 Girona, Spain  
 josepll.prades@udg.edu

### References

- Broome, J. 2002. Practical Reasoning. In *Reasons and Nature: Essays in the Theory of Rationality*. Edited by J. Bermúdez and A. Millar. Oxford University Press: Oxford.
- Bratman, M. 1987. *Intentions, Plans and Practical Reasoning*. Cambridge, Mass.: Harvard University Press.
- Davidson, D. 1982. Hempel on Explaining Action. In *Essays on Actions and Events*. Oxford: Clarendon Press.
- Moran, R. 2001. *Authority and Estrangement*. Princeton: Princeton University Press.
- Nagel, T. 1970. *The Possibility of Altruism*. Oxford: Clarendon Press.
- Quinn, W. 1993. Putting Morality in its Place. In *Morality and Action*. Cambridge: Cambridge University Press.
- Scanlon, T. 1998. *What we Owe to Each Other*. Cambridge, Mass.: Harvard University Press.
- Smith, M. 1987. The Humean Theory of Motivation. *Mind*, 96: 36-61.
- Smith, M. 1995. *The Moral Problem*. Oxford: Blackwell.
- Wiggins, D. 1987. Deliberation and Practical Reason. In *Needs, Values, Truth*. Oxford: Basil Blackwell.

<sup>7</sup> In no way do I try to question the possibility of a philosophical theory about what is usually described as 'substantial rationality.' Many times we have reasons to act. And sometimes we act for the reasons we have. I have just argued that the fact that an intentional action can be rationalised does not mean that the agent had, or believed she had, reasons to act.

# What is normativity?

**John Skorupski**  
University of St Andrews

## **Abstract**

The thesis that the concept of a reason is the fundamental normative concept is in the air. In this paper, I examine what it amounts to, how to formulate it, and how ambitious it should be. I distinguish a semantic version, according to which any normative predicate is definitionally reducible to a reason predicate, and a conceptual version, according to which the sole normative ingredient in any normative concept is the concept of a reason. Although I reject the semantic version, I examine its potential in some detail. And I claim that the conceptual version is plausible.

## 1. Normativity and reasons

The concept of a reason is the fundamental normative concept: this thesis is in the air.<sup>1</sup> One of its attractions, for me, is that it supplies illuminating terms for an old distinction: the distinction between the descriptive and the normative. It has been the fashion to deprecate any such distinction, but in my view its foundations in the critical epistemological tradition, whether empiricist or Kantian, remain as strong as ever. Putting the contrast in terms of propositions about the world and propositions about reasons, as against ‘is’ and ‘ought’, or ‘fact’ and ‘value’, is a helpful way of stating two ideas I would want to defend. The first is that normative propositions are not themselves descriptive or factual — they do not present more facts, they do a different job. The second, contrary to the non-cognitive (emotivist, prescriptivist, voluntarist etc.) strains in the critical tradition, is that

<sup>1</sup> I am using the word ‘normative’ broadly to contrast with ‘descriptive’ — not narrowly, as some writers use it, to contrast with ‘evaluative.’ As to the thesis, although it is in the air it is difficult to attribute definite versions of it to particular people. I have been helped in thinking about it by Gibbard (1990) and Scanlon (1998). Gibbard works with the basic notion of what is ‘rational,’ or ‘what makes sense,’ while for Scanlon the basic notion is explicitly that of a reason. (Substantively this may not be a big difference, though other substantive differences between them are indeed big).

they are nonetheless straightforward, genuine, propositions. Both ideas become, I think, a little more persuasive when we focus directly on propositions about reasons, in their full generality: reasons for belief, reasons for feeling, reasons for action — although I should emphasize that neither of them is thereby forced.

This semantic and epistemological background will not however be our topic here.<sup>2</sup> The thesis that something's being a reason for someone is the fundamental normative concept — as against for example obligation, or value or 'rationality' — is independent and can be assessed in its own right, whatever one thinks of those background issues. One can favour this thesis on more specific grounds; for example that it clears up the obscure relation of normative supervenience (which simply reduces to the reason relation) or that it opens up a plausible sentimentalist account of value in terms of reasons for feelings. And then there is the underlying, stage-setting thought that 'normativity' can be nothing more than that by which autonomous — reason-sensitive — agents steer. It must come down to their reasons for belief, feeling and action: to the normative relation between facts on the one hand and those reason-responsive acts or states on the other. The elaborations of normative discourse must issue in propositions about reasons or they are idle wheels.

But I shall not discuss these persuasive considerations (as they seem to me) either. My aim in this paper is simply to examine what the thesis amounts to, how to formulate it, and how ambitious it should be.

I should note that we cannot simply assume there is only one way of interdefining the circle of normative concepts. If it turns out that reductive analysis of normative concepts can be done in more than one way, of which reduction to reasons is one, we will not have shown that reasons are *the* fundamental normative concept. I myself do not believe it can be done in any other way, but I am only going to examine here whether a reduction to reasons is possible. Call this claim the *Reasons Thesis*. If it is sound, I leave it as a challenge to others to show that some other reduction, for example to obligation, value, or rationality is possible.

<sup>2</sup> I discuss it in Skorupski, 2006.

## 2. Semantic reduction and concept-possession

The simplest and sharpest version of the Reasons Thesis is explicitly semantic: any normative predicate is definitionally reducible to a reason predicate (as well as non-normative predicates if it is not itself wholly normative). Call this the *semantic thesis*.

An alternative is to put the Reasons Thesis at the level of concepts: the sole normative ingredient in any normative concept is the concept of a reason. This claim is weaker, in that the semantic thesis entails it but the converse need not hold.

To illustrate the difference, consider the moral concepts, which are obviously a major case in any general discussion of normativity. Let us allow that the basic concept here is that of moral wrongness: a moral obligation, for example, is by definition something with which non-compliance is morally wrong. And let us also allow for the sake of argument that the morally wrong is that which (absent extenuating circumstances) there is reason to blame. So that promises a reduction of moral claims to claims about reasons. But it may plausibly be objected that it cannot yield a *definition* of ‘morally wrong’, since ‘blame’ in the relevant sense should be defined as that sentiment which is appropriate — which there is reason to feel — towards the *morally wrong*. A good objection, but not the end of the story. Suppose that there is a distinctive sentiment of blame whose object is the morally wrong; suppose also that we can be given an independent characterization of this sentiment, say in terms of the actions to which it disposes, or by having it explained as the sentiment we feel when we consider some specific paradigm cases. Suppose, finally, that when we have been made familiar, in the first-person way, with the sentiment, we find ourselves able to go on spontaneously, and reasonably convergently, making confident new judgements about when there is reason to feel that specific sentiment. In that case we have everything we need to grasp the concept of moral wrongness. The morally wrong is that which is blame-sentiment-worthy in the absence of an excuse. The concept can in this sense be exhaustively captured in terms of the concept of a reason and the concept of a certain sentiment, even though ‘morally wrong’ cannot be defined in terms of ‘reason’. Furthermore if we can individuate the sentiment by these methods we can also use them to introduce a term to refer to the sentiment, say ‘BS’. The morally wrong is that which is BS-worthy

(which there is sufficient reason to respond to in that way) in the absence of an excuse.<sup>3</sup>

Colour concepts provide a partial analogy. To grasp the concept of yellow you have to be familiar in the first-person way with the sensation of yellow, and take the presence of the sensation as a reason to think it indicates a particular objective feature, while also being able to recognise cases in which that reason is defeated. If you have all that, you have what's needed to possess the concept. Nonetheless, 'yellow' is not definable in terms of 'sensation of yellow'. 'Yellow' is semantically simple, 'sensation of yellow' is complex; you need to know what the former means to know what the latter means. G. E. Moore was right about the indefinability of 'yellow' (if not about the indefinability of 'good').

The partial analogy with moral wrongness is that a being which has the blame-response, and an ability to distinguish between having it and having reason to have it, has what is needed to possess the concept of moral wrongness, even though 'morally wrong' is not definable in terms of 'what there is reason to blame'.

This is an approach that captures a concept by saying what is required for possession of the concept, rather than by the semantic route of explicit definition. Nonetheless, if the story is right we have shown that the normative concept of moral wrongness can be fully captured in terms of the normative concept of a reason, and in this sense made progress in defending the Reasons Thesis. (Attributing full possession of *either* concept, that of wrongness or that of yellowness, involves attributing reason-sensitivity to its user, in that it is a condition for grasping a concept that one is sensitive to the reasons that warrant applying it. What distinguishes moral wrongness as a *normative* concept, unlike yellowness, is that you can have the concept only if you are sensitive to reasons for the blame-sentiment itself. It always makes sense to ask whether there are reasons to feel the blame sentiment; with the sensation of yellow the analogous question makes no sense.)

If these points about the moral concepts are right, their normativity cannot be captured through a direct semantic reduction of moral predicates to reason predicates within our actual language. I shall come back to this in section 9. On either approach, however, we are

<sup>3</sup> See Skorupski, 2005, for further discussion of blame.

going to need a definition of normative predicates in terms of reason predicates. In sections 2 — 8 I want to consider how to construct such a definition, and how far it can take us in accounting for predicates that strike us intuitively as normative. In other words, I want to explore how far we can take the semantic thesis.

### 3. The semantic thesis: preliminaries

This project encounters two difficulties at the outset — controversy about what terms are normative, and their indefinite extent on any view. The circle of terms that most people would recognise as normative contains a vocabulary that is diverse and wide. There are normative terms that span all fields — including epistemology, ethics and aesthetics — and whose normativity is pretty uncontroversial, at least in their most salient uses: ‘good’ and ‘bad’, ‘right’ and ‘wrong’, ‘ought’ and ‘should’. But then there are terms special to a field, as to whose normativity or otherwise there can be significant dispute. In epistemology, for example, we have ‘evidence,’ ‘a priori’, ‘valid’, and perhaps a certain sense of ‘probable’. Whether these are normative is a question of fundamental significance, lying at a philosophical crossroads. Suppose we take them as normative, and apply the semantic thesis. Evidence that  $p$  consists of accessible facts that gives one reason to believe that  $p$ . A proposition is a priori if there is outright reason to believe it, that is, on the basis of no facts. Validity is a priori truth-preservation. Probability, in one of its senses, is a matter of degrees of reason to believe. Clearly these suggestions are not philosophically neutral. A broadly ‘Kantian’ epistemology that takes normativity as a precondition of factual discourse is only just below the surface.

Still, we can distinguish, at least initially, between a dispute about whether certain terms are normative and a dispute about whether normative terms satisfy the semantic thesis. What then remains is the problem of indefinite extent. Proponents of the semantic thesis cannot be expected to consider all putative normative terms individually. The most that should be expected is that they define the ‘thin’, topic-neutral ones, and then consider representative classes of ‘thick’ or topic-specific ones. So the semantic thesis will inevitably have an element of the stipulative to it, however thoroughly it is carried

through. But if it is carried through convincingly enough, the final stipulation will seem natural and useful.

To advance the thesis, then, we need to show how certain prominent types of *prima facie* normative terms are definable in terms of reason predicates. What are these reason predicates? I think there are three. First, one can say that some particular facts are a reason, weaker or stronger, for some person at some time to  $\psi$ . Second, one can say that taking everything that counts for and against into account there is more or less strong overall reason for a person at a time to  $\psi$ . And third, one can say that some facts give a person at a time sufficient reason to  $\psi$ . That gives us the following relational predicates:

*Specific reasons of degree:* facts  $p_i$  are at time  $t$  a reason of degree of strength  $d$  for  $x$  to  $\psi$  —  $R(p_i, t, d, x, \psi)$

*Overall reasons of degree:* facts  $p_i$  are at time  $t$  overall reason of degree of strength  $d$  for  $x$  to  $\psi$  —  $R_o(p_i, t, d, x, \psi)$

*Sufficient reasons:* facts  $p_i$  are at time  $t$  a sufficient reason for  $x$  to  $\psi$  —  $S(p_i, t, x, \psi)$ .

I do not think these are interdefinable, but I will not argue that here. If I am wrong, the semantic thesis will say that all normative terms are definable in terms of less than three reason predicates, if I am right, then it is these three reason predicates that count for the thesis.

It is also very important for our purposes to note clearly that we don't just talk about practical and epistemic reasons, but that we also talk about reasons for feeling, for affective responses, such as gratitude, resentment and so forth. We can call them evaluative reasons. That we talk in this way is indisputable; evaluative reasons are constantly in our conversations and our minds. So ' $\psi$ ' ranges over actions, beliefs, and feelings.<sup>4</sup> What can still be debated is whether we can get a reduction here, for example of evaluative to practical reasons or epistemic to practical reasons. I am a 'trichotomist' in the sense that I believe these three kinds of reason are not reducible. Thus

<sup>4</sup> With epistemic reasons, by the way, we must introduce a reference to epistemic fields: sets of facts that are discoverable by the inquirer and which give the inquirer reasons to believe this or that. Epistemic reasons are reasons relative to epistemic fields. This will not affect what follows. For more on this see Skorupski, 2006, section 2.

our reduction base consists of three irreducible reason predicates which range over three distinct kinds of state, attitude or act. With these preliminary points in hand we can set about formulating the semantic thesis.

#### 4. The semantic thesis: normative sentences

The three reason predicates yield sentences of the form

$$\begin{aligned} &R(p_i, t, d, x, \Psi) \\ &R_o(p_i, t, d, x, \Psi) \\ &S(p_i, t, x, \Psi). \end{aligned}$$

Call such sentences *atomic* normative sentences, and define an *explicitly* normative sentence as one that is either

- (i) an atomic normative sentence

or

- (ii) a sentence which is built from sentences which include at least one atomic normative sentence, by means of the connectives of propositional logic, quantifiers, and the truth operator ('it is true that').

Now define the class of normative sentences as follows:

- a. Any explicitly normative sentence is normative.
- b. Any sentence which has a normative sentence as a *definitional* consequence is normative.
- c. No other sentences are normative.

By a 'definitional consequence' I mean something narrower than analytic consequence in the usual sense, i.e. derivability by principles of pure logic plus definitions. Definitional consequence is derivability by means of the natural deduction rules for conjunction and the universal quantifier alone, plus substitution of explicit definitions.

Before moving on let me note two points. First, our definition of normative sentences commits us to counting the negation of a normative sentence as a normative sentence. To count these wide-scope negations as normative sentences is admittedly somewhat artificial; it is not obvious, for example, that the denial that there is a reason to  $\psi$

is itself a normative claim. One might accept the truth of such propositions on nihilist grounds. Yet the nihilist about reasons is presumably not taking a normative stance, whereas our definition attributes to him acceptance of some propositions expressed by normative sentences. This is an awkwardness. But we can say that the normative propositions a nihilist accepts are all expressible as wide-scope negations. It is in this sense that he takes no affirmative normative stance: he thinks that there are no non-negative normative truths.

Second, our definition commits us to counting a conditional with a normative antecedent or a normative consequent as a normative sentence. So consider 'If Michelangelo was not Italian, there is reason for you to take up golf'. This normative sentence can be deduced from non-normative sentences alone — in this case, from the non-normative sentence that Michelangelo was Italian. Thus we cannot maintain that no normative sentence is deducible from non-normative sentences.<sup>5</sup> However, we can still maintain two truths that lie behind that thesis, and I think capture what is intended by the 'is/ought' or 'fact/value' distinction. The first truth is that there is no deductively valid argument (relying only on principles of logic) from *true* non-normative premises alone to an *atomic* normative sentence as conclusion. This holds on our account. As to the second truth: that, I suggest, is that no normative sentence can be the *definitional consequence* of non-normative sentences alone. This follows from our definition of 'normative sentence', together with another claim (which I take to be true), that the reason predicates are not themselves definable 'naturalistically', or in descriptive terms.

## 5. 'Ought', 'should' and 'right'

Let us say that any predicate formed from a normative sentence by dropping singular terms is normative. The project is then to show that all and only the predicates we are inclined to class as normative can be shown to be normative in this sense. A first stumbling block,

<sup>5</sup> If we said that the conditional was non-normative, a normative sentence could still be deduced from non-normative sentences alone. For this conditional, together with the non-normative premise that Michelangelo is not Italian, entails the normative conclusion that you have reason to take up golf. The point goes back to Arthur Prior.

however, lies in the very varied way that many of the ordinary-language predicates that come to mind as normative are in fact used.

Take 'should.' Suppose A says 'It should be getting colder soon,' B says 'You shouldn't park here — it's a double yellow line,' and C says 'You should show some consideration for other people's feelings.' Does the word 'should' in these three statements have different senses, or the same sense but different uses? Without meaning to take a firm view about these semantic questions, I am going to talk about normative and non-normative *uses* of ordinary-language 'should,' 'ought,' etc. The question I am interested in can then be put by asking whether what a person says on a particular occasion could have been said — neither more nor less — by using a normative sentence, as defined above.<sup>6</sup>

Take the case of A. If A could have said exactly the same thing by saying 'There's (sufficient) reason to expect it to get colder soon' then he's using 'should' normatively. But without further context it's not clear that that *is* what he's saying. What about B's statement? Suppose B could have said the same thing by saying 'It's illegal to park on yellow lines' — is what he did say normative? Defining normative predicates in terms of reason predicates helpfully focuses the question. On a positivist view of law, at any rate, his use of 'should' is not normative by that definition. For according to legal positivism the fact that a law exists in a population is a fact about certain attitudes, dispositions and actions in that population (and perhaps their causal antecedents). If so, then 'It's illegal to park on yellow lines' has no explicitly normative sentence as a definitional consequence. If what B says could be said by saying only that, then he is not using 'should' normatively. The same goes for statements of convention — 'You should only move the king one square at a time'. If this sentence is used just to state the rules of chess then again 'should' is not being used normatively. What if it simply is not determinate what a person said? It may be simply indeterminate, for example, whether B is just saying that it is illegal to park on double yellow lines, or saying that there is sufficient reason for you not to do so, because it's illegal. In that case it is indeterminate whether A's use of 'should' is normative.

<sup>6</sup> I am assuming that the reason predicates I have listed do not have the fluidity of use that I am allowing to 'should' etc. That seems right, however.

As for C, in a given context he may be saying that you have a moral obligation to consider other people's feelings. I have already suggested that moral concepts constitute an insurmountable objection to the semantic thesis. However that does not settle the question whether what C is saying, on this moral reading, *can* be said by means of normative sentence. I shall come back to the question in section 9.

'Should' has various uses, then, not all normative.<sup>7</sup> The same applies to 'ought'. But 'should' and 'ought' also have salient uses that are normative. In one of these uses, what one should or ought to do is simply *any one of the things there is sufficient reason to do*. In another use, what one should or ought to do is any one of the things that one has sufficient reason to believe there is sufficient reason for one to do. Let us stick with the first, so-called 'objective' use. The practical-normative case tends to be the one that people concentrate on, but we can talk in the same normative way about what a person should believe or feel. We can consider whether a person *should* feel resentment given the way he has been treated, or whether he *should* believe something in the light of the facts available to him. Here again we may be talking about what there is sufficient reason for that person to believe or to feel.

We can define a regimented sense of 'should' and 'ought', reserved for just that normative use. First, note that in the practical case, unlike the epistemic and evaluative cases, there can be sufficient reason for you to do any one of a number of strictly incompatible things. So let us say in the practical case that you have *uniquely* sufficient reason to  $\psi$  if  $\psi$ -ing is the only thing you have sufficient reason to do. What you have uniquely sufficient reason to do may thus be stateable only in a disjunction — to  $\alpha$  or  $\beta$  or  $\chi$ . In the epistemic and evaluative cases you have uniquely sufficient reason to  $\psi$  if and only if you have sufficient reason to  $\psi$ . We can now define the regimented sense of 'should' and 'ought':  $x$  should/ought to  $\psi$  if and only if there is uniquely sufficient reason for  $x$  to  $\psi$ . And we can define a regimented sense of 'right'. A right thing to do, believe or feel is a thing there is sufficient reason to believe, do or feel.

<sup>7</sup> Another example is its use as an auxiliary verb — 'If I should die, think only this of me.'

## 6. 'Good'

While 'ought', 'should', 'right' have varied uses in ordinary language, the value terms 'good' and 'bad' are more stable. And here there is a well-known reduction to reasons — what Scanlon has called the buck-passing account.<sup>8</sup> According to this, if we say that  $y$  is good we imply that there are some facts or other,  $p_i$ , which are sufficient reason for some pro-act or attitude towards  $y$ . (For 'bad' it would be a con-act or attitude.) But we do not say what those facts are. Nor do we say what the relevant pro-act or attitude is. However the buck-passer holds that a list of the pro-acts and attitudes can be provided, allowing the analysis to apply to the whole range of things that can be said to be good, including good arguments, good picnic tables, good people and good music. So 'good' is a normative predicate, as are 'right', 'should' and 'ought.'

I think that the buck-passing view is correct, but it has proved controversial and needs to be defended. There is the question of what the pro-attitudes are, how to formulate the view for all cases, including, for example, functional uses of the term 'good', and how to extend it to *better*, *worse* etc. I think these questions can be satisfactorily answered; but here I want to discuss very briefly two more general objections that have been made to the buck-passing account, since we can illuminate the Reasons Thesis by considering how to answer them.

Consider the following definition of 'good F':

$y$  is a good F in certain respects C, to degree  $d$  at  $t$  if and only if there is sufficient reason at  $t$  to pro  $y$  to degree  $d$  as an F in those respects C.<sup>9</sup>

Two objections may be raised to it.<sup>10</sup> It may be objected in the first place that the equivalence fails to hold. And secondly, even if the equivalence holds it may be objected that the definition gets things the wrong way round. If there is sufficient reason to pro  $y$  as an F, that is

<sup>8</sup> Scanlon, 1998, p. 96. Rabinowicz and Rønnow-Rasmussen, 2004, provide some earlier history.

<sup>9</sup> I discuss buck-passing more fully in Skorupski, 2007.

<sup>10</sup> See for example D'Arms and Jacobson, 2000 (especially section IV), Crisp, 2000, 2005.

because  $y$  is a good  $F$ . I *explain* why there is reason to pro  $y$  by pointing out that the reason is that it is a good  $F$ . If saying that it was good just *was* saying that there is sufficient reason to pro it, the objection runs, that would not be an explanation.

As to the first objection, ‘evil demon’ examples illustrate why the equivalence may be thought to fail. Suppose the violin performance is not good, but the evil demon will punish me with eternal torture if I fail to admire it. Is that not sufficient reason for me to admire it, even though it is not good? Similarly, suppose that some purported evidence to the effect that  $p$  is not good evidence, but the evil demon will again punish me if I fail to believe that it probabilises the conclusion that  $p$ . Is that not sufficient reason for me to believe that it does probabilise that conclusion, even though it is not good evidence?

In reply to this objection, we can apply the distinction between reasons to believe or feel on the one hand and reasons to bring it about that one believes or feels on the other. Thus: there is no sufficient reason for me to admire the performance, though there certainly is sufficient reason for me to bring it about that I admire the performance, if I can. In other words, in this case there is reason for me to bring it about that I admire something which there is no reason for me to admire. Likewise: there is no sufficient reason for me to believe that this purported evidence to the effect that  $p$  probabilises that conclusion, but there certainly is sufficient reason for me to bring it about that I believe it does, if I can. Bringing about these things is undoubtedly choice-worthy, if it is possible. So what is good in these cases, according to our definition, is the choice or policy or action of (trying to) bring it about that I admire the performance, or that I believe the evidence to be probabilising. It does not follow, on the buck-passing analysis, that the performance or the evidence itself is good.

Is this response *ad hoc*?<sup>11</sup> It does not seem to me that it is. The distinction between epistemic, evaluative and practical reasons is independently based on what kind of act — belief, feeling or action — a given reason is a reason for. So the response is just an automatic consequence of identifying what exact reason relation we are discussing. In the case of the violin performance, the fact that the evil demon has his evil plans is a sufficient reason for me to *do* something —

<sup>11</sup> As suggested in Rabinowicz and Rønnow-Rasmussen, 2004, p. 412.

namely, bring it about that I admire the performance, if I can.<sup>12</sup> In the circumstances, that would be a very good thing to bring about. Over and above that uncontroversial point, there is then the question of whether an *evaluative* reason relation also holds. Does the self-same fact about the evil demon stand in that *distinct* reason relation to me and a certain feeling of mine, namely, *admiring the performance*? The two relations are distinct, since their *relata* are distinct. And once they have been distinguished, there seems to me to be no case for holding that the second relation holds as well as the first.

This may be denied. Take the case of belief. Someone may agree that there is a type difference between epistemic reasons to believe and practical reasons to make yourself believe, or bring it about that you believe, but still insist that facts about the usefulness of believing that *p* are not just a reason to make yourself believe that *p* but *also* a reason to believe that *p*. Against this one can object that there is nothing self-contradictory in saying that a person has reason to *make* himself believe something, for example, that he will survive the dangerous mission, even though he has no reason to believe it. The objection seems to me decisive, but suppose that it too is denied. The denier would have to have, it seem to me, an at least partly pragmatic theory of truth. A pragmatist about truth can hold that the usefulness of believing a proposition is a priori indicative of the truth of that proposition. For him, therefore, the usefulness of believing that *p* is unproblematically both a reason to believe that *p* and a reason to believe that it is true that *p*. If, in contrast, one denies that the usefulness of believing a proposition is a priori indicative of its truth, then how can this usefulness constitute, in and of itself, a reason to believe that *p*?

Let us turn to the second objection. Does our definition get things the wrong way round? I can *explain* why there is reason to favour something by pointing out that it is good. I do not *explain* why it is good by pointing out that there is reason to pro it.

This is a weak objection. In the first place, there may be reason to favour it in one way because there is reason to favour it in some other way. Thus, the reason to *choose* this CD performance may be that it is the best performance — the one there is most reason to *admire*. We might have reasons to choose this particular CD for reasons other

<sup>12</sup> Rabinowicz and Rønnow-Rasmussen agree.

than the goodness of the performance: because it is cheap for example. If we are working to a tight budget the cheapest CD performance may be the best choice, even though the performance is not the best. So it has explanatory force to say that there is reason to choose it because it is the best performance.<sup>13</sup>

Further, how do we explain why something is a good F? We do so by pointing out the facts about it that make it good as an F. Those are the very facts that give one reason to favour it as an F. In that sense we do explain why it is a good F by stating the reasons for favouring it as an F. We do not of course explain why a thing is good by simply saying *that* there are reasons to pro it, but we do explain by saying what those reasons are.

My conclusion is that if one clearly distinguishes and acknowledges the three basic kinds of reasons — epistemic, practical and evaluative — and looks in detail at the structure of normative explanations, or justifications, what emerges is a perfectly sound defence of the buck-passing view.

## 7. Thick evaluative terms

We have now considered the thin or topic-neutral normative terms that make up the central circle of normative vocabulary. We also have to consider the difficult and ramified outer circle. Significantly, most terms in this outer circle are in fact evaluative, having to do with reasons for feeling this or that — I will call them ‘thick evaluative terms.’

When we say that something is good or right, we say there are reasons without saying what the facts are that constitute those reasons, or for what act or attitude they provide a reason. Thick evaluative terms effectively impose tighter restrictions. How do they do that?

One way they might do so is by being more specific about what attitude there is reason for. Consider in particular predicates of the form

$$A: (\exists p_i)S(p_i, \phi(y)).$$

<sup>13</sup> You might have reason to choose a CD which has a bad performance just to illustrate in your consumer survey what poor value there is on the market. In this case the bad performance is a good choice.

where ' $\phi$ ' names some specific affective reaction. What they say about  $y$  is that there is sufficient reason to take attitude  $\phi$  to  $y$ .<sup>14</sup> They say no more than that. However, that may tell us quite a lot. From our inner understanding of that particular reaction — admiration, say, or disgust — we can work out quite a bit, indirectly, about what sort of facts those facts must be, even though we are not told what specific facts obtain.

The number of potential practical normative or epistemic normative predicates is as large as the number of action and belief contents, i.e. indefinitely large. In the evaluative case the affective *attitudes*, as well as their contents, ramify indefinitely. Furthermore, and very importantly, evaluations of type A will not all be neatly classifiable as asserting that there is sufficient reason for a *pro* or *con* attitude toward the object, and thus, as implying that the object is good or bad in some respect. Does the fact that there is reason to be disturbed or shocked by a performance, for example, make it good or bad? That may well depend on context. Disgust is generally a con-attitude. But even this need not always be evident. An object may give sufficient reason to respond with a complex of attitudes of which disgust may be an element. Yet the overall complex may justify a pro-attitude. This openness between making a specific evaluation in terms of an affective attitude and determining whether that evaluation counts towards the goodness or badness of the object is one of the things that gives critical discourse in the arts its depth, complexity and elusiveness. Specific evaluations can be indefinitely subtle, quite informative (though inexplicitly) about the facts, and linked in no neat or simple way to an assessment of their objects as good or bad.

Does ordinary language contain any terms analysable as type-A predicates? One set of candidates for the role consists of terms like 'admirable,' 'lovable,' 'enviable,' 'desirable' and so on. Arguably, 'This is admirable' means 'there is sufficient reason to admire this.' If that is right, ' $\psi$  is admirable' is a purely normative predicate.

It may be replied that predicates of this kind — ' $\phi$ -able' — can have the non-normative meaning 'capable of inspiring admiration/desire/love.' Maybe they are useable in either or both ways, the

<sup>14</sup> This is the simplest case. Where the reason is relative to an agent the agent variable can come into play: there is reason for me but not for you to desire  $y$ . Where it is temporally relative the time variable will come into play: there was reason then, but there is no reason now, to hope that  $y$  would come.

normative and the non-normative. But we need not accept this. The appearance that such predicates have non-normative use may be explained away, by reference to their epistemology and their function. The fact that something does inspire admiration, desire, love or whatever (lastingly, on reflection, after discussion) is the fundamental epistemic criterion for its being worthy of admiration, desire, love etc. Now just because the fact that  $p$  is a criterion of, that is, constitutes a defeasible warrant for, the assertion that  $q$ , it does not follow that in asserting that  $q$  we are asserting that  $p$ . Nonetheless, when a proposition is very tightly tied to a dominant criterion, as in the case of these ' $\phi$ -able' predicates, it is easy to think (though strictly it is incorrect) that asserting the proposition is asserting that the criterion obtains. Furthermore, given the use of these predicates in criticism and persuasion, it is very natural, if one is challenged as to whether a performance is really admirable or a house is really desirable, to fall back on the fact that many people do admire it or desire it. But neither point shows that 'admirable' or 'desirable' have a non-normative ingredient in their meaning. So the view that these are purely normative predicates can be defended.

At first sight terms such as 'frightening', 'moving', 'boring' and so on seem more clearly to fall on both sides of the normative/non-normative border line. Again, however, one can argue by reference to their epistemology and their function that the appearance of non-normative content is misleading. Take 'frightening'. Can it *mean* 'fear-causing' as well as 'fear-worthy'? It is not easy in fact to come up with an example of a use that is clearly non-normative. Consider 'These steps would be very frightening to someone with poor balance and short sight.' 'Frightening' here could mean 'fear-causing', but it could also mean 'fear-worthy'. Or again: 'Spiders are frightening to many people but there's really nothing about them to be frightened of.' Is the first occurrence of 'frightening' in this sentence a non-normative use? Not necessarily. The sentence could mean 'Spiders seem to many people to be fear-worthy but they are not really.'

In the same vein, suppose I attend a lecture on quantum physics and find it boring. Note the typical locution: *finding* something boring, amusing, irritating etc. This looks as though it might be factive, but it is not: from the fact that I found something boring it does not follow that it *is* boring. Rather, to say I found it boring is equivalent to saying that it seemed boring to me. Someone may reply 'It was not *really* boring — it is just that you do not know anything about the

subject. The experts who were there found it really interesting.' Likewise, if one says that something was genuinely moving, or really or truly moving, one makes a contrast between what merely seems moving and what really is. How does the contrast between appearance and reality work here? It is not that something can seem to move me without really moving me. Rather, the contrast is between apparent and real normativity. When one is moved by something, a musical performance say, one doesn't simply experience an emotion, one experiences it as appropriate, fitting to its object. But it can happen that this normative dimension in one's feelings strikes one as misguided. I am moved by the performance, and that emotional reaction comes packaged, so to speak, with an impression of itself as reasonable — but at a level more detached from the reaction itself I think to myself that I am being sentimental. The performance was not genuinely, really, moving; I am just a sucker for that sugary kind of violin sound. So 'genuinely moving', 'truly moving' are normative.

So far, so good. Yet many terms used in ordinary-language evaluative discourse may seem normative even though they carry no connection to an affective reaction on their face. In aesthetic evaluation they abound: 'bizarre', 'beautiful', 'sublime'. Let us examine a few cases.

Since our linguistic taxonomy of affective responses is less ramified than our taxonomy of evaluations one should expect difficulty in finding a word to characterise, for each evaluative predicate, the *specific* affective response that its application says there is sufficient reason for. Take the example of 'bizarre.' It has reference to some form of surprise. To say that something is bizarre is to say, according to the *Oxford English Dictionary*, that it is 'extravagant, whimsical, strange, odd, fantastic.' But also: 'At variance with the standard of ideal beauty or regular form; grotesque, irregular.'<sup>15</sup> In each case, the surprise is justified by the fact that the item in question goes beyond what there is reason to expect from a 'regular' item of that kind. The whimsical, fantastic, etc. is surprising in *that* way; it gives reason for an attitude of surprise directed to *that* content. What content? Something that is bizarre. So it seems that we have the same circularity here as we found with 'morally wrong'.

<sup>15</sup> The OED also has: 'At variance with recognized ideas of taste, departing from ordinary style or usage; eccentric ...' But recognized ideas of taste, and ordinary style and usage, may themselves be bizarre, in which case the eccentric may precisely not be bizarre. None of these, of course, are strictly *definitions*.

With the beautiful and the sublime we arrive at terms which have long been recognised as difficult to define. Many of the traditional discussions in aesthetics have in fact been concerned with what *makes* a thing beautiful (in the sense in which one might discuss what makes a thing annoying.) It seems uncontroversial that the beautiful is that which is admirable in a certain way. But what way? By being delightful, charming, graceful, by giving a certain sort of pleasure, in virtue of a certain sort of rightness of combination, of proportions, colour-combinations or whatever. So: if  $y$  is beautiful then  $y$  has properties combined in such a way as to give satisfaction or pleasure of the kind taken in this perceived *rightness* of combination, and in virtue of that pleasing combination, to give sufficient reason for admiration. Thus, the concept of the beautiful involves a double layer of normativity. (The concept of the sublime seems simpler in its logical structure: it is that which is uplifting, exalting, ennobling, awe-inspiring.) The properties of a beautiful thing have a rightness of combination, they combine as they should. And it is this rightness of combination that gives sufficient reason for delight and admiration. This account of the notion also highlights one reason why the concept of the beautiful remains a problem for aesthetic theory. What is it for properties to combine as they should, or as there is reason for them to combine? What can it mean to talk about whether properties are doing what they should — how can the normative concept get an application here? There seems to be a simulated functional assessment involved in a judgement of beauty, captured in Kant's formula of 'purposiveness without a purpose.' We can say that an object with a function is working as it should, arranged as it should be, by reference to its function. But with beauty we have the idea of rightness of arrangement without a function to underpin it.

## 8. Mixed predicates

There might be terms that predicate both a non-normative property and a reason constituted by possession of that property. Where 'F' is the non-normative property in question, such terms will have the form

B: S(the fact that  $Fy$ ,  $\phi(y)$ )

Call this a mixed predicate. Are there any such predicates?

We have just suggested that various predicates which may look as though they have this form don't really. Term like 'admirable' and 'boring' are purely normative. It's true that the fact that something is boredom-causing is, epistemically speaking, a reason to believe it is boring, but that is a different point.

Then again, some predicates which may be thought to be mixed are probably not normative at all. Consider 'That was a cruel thing to do'. This means (at least) 'That was an action which showed in its motivation the presence of a non-instrumental desire to inflict suffering'.<sup>16</sup> So the assertion makes a factual claim. But does it also, as part of its meaning, make a normative claim? That would make it a mixed predicate: e.g. 'y shows in its motivation the presence of a desire to inflict suffering, and that fact is a sufficient reason to disapprove of y'.

It is not clear to me that this is the right account. Is it self-contradictory to deny that there is reason to disapprove of cruelty? I think not: it is just false. Likewise with kindness, courage, resentfulness, laziness, arrogance, coldness and so on. These are names of character traits which we take to be virtues or vices, excellences or defects, but it does not seem to me to be part of their meaning that that is what they are. Similarly with such terms as 'adultery': there seems to be nothing self-contradictory in saying that there is nothing wrong with adultery. (One might avoid that word, meaning to avoid unintended conversational implications, or one might choose it precisely to rid the word of those implications.)

But now consider 'murder'. To call something a murder is at least to say that it's an intentional killing with 'malice aforethought'. Is it a definitional truth that there is reason to condemn murder? If it is, that's because *malicious* killing means 'killing there is reason to condemn.' Not every intentional killing is malicious, however, and it does not seem possible to narrow down the intentional killings that are malicious by a wholly non-normative criterion. A murder in an intentional killing where the intention is of such a kind that there is sufficient reason to blame the person who does the act with that intention.<sup>17</sup> This gives us as a descriptive as well as a normative com-

<sup>16</sup> Cruelty is more than callousness, or mere lack of care. 'It wasn't just thoughtless, brutal, stupid, reckless etc. — it was cruel.'

<sup>17</sup> There can of course be further distinctions, say between murder and culpable homicide, or murders of various degree.

ponent. Importantly, however, it does not give us a wholly *non-normative* characterization of the fact on which the blameworthiness supervenes, so it does not fit form B. It has this structure:

$x$  is an intentional killing of such a kind that there is sufficient reason to blame the person who does  $x$ .

Predicates with this structure have a non-normative component, but one that is not strong enough on its own to characterise the reason purely descriptively. Consider courtesy and discourtesy. Courtesy is a matter of paying due care to putting people at their ease in a social context where that is relevant. Discourtesy is giving insufficient attention to that. 'Insufficient' is normative: the attention (if any) the action showed to placing others at their ease was small enough to constitute sufficient reason for disapproval of that aspect of the action. So the predicate ' $x$  is discourteous' has a reason predicate as a definitional consequence — 'there is sufficient reason to disapprove of  $x$ '. But it is not possible to factor out a wholly non-normative component that constitutes the sufficient reason for disapproval. We are assessing the action by reference to how much attention it showed, in its context, to putting others at their ease, and saying that it did not show enough. Saying that is assessing the action normatively. Equally if we say the action was courteous, we are assessing it in the same dimension and saying that it showed a degree of attention to putting people at their ease great enough to give sufficient reason for approval.

So what we are saying, in these two examples, is of the form

C:  $y$  has property F to a degree or in a way which is such as to give sufficient reason to  $\phi(y)$ .

Various other terms work like that: 'careless', 'reckless', 'negligent', 'inattentive', 'overbearing'. We characterise the degree of care, attentiveness, etc. by saying that it is such as to satisfy a normatively expressed condition. So we cannot factor out a non-normative sentence which states the fact that constitutes the sufficient reason.

I therefore agree with people who hold that there are thick evaluative terms from which a reason-giving factual component is inextricable. We may not have a way of exactly characterising just those facts that merit a particular attitude except by reference to the attitude in question. And we may not have neat vocabulary to characterise

exactly the attitude in question. However these points do not undermine the two truths which I said lay behind the *is/ought* distinction.

## 9. Normative content and the reason predicates

If the semantic thesis were correct, we could draw the conclusion that anything normative our actual vocabulary allows us to say could be said while using only descriptive predicates and reason predicates. All normative content could be expressed that way. Other normative predicates would be expressively, if not practically, redundant.

Does this conclusion still apply if we have to supplement the semantic thesis with the concept-possession thesis, as with the predicates ‘morally wrong’ or ‘bizarre’? Take ‘morally wrong.’ Would it be possible, without expressive restriction, not to have the term ‘morally wrong’ but only a word for a certain affective response to actions, plus a reason predicate that allowed one to distinguish between actions that provide sufficient reason for that response (‘merit’ it) and actions which do not? More generally, could terms that are not normative by the semantic thesis, but nevertheless express normative concepts — as is the case with ‘morally wrong,’ and arguably, ‘bizarre’ — be replaced without loss of expressive power by explicitly normative predicates? If so, then the conclusion would still apply.

Consider again the analogy with colour words. Would it be possible, without expressive restriction, to have no colour words but only colour-sensation words, and to talk only about the powers of objects to cause those sensations, and the grounds of those powers? This question should be distinguished from a question about the order of language-learning, or of concept-acquisition. Plausibly, the concept of a colour-sensation is further on in that order than the concept of a colour, and it might even be argued that it cannot be grasped except through a prior grip on the concept of a colour. This might explain the semantic priority of colour words in our actual language. However, this impossibility would not show that a language which had colour-sensation words but no colour words would be *expressively* restricted in comparison to our actual language. To assess the issue, we would have to consider more carefully the notion of ‘saying the same thing,’ ‘expressing the same content’.

The same goes for the moral concepts. Perhaps they lie earlier in the order of learning than the concept of a blame-sentiment, and

perhaps that explains their semantic priority in relation to terms denoting the sentiment. But again that does not show that a language which included the reason predicates, and terms for the sentiments or attitudes, including the blame sentiment, would be expressively restricted in comparison to the one we have.

This has been an illustrative rather than a comprehensive survey, but we have turned up no good objection to the Reasons Thesis — the thesis that all normative concepts are reducible to the concept of a reason. The question on which I have ended is whether a language which added only the reason predicates to its descriptive vocabulary, and contained a sufficiently ramified and nuanced vocabulary of sentiments, would be able to express any normative contents that we can express. I have found no decisive objection to that idea, and some reason to consider it plausible. Furthermore, taking the concept of a reason as basic seems to me to fit in with other persuasive philosophical considerations. All in all, it seems to me that the case for thinking the normative domain is the domain of propositions about reasons is pretty strong.<sup>18</sup>

John Skorupski  
 Department of Moral Philosophy  
 Edgecliffe, The Scores  
 The University of St Andrews  
 St Andrews, KY16 9AL  
 United Kingdom  
 jms2@st-andrews.ac.uk

### References

- Crisp, Roger. 2000. Review of *Value ... And What Follows*, by Joel Kupperman, *Philosophy*, 75.
- Crisp, Roger. 2005. Value, Reasons, and the Structure of Justification: How to Avoid Passing the Buck. *Analysis* 65.
- D'Arms, Justin and Jacobson, Daniel. 2000. Sentiment and Value. *Ethics* 110.
- Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Oxford: Oxford University Press.
- Rabinowicz, Wlodek and Rønnow-Rasmussen, Toni. 2004. The Strike of the Demon: On Fitting Pro-attitudes and Value. *Ethics* 114.

<sup>18</sup> I am grateful to Jonas Olson, Wlodek Rabinowicz and Toni Rønnow-Rasmussen for their helpful comments.

- Scanlon, T. M. 1998. *What we Owe to Each Other*. London: Harvard University Press.
- Skorupski, John. 2005. Blame, Respect and Recognition: A reply to Theo van Willigenburg. *Utilitas* 17.
- Skorupski, John. 2006. Propositions about Reasons. *The European Journal of Philosophy* 14.
- Skorupski, John. 2007. Buckpassing About Goodness. In *Hommage à Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz*. Ed. T. Rønnow-Rasmussen, B. Petersson, J. Josefsson & D. Egonsson, 2007, where I also give definitions of good for *x*, good of *x*, and functional goodness. (<http://www.fil.lu.se/hommageawlodek/site/papper/SkorupskiJohn.pdf>).