Normativity

Edited by

JONATHAN DANCY

Normativity

Normativity

Edited by

Jonathan Dancy

Normativity

Normativity

Edited by

Jonathan Dancy

First published 2000

ISBN 0–631–22041–0 (Pbk)

Blackwell Publishers
108 Cowley Road
Oxford OX4 1JF, UK

Blackwell Publishers Inc
350 Main Street
Malden, Massachusetts 02148, USA

British Library Cataloguing in Publication Data has been applied for

Library of Congress Cataloging in Publication Data has been applied for

Typeset by Cambrian Typesetters, Frimley, Surrey
Printed in Great Britain
By MPG Books, Bodmin, Cornwall

This book is printed on acid-free paper

Contents

Editor's Introduction

The 1998 RATIO conference was on normativity, and this volume is the ostensible result. The speakers at the conference were Peter Railton, Derek Parfit, and John Skorupski, who at the last minute kindly took the place of Joseph Raz, whom illness had prevented from giving his paper. The paper that Derek Parfit gave at the conference has sadly ceased to exist in any recognisable form, and so cannot be included here. But the collection is enhanced by the contributions of Frank Jackson, John Broome and Christopher Hookway, and I am grateful to them for allowing their work to appear in this context.

My introduction starts with some general remarks about normativity, and then turns to discuss a theme that links several of the essays in this volume.

I

It is often said that normativity is the characteristic common to everything that appears on the "ought" side of the distinction between what is and what ought to be. This is true, however, only if our "ought" here is not particularly a moral ought, nor even just a practical ought; and true only if we include what is good and bad under the general heading of what ought to be or not to be. For the notion of value (good and bad) is held to be as normative as the notion of the right; in other terms, the class of the normative has two distinct sub-classes, the evaluative (good, bad, etc.) and the deontic (right, wrong, duty, obligation, permission – and perhaps also "ought"). Perhaps, then, it would be better not to have any single term like "ought" as the mark of the normative, and to say merely that normativity is a feature common to both sides of the evaluative/deontic distinction. I confess, however, that I find it helpful to keep an "ought" in mind when thinking about normativity.

We find such "oughts" in ethics, of course, and in practical deliberation (you ought to turn left now); and we find them in

what is called "theoretical reasoning", that is, in the attempt to use available evidence to determine what is the case (either in doing science or just in investigating the everyday nature of our surroundings). We can say "you ought to have realised that it was going to rain, because the clouds were so very black", and "you ought not to deny that *q* while you think that *p* and that *p implies q*".

It is common to think that these sorts of oughts can all be understood in terms of one basic notion, that of a reason. If you ought (morally or not) to do it, this can mean no more than that there is good reason for you to do it and inadequate reason against. On the theoretical side, to say that you ought to have realised that it was going to rain is to say that there were strong and evident reasons for thinking so. Now nobody denies that the notion of a reason is central to that of normativity. The question is whether we need more than that – whether the whole story of the normative can be told in terms of this one notion. We can address this question by considering a common way of thinking about reasons. When we say that her need is a reason for us to help her, this is to be understood as saying that her need *favours* our helping her; it speaks in favour of our acting in that way. A reason for doing an action is a feature that stands in the "favouring" relation to that action. As John Broome puts it (p. 80), it "reasons" the action. Our question is whether this favouring relation is the common core of the normative. There are at least two reasons for thinking that it is not.

The first of these is that we sometimes think of action, not so much as favoured by the situation, but as *demanded* by it, and this "demanding" relation does not seem like a specially strong version of the "favouring" relation we started out with. Favouring comes in degrees of greater and less strength. We might have two alternatives to consider, one more favoured than the other. But this means that if the other were to become more favoured, it would eventually become the one that we ought to adopt. If an action is *demanded* by the situation, however, things are different. "Demanded" does not just mean "most/more favoured". (For further reasons, see Broome's distinction between strict and slack at pp. 80–3.)

One might allow this point as a necessary correction of an overnarrow conception of reasons as capable only of favouring. Some reasons favour, we could say, and others demand. When Joseph Raz writes "The normativity of all that is normative consists in the

way it is . . . related to reasons" (p. 34), I suspect that he has such a more flexible conception of reasons in mind. All that we have learnt so far, on this showing, is that our initial characterisation of a reason as something that stands in the favouring relation to something else was too limited. But there is reason to think that we need more than these two normative relations. Favouring is a relation between something that favours and something else that is favoured (between the reasons or ground, on the left, and the thing "reasoned", on the right). When an action is made right by the features that favour doing it, those features can be understood as the ground for the rightness of the action; the same is true when an action is made sensible or rational, and when a belief is made justified by the evidence that speaks in favour of it. The "demanding" relation is also a relation between ground and action (or belief). Our question then is whether all normativity is to be thought of as *relational*, and if relational, relational in just this sort of way. Most radically, we might need something that is not a *relation* at all. Is there, we might ask, something that is demanded, but not demanded *by* anything? What, for instance, demands that you do not believe that you do not exist?

We have, then, two ways in which we might want to move beyond the favouring relation. We need a rather different relation, one of demanding, I think, and we *might* need something that is not a relation at all. The latter suggestion does not, I admit, have many supporters. There is much stronger reason to think that we might need something that is relational but different from the relations we have so far. This is what Broome argues in his contribution here. He focuses attention on what he calls "normative requirements".[1] The difference between this relation and the ones we already have in place is that they offer what is called "detachability"; it does not. What this means is that if we know that her need favours (or even demands) your helping her, and that she is in need, we know that you have a reason (or even ought) to help her. Where the left hand side of the relation is satisfied, normativity passes to the right hand side. We can see this most obviously in terms of the "formal consequences" that Broome ascribes to each relation. Where *p* favours *q*, the formal

[1] The terms "require" and "demand" are not really distinct in ordinary usage. I have chosen to use the term "demand" to characterise the stronger relation that is of the same general type as that of favouring, and to accept Broome's term "requirement" for the different, though still relational, type of normativity that he discusses in his paper.

consequence of this instance of the favouring relation is $p \rightarrow Rq$, and this reveals that where it is the case that p we can infer that Rq. Normativity has passed to the right hand side on its own (it is "detached"). With Broome's new normative relation of requirement, this does not happen. One thing can normatively require another without this meaning that where we have the first, the second is normatively demanded (by *anything*).

This new notion of a normative requirement is worth careful consideration, so that we can be sure that it is genuinely something beyond what we had already. Matters are not entirely helped by the formulations that Broome offers of what he calls the "logical factors" or formal consequences of the various relations we have been considering. In the case of the demanding relation he writes $p \rightarrow Oq$, with $p \rightarrow Rq$ as the logical factor for the favouring relation. For normative requirement he offers $O(p \rightarrow q)$. And he remarks that with normative requirement, "the normativity is attached to the relation", whereas with the demanding and favouring relations "the normativity is attached to the consequent" (p. 82). This contrast seems exaggerated to me. In the case of the demanding and favouring relations, it is certainly true that normativity passes to the consequent; but there is nonetheless a normative *relation* there as well; in the case of demanding it is the right-making relation; in the case of favouring it is the reason-giving relation. These relations are there and they are normative in their own nature. The difference between them and a normative requirement is that the latter never gives its consequent a normative status of its own.

Why then should we not understand a normative requirement simply as a complex that is demanded? This would be in line with Broome's account of the "logical factor" of normative requirement, which uses the same operator O that we see in the logical factor of demanding, but has a more complex right hand side and fails to specify the relevant ground. Suppose that we reject the suggestion that there is such a thing as groundless normativity. There must be a ground, then, even for Broome's normative requirements. Suppose, for instance, that one example of a normative requirement is that one is required not to believe that others would be wrong to do this while blithely doing it oneself. The ground for this might be that there is no relevant difference between oneself and others. Call that ground r. This gives us, as logical factor of the whole situation, $r \rightarrow O(p \rightarrow q)$. This has the form of a demand, but with a more complex right hand side. Of

course Broome's point is that even given *r* and *p*, we cannot detach *Oq*. But this can remain true without generating another *form* of normativity. We don't need a new form of normativity to stop $O(p \rightarrow q)$ & *p* from entailing *Oq*.

Broome does say some other things that we should consider here. Take a normative requirement whose logical factor is $O(p \rightarrow q)$; the difference between the normative requirement itself and the logical factor is that the latter does not tell you that it is *p* that requires you to *q* (as he puts it on p. 82). So the normative requirement specifies one thing that requires another. But, and this is the crucial point, it does not require it in the sort of way that a ground demands (or favours) something. The requiring is done by something that is not a ground. This does not mean that requirements are all groundless, exactly; there may be a ground for the normative requirement, as we saw above. But it means that we have to try to understand a sort of requiring that is different from anything that a ground can do. This is the crucial point. To attempt an example: believing that she is in trouble and needs help may (on occasion) normatively require that you help, without making it the case that you ought to help. Perhaps she is not in trouble and does not need help; if so, it is not the case that you ought to help. But even so, Broome would say, if you believe that she is in trouble and do not help, you are not as you ought to be.

However we are to understand this notion of a non-grounding requirer, Broome's discussion does appear to reveal a difficulty with Frank Jackson's paper. Jackson argues that non-cognitivists cannot give any account of what he thinks of as an obvious truth, namely that someone who believes that *p*, and that *if p then q*, ought to believe that *q*. If Broome is right, this is not an obvious truth at all, but in fact a falsehood (or at least an invalid inference). For if you ought not to believe that *p*, the fact that you do believe this and also believe that *if p then q* does nothing to show that you ought to believe that *q*. Analogously, in the practical sphere, if you ought not to have promised to do it, the fact that you did promise may do nothing to show that you ought to do it. What Jackson has in mind is really a normative requirement, in the terms we have been using, not a demand at all. Once we make this change, the obvious truth which the non-cognitivist cannot capture would now be that you ought not to deny that *q* while believing that *p* and that *p implies q*. Most, at least, of Jackson's discussion can be recast in this way without loss.

II

I now turn to a quite different issue that surfaces in several of the contributions to this collection. John Skorupski argues that normative claims may be true or false without there being any metaphysically robust (as he calls them, "worldly") normative facts. This is what he calls "irrealist cognitivism". We can contrast it with the sort of cognitivism espoused by Frank Jackson (p. 113), which is a kind of reductive naturalist realism. Jackson argues for a realist understanding of normative constraints on belief, though he does not take his arguments on this point to be any direct support for his favoured naturalist version of realism. He clearly has not considered the possibility of the view that Skorupski is offering, which allows truths about how we should believe and act without there being any "substantial" fact of the matter. As Jackson presents the matter, if you have truths you have facts, and substantial facts at that. "Truth-apt" sentences are those suited for the conveyance of information (p. 107), and information (at least when true) consists in substantial facts – worldly facts, as Skorupski would put it.

Now one thing that is driving Skorupski here is expressed in terms of a distinction between receptivity and spontaneity. As he puts it "the very idea of a worldly fact is the idea of a fact that can only be known to obtain by *receptive* awareness . . . Irrealism about the normative says that normative knowledge of fundamental normative propositions *rests on no receptive awareness.* The only capacity it requires is the non-receptive cognitive capacity of rationality, a capacity which involves spontaneity and regulation by the universality of reasons, not receptivity" (p. 136).

The matter I want to pursue is the nature of this distinction between receptivity and spontaneity and its relevance to the study of normativity. Earlier on in his paper, Skorupski says that the receptivity/spontaneity distinction is the same as the distinction between descriptive and normative judgement (p. 117). If that were true, it would immediately follow that fundamental normative facts (those with no admixture of the descriptive) cannot be known in any "receptive" way. It would presumably also follow that the "purely descriptive" cannot be grasped in any way that involves spontaneity.

This matter is of interest in its own right, but especially here because what Skorupski claims here may be at odds with some things said by other contributors, namely Joseph Raz,

Christopher Hookway and Peter Railton. My purpose here is to try to work out whether this is so or not, in a way that might help readers to sort out a pretty tangled issue.

Raz and Hookway present a fairly united front. They agree that our response to reasons is at least partly active even if also partly automatic. Raz, for instance, claims that there is no conflict between (certain aspects of) the exercise of our rational capacities being automatic and its being under our control. He writes, "Control is manifested when a belief is adopted, or endorsed in a process in which the ability to recognize reasons and respond to them . . . is active. It can be active even when beliefs are formed without deliberation or awareness, but where the agent's critical faculties would have stopped their formation had they been rationally suspect" (p. 44). He continues, "People's beliefs . . . are rational only if they are formed and maintained while the people involved are in control of their formation and continued endorsement" (p. 45). Hookway seems to agree. He writes, "perhaps through a process of habituation, we acquire abilities to make these evaluations in a largely automatic and unreflective way; making them becomes second nature" (p. 64). He goes on: "it seems problematic that we have no control over these habits and dispositions. . . . Questions 'occur to him'; suggestions 'pour in on him': such idioms indicate that we are the passive recipients of these questions and suggestions" (p. 67). But, he claims, "This passive influx is a kind of sensitivity to the normative demands of reason and . . . an expression of our freedom of mind" (*ibid.*). It is an expression of freedom of mind (and hence, presumably, of spontaneity rather than of pure receptivity) because our acceptance of a suggestion is not merely passive; it signals our "active endorsement" of the suggestion (p. 72), and is a "manifestation of our values" (p. 71). For Hookway, in our acceptance of many beliefs about our surroundings, descriptive though they are, there is an "interplay" of passivity and active endorsement.

The main point to be extracted from all these quotations is that even in ordinary descriptive judgement, we are as much active as passive. In fact, both Raz and Hookway really want to say (I think) that what looks like passivity is in fact merely the automatic functioning of active habits and dispositions, in which we at no time lose our status as agents in favour of a status as mere recipients of information. In "receiving" information from our surroundings we are always assessing and endorsing reasons, coming to some view about the interplay between different reasons present in the

case, and thus eventually forming an opinion about how things add up.

The question then is whether this is at odds with the picture to which Skorupski is appealing. We might allow that, if Raz and Hookway are right, there is no such thing as pure receptivity. If so, and if, as Skorupski says, the notion of receptivity is the same as that of descriptive judgement, there is no such thing as purely descriptive judgement. This conclusion seems both improbable in itself and unlikely to follow from a correct account of receptivity. I conclude that the identification of descriptive judgement with receptivity is a mistake. What is important to Skorupski's argument, however, is not this, but that there should be a notion of pure normative judgement, with no tincture of receptivity at all. If there is such a thing, it would immediately follow that fundamental normative facts (those with no admixture of the descriptive) cannot be known in any "receptive" way. But it may well seem to us that the entire spontaneity/receptivity distinction has to be rethought now. The idea that there can be pure spontaneity even though (if Raz and Hookway are right) there cannot be pure receptivity should itself seem dubious. Our reasons, after all, are not chosen so much as given, even though their status as reasons is something that we are active in assessing. If so, the supposedly fundamental normative judgements cannot be the product of pure activity, since even these must rest on something if we are to have any reason to believe them true rather than false. But if they rest on something which is *presented* to us for assessment of its relevance as a reason, then even here there must be receptivity as well as spontaneity.

I present this as a sort of temporary conclusion, one intended to challenge Skorupski to say more about how he intends to run the contrast between spontaneity and receptivity. There are occasions, after all, where he seems to accept the sort of point I have just been making. He says, for instance, that rationality is the capacity to assess reasons of any kind, and that it "can be contrasted with *receptivity*, the capacity to receive information". He adds that "[rationality] may be diminished by . . . defects in what I will call *spontaneity*, that is, in one's capacity for spontaneously appropriate normative responses" (p. 119). It remains, then, an open question where exactly he stands on this issue.

I turn finally to Peter Railton's contribution. Railton's main theme is the way in which the normative combines two elements that are apparently in tension with each other: force and freedom.

How does Railton's contrast between force and freedom relate to Skorupski's distinction between receptivity and spontaneity, and to Hookway's contrast between passivity and activity? Railton explains his contrast on pp. 3–5. Despite their authority, he says, norms have no "coercive power"; the "must" or "ought" they involve is clearly resistible, and in that sense the normative domain is a domain of freedom. However it is not pure freedom, if by that we mean a matter of pure willing. For many of the attitudes that come under normative constraint (belief, desire, admiration, approval and so on) "appear not to be wholly within the scope of direct willing" (p. 5). There may be areas where the will is "directly" effective, where things happen just because we will that they should (one's arm's rising, for instance). But even in these cases the will is guided by something other than itself. An act of will of this sort is not a *fiat*, not a mere "let it be so"; it is a *choice*, guided by reasons (*ibid.*).

These remarks give us to understand that although there is a distinction between force and freedom there is no such thing as pure freedom, and no such thing as irresistible force. The sort of force we are dealing with requires the consent of the will if it is to be effective, and the sort of consent we give must itself be a response to something that guides, even if it does not control. This might remind us of the contention of Raz and Hookway that there is no pure receptivity; all receptivity involves an element of spontaneity, since there is always the element of active endorsement. Does this mean that Railton's contrast is the same as theirs? I am inclined to think that the answer to this question is yes. The space left by the fact that normative force is resistible is there to be occupied by the sort of active endorsement that Hookway and Raz are thinking about. And the fact that we are not here dealing with "pure" agency is capturable by the thought that the reasons to which we respond are guides for judgement, guides whose status as such is not the product of our will even though they can act as guides only if we consent and endorse them.

If this is right, we have here a theme that permeates the contributions of Railton, Raz, Hookway and Skorupski, in a way that I certainly did not particularly expect when inviting contributions to this collection.

Jonathan Dancy

1

NORMATIVE FORCE AND NORMATIVE FREEDOM: HUME AND KANT, BUT NOT HUME *VERSUS* KANT

Peter Railton

Introduction

'Normativity' is, for better or worse, the chief term we philosophers seem to have settled upon for discussing some central but deeply puzzling phenomena of human life. We use it to mark a distinction, not between the good and the bad (or between the right and the wrong, the correct and the incorrect), but rather between the good-or-bad (or right-or-wrong, . . .), on the one hand, and the actual, possible, or usual, on the other. Ethics, aesthetics, epistemology, rationality, semantics – all these areas of philosophical inquiry draw us into a discussion of normativity. And they do so not because we philosophers import this notion into our inquiries, but because – sometimes rather belatedly – we discover it there whether we went looking for it or not.

I said 'for better or worse' because, while it is useful to bring these various normative phenomena together, the term 'normativity' itself bears the stamp of but one aspect of such phenomena: *norms* – rules or standards. The etymology of the English term *norm* traces it back to the Latin *norma*, a builder's square. The term *rule* also seems to come to us from the building trade – it descends from the Latin *regulus*, a straight-edge or ruler. Now anyone who has sawn a board or chiseled a stone recognizes what it is to take a square or a ruler as a guide in cutting, and thus to treat gaps between the actual cut and the square or ruler to show there is something to be "corrected" in the cut rather than the tool. So we have here a seemingly concrete example of "action-guidingness" and an associated "standard of correctness", different from the merely actual, at work.[1]

[1] Moreover, we have an equally concrete way of illustrating part of what Kant had in mind in insisting that the normative is *a priori*. A *norma* (or *regulus*) has its form "before the fact", giving the builder a "standard of correctness" for the cut, but not staking a claim

Because the *norma* (or *regulus*) is a tool whose application is so transparent to us, it can prove a useful example. But there is a danger as well as an aptness in using such a model when we attempt to construct a philosophical account of normativity. A builder can consult his *norma* to guide himself in making cuts and to judge whether his work "measures up", but does this tool, or any tool, tell him why or when his cuts should measure up to the *norma*? In most cases it is of course evident why they should, and there certainly is no mystery why the builder's square is ubiquitious in the building trade. But what if an arch is needed, or a compound curve – is it still the case that cuts are always *to be made* following the *norma*?

Understanding how a *norma* or a norm could possess legitimate regulative standing thus also requires us to ask: What is it in general for a rule or standard to *apply*? There is no special difficulty about saying what it is for a rule to apply in (what we might call) a "formal" sense. A *norma* can be applied to a cut and we can find the cut to fit or not. But in this sense the *norma* applies even when we needed to cut a curve. So when do we say a rule *applies* or is *in force* in the sense that it is *to be followed*? Clearly, we have simply re-encountered the question of action-guidingness, now in the form of a distinction between "formal" and (and what we might call) "normative" applicability. If at this point we ask for another rule, a "rule of application", the threat of regress emerges at once – for how to distinguish those cases in which the rule of application itself normatively applies among those in which it merely formally applies?

We could block the regress if there were a super-rule (rationality?) that always normatively applies and that directs us regarding the applicability of all other rules. Unfortunately, however, the useful transparency of anything like the *norma* – or of such familiar examples as rules of a game – is lost once we speak of super-rules. For we can intelligibly ask when to use the *norma* – or when to play a game – and why. But somehow, a super-rule is supposed to prevent such questions about itself from arising. Even as strong a proponent of rules and rationality as Kant seemed able to see the sense of asking what might be "the purpose of nature in

as to how the cut will in fact be made. His subsequent cutting performance is "guided" but not "predicted" by it, so actual failure on his part to conform to the *norma* does not impugn or discredit the *norma a posteriori*. For further discussion of these examples, and their relation to the *a priori* status of norms and rules, see P. Railton, "*A Priori* Rules: Wittgenstein on the Normativity of Logic", forthcoming.

attaching reason to our will as its governor" (G 305).[2] This is a question about the *normative* applicability of "rules of reason", that is, a question about the source of reason's normative authority.

Normative authority

Authority is an impressive thing. At least, it is when it works. We speak of rules *binding* us, or being *in force*, even when we would rather not comply. This suggests a certain image of what it would be to explain or ground normative authority. But though sheer force is sometimes called upon to enforce norms, it is not much of a model of the "coercive power" of norms as such. Rousseau noted that "If force compels obedience, there is no need to invoke a duty to obey".[3] A sufficiently great actual force simply *is* irresistible. Familiar rules and *oughts*, even stringent ones, are not like that – we can and do resist them, as Kant noted:

> The moral law is holy (inviolable). Man is certainly unholy enough, but humanity in his person must be holy to him. [CPrR 87]

Clearly the *must* here is not the *must* of something irresistible – the moral law is normatively, not actually, "inviolable". Since an *ought* is to apply to us even when we fall short, its force (and recognition thereof) must leave that option open. If "guidance by norms" is to play a nontrivial role in the explaining of an individual's or group's behavior, then the normative domain must be a domain of freedom as well as "bindingness".

This need for a "possibility of incorrectness" is often remarked

[2] Herein I will use the following abbreviations in citing work of Immanuel Kant: CJ = *Critique of Judgment*, trans. by Werner S. Pluhar (Indianapolis: Hackett, 1987); CJm = *Critique of Judgment*, trans. James Creed Meredith (Oxford: Clarendon, 1952); CPrR = *Critique of Practical Reason*, trans. by Lewis White Beck (Indianapolis: Bobbs-Merrill, 1956); CPrRm = *Critique of Practical Reason*, trans. by Mary Gregor (Cambridge: Cambridge Unviersity Press, 1996); G = *Groundwork of the Metaphysics of Morals*, trans. by H. J. Paton, 3rd ed. (New York: Harper & Row, 1956); LoE = *Lectures on Ethics*, ed. by P. Heath and J. B. Schneewind, trans. by P. Heath (Cambridge: Cambridge University Press, 1997); MM = *Metaphysics of Morals*, ed. and trans. by Mary Gregor (Cambridge: Cambridge University Press, 1996); OBS = *Observations on the Feeling of the Beautiful and Sublime*, trans. by John T. Goldthwait (Berkeley: University of California Press, 1960); SRL = "On a Supposed Right to Lie from Philanthropy", in *Immanuel Kant: Practical Philosophy*, trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996). All page numbers are to the Academy edition; Academy volume numbers are given only for the *Lectures on Ethics*.

[3] Jean-Jacques Rousseau, *The Social Contract*, trans. Maurice Cranston (Baltimore: Penguin, 1968), Bk. I, ch. 3, p. 53.

upon in philosophical discussions of normativity, usually in connection with physical or causal possibility. But it is no less important to make room for the *logical* or *conceptual* possibility of error. It is sometimes said, for example, that a *free agent* is by definition guided by rationality or a good will. There is no objection to this kind of definition as such, but it does not capture the sense of 'freedom' we need here.

Consider a more mundane example. Suppose that I have written you a letter and have spelled 'correspondence' correctly, rather than as the often-seen 'correspondance'. You, the reader, aware that my spelling is at best uncertain, remark upon my unexpected success to a colleague and wonder aloud whether it was accident or competence. You are, in effect, assessing two explanations, according to one of which I spelled it with an 'e' by chance, while according to the other I did it on purpose (though perhaps without explicit deliberation) – as a manifestation of my internalization of, and deference to, this particular norm of English spelling. Suppose your friend replies, "No, there simply is no question of why Railton spelled 'correspondence' with an 'e'. *Spelling* is a normative concept – acts of spelling constitutively involve satisfying the norms of spelling. So he *couldn't* have spelled the word with an 'a' – to have written 'correspondance' wouldn't have counted as a spelling of 'correspondence' at all."

Now there certainly is a "normative sense" of spelling, according to which 'correspondance' cannot count as a spelling of 'correspondence'. In this sense, it is analytic that spelling is correct, and even losers in spelling bees never spell incorrectly. That's why, though it may sound odd to say so, when we ask why or how someone spelled correctly we typically are *not* using the term in this "normative sense". As you intended your question to your colleague, my spelling 'correspondence' with an 'e' was either a happy accident or a pleasant surprise, not an analytic truth.

If a normative *must* is to have a distinctive place in the world, then, it cannot be the *must* either of natural law or of conceptual necessity. Natural law and conceptual necessities are "always at work", even when we're tired, weak-willed, lazy, disobedient, evil, or ignorant. No worry about anyone violating *them*. But normative guidance requires some contribution on our part, in a domain where freedom in the "non-normative" sense makes some vigilance or effort necessary.

However, having escaped the danger of missing the phenomenon of normative guidance altogether by assimilating it to a kind

of unfreedom, we had better be careful not to think of it as simply a matter of free willing. First, many of the *attitudes* (and associated motives and emotions) basic to normative conduct – attitudes of belief, desire, admiration, regret, approval, anger, and so on – appear not to be wholly within the scope of direct willing.[4] Kant, for example, distinguishes attitudes of love and reverence (*reverentia*), which are not directly subject to the will and cannot strictly be objects of duty (MM 401–403), from attitudes that accord to others a respectful observance (*observantia*) of their rights or goals, which can be required of us as a duty (MM 449, 467–468; compare G 399).[5] Kant does not conclude that attitudes of the first sort are therefore irrelevant to the domain of normative governance – on the contrary, according to the interpretation to be discussed below, they are to be found at the very bottom of his view, as a source or "basis" of duties (cf. MM 402–403).

Second, even if we restrict attention to those areas of normative governance in which the will seemingly can be effective – in selecting among acts, in regulating the more voluntary attitudes (such as acceptance or acknowledgement), and in shaping indirectly over time the less voluntary attitudes and motives (such as esteem, reverence, or liking) – it seems we cannot capture all of normative guidance with the notion of freely willing. For though the will may guide us, what guides the will? If we say, simply, "*We* do – we exercise our normative freedom and choose", this appears to get at only half the truth. For what makes an exercise of will a *choice*, rather than a mere *fiat*? And what would make a choice a moral one – or a rational, aesthetic, prudential, or epistemic one? Could the bare fact that a will is *my* will make it (say) a *good* will?

Reason and normativity

Kant tells us that reason's "highest practical function" is to enable us to discover and "establish" the good will (G 396), but speaking

[4] Perhaps *judgments* concerning these attitudes are more directly within the scope of will, but it is one thing to form a belief or feel an emotion, and another to form a judgment of it. Although our judgment is supposed to guide our belief, our beliefs might in fact prove recalcitrant. Thus we say: judgment is *normative for* attitudes like belief or feelings like appreciation. For a seminal discussion of evaluation as normative for attitudes, see Elizabeth Anderson, *Value in Ethics and Economics* (Cambridge: Harvard University Press, 1993), ch. 2.

[5] I am grateful to Peter Vranas for bringing to my attention this discussion in Kant of *reverentia* vs. *observantia*.

of reason and rationality can be ambiguous, at least in ordinary discourse. Let us distinguish, roughly, two senses of 'rational choice'.

In the first sense, a rational choice is a *well-reasoned* choice, one that is (or, perhaps, could in principle be) supported by a chain of deliberation in accord with norms of good reasoning. In the second sense, a rational choice is a choice *appropriately responsive to reasons*, whether or not it is (or, perhaps, even could in principle be) supported by such deliberation.

A simple example might help here. Consider a circumstance in which it would be best to pick an option from among those saliently available, rather than to deliberate – perhaps time is short, or perhaps the question is of little significance. To be "appropriately responsive to reasons" would involve prompt and decisive selection of one option and moving on. If we were even to stop and deliberate about *whether* to deliberate, we might miss our chance, or waste valuable time. In such cases, the two senses of 'rational choice' come apart in practice.

Yet we might hold that this represents no deep ambiguity in our basic thinking about practical rationality. For it seems we could, in principle, in a restrospective "context of justification", give a well-reasoned argument in favor of selecting without deliberation in certain circumstances. Indeed, it is not uncommon to find philosophers supposing that the two senses of 'rational choice' always come to the same thing, at least once we understand "well-reasoned" in terms of an in-principle constructable argument in the context of justification rather than a piece of actual cogitation in heat of the moment. And in this coming together of "well-reasoned" and "responsive to reasons" we might hope to find the secret to explaining how the free and forceful elements of normativity can be combined. Perhaps we can understand normative force on the model of appreciating *the force of argument.*

The force of argument has many features that make it an appealing general model for normative guidance. Unlike an irresistible coercive or natural force, the force of argument is one we can fail to follow. We have all departed from laws of logic by reasoning fallaciously, and we have all had the experience of finding our actual belief tendencies somewhat recalcitrant in the face of an argument whose validity and premises we cannot fault. The connection between the force of argument and belief is a normative one, rather than a matter of nomic or conceptual necessity.

At the same time, our response to the force of argument seems appropriately free without being arbitrarily willful. When we feel "trapped" by an argument or "caught" in a contradiction, we want out, but we are not inclined to think that we can, with sufficient power of will or strength of desire, bend the logical relations and escape. Moreover, even though logical relations thus stand independent of our will and wishes, recognition of them does not seem to be at odds with our capacity for autonomy in thought and belief. Since we take our beliefs to aim at truth and to be responsive to logic and evidence – one might even say this sort of commitment is *constitutive* of belief as an attitude[6] – we do not need to be subject to some further coercion or external sanction in order for self-acknowledged logical implications to be felt as putting normative pressure on us. We think we can see responsiveness to argument as a form of *epistemic attunement* of just the sort belief presents itself as having – attunement to content, to relations of implication and evidence, and so on.

"The force of argument" is indeed a central example of the peculiar mixture of force and freedom that we take normative guidance to involve. If it were possible to understand all normative guidance on this model, then we might hope that the two senses of 'rational choice' would never lead to genuinely divided loyalty and that we had gotten to the bottom of things normative. No doubt the lasting appeal of rationalism in philosophy is partly explained by this.

But I will spend most of the balance of this chapter discussing – in a very preliminary way – some ways in which the force of argument seems unable to afford a general model of normative guidance, or to take us to the bottom of all things normative. I will look first at what might seem the most hospitable territory for the force of argument: epistemology, or reasons for belief. Second, I will look at another domain of judgment, which might at first strike us as peripheral but instead emerges as central: aesthetics. Third, I will consider the classic turf for normativity: morality.

[6] For discussion, see David Velleman, "The Guise of the Good," *Nous* 26 (1992): 3-26, and "On the Possibility of Practical Reason", *Ethics* 106 (1996): 694-726; also, P. Railton, "On the Hypothetical and Non-Hypothetical in Reasoning about Belief and Action", in G. Cullity and B. Gaut (eds.), *Ethics and Practical Reason* (Oxford: Clarendon, 1997).

Normative authority for belief

We face a problem at the very outset attempting to understand normative authority in the domains of theoretical or practical reason in terms of the force of argument. For arguments and the logical relations they involve operate on, and conclude in, *propositions*. But according to a long tradition that seems worth maintaining, the conclusion of a piece of practical reasoning is an action and the conclusion of a piece of theoretical reasoning is a belief, and neither a belief nor an action is a proposition. If we are somehow to connect the propositional conclusion of an argument to a phenomenon like belief or action, it seems as if some non-argumentative but nonetheless *justifying* or *"rationalizing"* relationship must be found. Can we do this without already introducing a species of normative authorization not encompassed by the power of argument?

This is a contested matter. For example, we are inclined to speak of sensory experience as paradigmatically justifying perceptual belief, yet it is far from obvious that the content of experience itself is propositional, or that the justificatory relationship of this content to perceptual belief can fully be captured in deductive or inductive relations among propositions. To explore these questions would take us into deep waters. But perhaps we can give a less controversial example of justified belief to illustrate how difficult it would be to reconstruct all epistemic justification propositionally.

So as not to prejudice matters against "propositionalism", let us make some favorable assumptions. Suppose that we were able to give an uncontroversial account of "the force of argument" in the inductive case, that is, of what it is for a hypothesis to be inductively supported to a certain degree by a given body of evidence. And suppose as well that we can state the "rationalizing" relationship linking justified belief to inductive argument by a simple formula: a *belief that h* of strength r is justified in epistemic context C if h is inductively supported in C to degree r.

Focus now on beliefs that ascribe self-identity. Some such beliefs, I trust, are in fact epistemically justified. Can we give an account of this justification in propositional terms, even under our favorable assumptions? Perhaps, one might suppose, they are justified on the basis of an inductive inference from certain coherences and continuities among one's experiences. Consider an argument of the form:

(SI) I have experience e1 at t-3
I have experience e2 at t-2.
I have experience e3 at t-1.
I have experience e4 at t.
Experiences e1–e4 exhibit coherence and continuity.

I therefore conclude (with strength *r*) that I am self-identically me throughout the time interval (t-3) to t.

Yet it is clear that this argument simply *presupposes* self-identity, since it is formulated in terms of (a presumably unequivocating) first-personal 'I'. Now propositions are essentially third-personal, so we would have to reformulate the argument replacing 'I' and 'me' with 'Peter Railton'. Suppose this done, and suppose there is no doubt about the truth of the premises or the argument's inductive legitimacy. We now have a conclusion about Peter Railton, but it tells me nothing yet about *my* identity. That is, it does not yet sustain a conclusion licensing a *de se* self-identity ascription on my part.[7] It does not tell me that 'Peter Railton' refers to me.

If experiential induction, propositionally construed, will not suffice, where does my sense of self-identity and my entitlement (if any) to the first-personal 'I' come from? Presumably I arrive at a *sense* of being me (and here, and now) in part from something like what has been called *proprioceptive* aspects of my experience (both conscious and nonconscious) – a kind of feeling or expectation that pervades my mental life and which, so far as I can see, cannot in principle be rendered as a third-personal propositional content.[8] Now, if we dismiss this as no more than my "sense" of self-identity, and insist that we would need *evidence* reconstructable in argumentative form in order to *warrant* such a conclusion, we will find ourselves cut off from any possible avenue of

[7] See David Lewis, "Attitudes *De Dicto* and *De Se*", in his *Philosophical Papers*, vol. I (New York: Oxford University Press, 1983).

[8] There is some experimental evidence in the literature on autism that autistic individuals may experience deficits in developing a feeling for the self, much as individuals can experience color deficits in ordinary perception. Autistic individuals, for example, experience difficulty with first- vs. third-person asymmetries in so-called "false belief tasks", and are known to lose track of first- and second-personal pronouns in conversations, as in the phenomenon of "echo-locution". After reviewing a description of a cognitively very high-functioning autistic individual, Temple Grandin, who herself professes finding ordinary social language and exchange baffling, but technical or scientific language much clearer, Simon Baron-Cohen writes:

> And her own explanation . . . ? "She surmises that her mind is lacking in some of the 'subjectivity,' the inwardness, that others seem to have.

From *Mindblindness* (Cambridge: MIT Press, 1995), pp. 142–143.

justification. This could leave us stranded as theoretical reasoners, since without any entitlement to the 'I', how am I ever to be responsive in my belief to the evidence *I* have? – A lot of people have a lot of evidence, much of it conflicting, but whose should weigh with me? To justify my beliefs I need to identify myself in the space of epistemic reasons.

Hume himself seems to have become sensible of such a defect in any purely continuity-and-coherence-based approach to personal identity, such as the one he experimented with in the *Treatise*. He reflected in an Appendix:

> If perceptions are distinct existences, they form a whole only by being connected together. But no connexions among distinct existences are ever discoverable by human understanding. We only *feel* a connexion or determination of the thought . . . the ideas are felt to be connected together, and naturally introduce each other. [T 635][9]

He is at a loss to describe this feeling, or to explain it as based upon principles. "[T]his difficulty," he concedes, "is too hard for my understanding" [T 636].

Just what a fix we could end up in is seen at the end of Part I of the *Treatise*, where Hume gives a perhaps inadvertent intimation of the problem his later reflection brought clearly into focus. Hume is describing the depths of the mental distress he reaches as a result of an "*intense*" commitment to following the rationalistic maxim to restrict belief to those matters where we can give a reasoned justification. He finds that, as a result, he loses any entitlement to confidence in induction, memory, external body, or even deduction. Eventually he "can look upon no opinion even as more probable or likely than another", and calls out in desperation, "Where I am I, or what?" (T 269). Rigorous adherence to the self-imposed rationalist maxim prevents him from attibuting any epistemic authority to his "natural introduction" to the self via an unreasoned "feeling" of it – and he thus loses his grip on self-location and self-identity.

Having seen what it would be to reach this point, Hume cannot convince himself that epistemology would be well-served by

[9] Here are the abbreviations used in the text for Hume's writings: Inq = *Inquiry Concerning the Principles of Morals*, ed. by C. W. Hendel (Indianapolis: Bobbs-Merrill, 1957); T = *Treatise of Human Nature*, ed. by L. A. Selby-Bigge (Oxford: Oxford University Press, 1888); ST = "Of the Standard of Taste", in *Of the Standard of Taste and Other Essays by David Hume*, ed. by John W. Lenz (Indianapolis: Bobbs-Merrill, 1965).

unqualified obedience to the rationalistic maxim. Why is it, he wonders, that

> . . . I must torture my brain . . . at the very time I cannot satisfy myself concerning the reasonableness of so painful an application, nor have any tolerable prospect of arriving by its means at truth or certainty? Under what obligation do I lie . . . ? [T 270]

Hume remains concerned with reasonableness, truth, and probability. He is, however, "sceptical" that trusting *only* the force of argument will enable us to be fully responsive to these concerns.

> . . . understanding, when it acts alone, and according to its most general principles, entirely subverts itself, and leaves not the lowest degree of evidence in any proposition, either in philosophy or in common life. . . . I am ready to reject all belief and reasoning . . . . Whose favor shall I court, and whose anger must I dread? What beings surround me? and on whom have I any influence, or who have any influence on me? I am confounded . . . and begin to fancy myself . . . utterly depriv'd of the use of every member and faculty. [T 268–269]

Far from consolidating belief around a core of rational certainty like the Cartesian *cogito*, Hume finds himself in a complete collapse of normative epistemic guidance – there remains no discernment concerning evidence or probability, no sense of anyone's authority, even one's own. His "distribution of credence" has become entirely undiscriminating, even with respect to logical relations and "the force of argument". How, for example, are we to *reason* in the "context of justification" about the relationship between our beliefs and their grounds if we accord immediate experience no *prima facie* authority to support belief even concerning the content of our own thoughts?

If belief and reasoning are to be resurrected, we will need to authorize ourselves to draw directly upon a wider base of epistemic resources, without asking for reconstructability as argument, even in the context of justification. But what to add? Belief, we've noticed, is not a bare proposition, but an *attitude* toward propositions. Hume puts it starkly: "*belief is nothing but a peculiar feeling, different from the simple conception* [of its object]" (T 624). If we consider *de se* belief, Hume's suggestion would seem to be that this attitude is a feeling that is *to be regulated* (at least in part) by

"self-introducing" (we might say "self-intimating") feelings. A feeling regulating a feeling? Hume writes that "*belief is more properly an act of the sensitive, than of the cogitative part of our natures*" [T 181]. Hume appears to apply this idea well beyond self-identifying belief, stressing the role of feelings in shaping belief concerning external objects, and observing:

> Nature has . . . doubtless esteem'd it an affair of too great importance to be trusted to our uncertain reasonings and speculations. [T 187]

But what is such regulation of feeling by feeling like, and, if it cannot be reconstructed as a argument, how can it constitute justification? It seems we will need to supplement the normative "force of argument" in epistemology with something like a normative "force of feeling", if we are to resuscitate epistemic discrimination or even self-discernment. How can feeling be appropriately discerning to possess epistemic authority? To have some idea of how this might go, we will turn to another work of Hume's – on discerning, knowing, appreciative feelings.

Normative authority and appreciation

We encounter a structurally similar problem – of how to find the resources necessary to support a domain of appropriate discrimination in judgment – in Hume's late essay, "Of the Standard of Taste", which apparently is a survival of a systematic project he had undertaken on the nature of "criticism", to include morality as well.[10] After observing that we cannot ground aesthetic distinctions on "reasonings *a priori*" (ST 231), he begins to consider the possible contribution of sentiment. Yet he quickly finds that mere *acquiescence* in sentiment would equally leave aesthetic distinctions groundless:

> There is a species of philosophy, which cuts off all hopes of success in such an attempt, and represents the impossibility of ever attaining any standard of taste. The difference, it is said, is very wide between judgment and sentiment. All sentiment is right; because sentiment has a reference to nothing beyond

[10] See David Fate Norton, "Introduction to Hume's Thought", in his edited collection, *The Cambridge Companion to Hume* (Cambridge: Cambridge University Press, 1993), p. 27.

> itself . . . . . [E]very individual [therefore] ought to acquiesce in his own sentiment, without pretending to regulate those of others. . . . [And thus it is] fruitless to dispute concerning tastes. [ST 230]

This species of philosophy has the wholly "sceptical" result that we cannot even say that Milton is better than Ogilby, and any such philosophy effectively undermines the discrimination upon which taste must be based. Agreeable as this "levelling" sort of skepticism may be to some strands of common sense, common sense on the whole, Hume notes, does not really take it to heart:

> Whoever would assert an equality of genius and elegance between OGILBY and MILTON, or BUNYAN and ADDISON, would be thought to defend no less an extravagance, than if we had maintained a mole-hill to be as high as TENERIFE, or a pond as extensive as the ocean. Though there may be found persons, who give preference to the former authors; no one pays attention to such a taste; and we pronounce without scruple the sentiment of these pretended critics to be absurd and ridiculous. [ST 230–231]

Hume isn't personally threatened by a "species of philosophy" that would forced us to give up aesthetic distinctions. "The principle of natural equality of tastes", he believes, can hold sway only in disputatious or esoteric settings where we are not actively relying upon taste to guide us. In ordinary life, it is "totally forgot" (ST 231). Unlike the younger Hume, who wrestled nearly to the point of exhaustion with reason's normative force, worrying aloud "For my part, I know not what ought to be done" (T 268), the older Hume who wrote "Of the Standard of Taste" seems confident that he knows reason's place and unfraid of the world of normative discrimination tumbling into ruin around him. Any aesthetician – rationalist or sentimentalist – who cannot find a basis for distinguishing a Milton from an Ogilby will simply find himself without authority in Hume's eyes, or ours.

To whose taste, then, *do* we actually pay some attention, i.e., attribute some normative force, and why would this count as authority about *beauty*? Hume identifies two sources of authority, convergence of "expert opinion" among those with relevant knowledge and sensory discriminative capacities, and convergence of general, experienced opinion in the "test of time". In both cases, we are seen to accord some authority to these sources,

beyond our own simple likings. After all, we know that our own simple likings, convincing though they may be as feelings of attraction, may nonetheless be attributable to our own partiality, ignorance, fashion, novelty, lack of sensory discrimination, or distaste for (or perverse fascination with) the odd or déclassé. Why should this matter – isn't it up to us what we like? Yes, but when we judge beauty, we attribute something to an object or event, not merely to ourselves; and we accord ourselves authority concerning it. Partiality, fashion, lack of sensory discrimination, etc., are all ways in which the pleasure one takes in the experience of a landscape or of a work of art might simply be unrelated to the "beauties" (in Hume's terminology) it possesses – since we do not think self-interest, fashion, and the like are, or "make for", genuine beauty.

Well then, what sorts of features do we uncontroversially take to have a constitutive role in beauty-making, in both natural and man-made objects? Where do we expect to find the "beauties"? Surely, if there is anything at all to our notion of beauty, then among these features are: form, proportion, color, texture, composition, melody, harmony, rhythm, progression, and the like. When these features of an object are of a kind that our sensory and cognitive engagement with them seems reliably to yield experiences we find intrinsically enjoyable, we seem to have (to that extent) a candidate for beauty. That such features do figure in our assessments of beauty is reflected in ways we typically attribute lesser or greater aesthetic authority to our own likings or the likings of others. For example, I do not take my likings concerning Middle Eastern music to have much authority – I am inexperienced with it, unable to discern its shades of tonality, structures, progressions, or variety (the different pieces sound too much alike to me), I don't claim to be exercising taste or discernment in when I express sporadic likes and dislikes of what I happen to hear. And I certainly claim no authority over others. By contrast, there are those whose likings in Middle Eastern music I find much more authoritative than mine, and whom I would consult for guidance. Now someone I take to be expert could lose some standing in my eyes if I came to learn that he plays favourites, judges music by its ideological content, lacks sensory discernment, or cannot find other individuals seriously engaged in making or judging such music who take his judgment seriously. Our practices – including our patterns of normative deference – reveal that we do have some idea of what it would be

for a feeling (an appreciative delight) to be more or less attuned to objective, beauty-making features of objects, even though this attunement is effected in part via careful cultivation of, and attention to, subjective feelings or sensations.

A degree of deference to experts who possess demonstrable skills of discernment, greater knowledge of genre or context, wider experience, and so on, enables me to extend my "critical" power in detecting beauty-making features – they help me form a better idea of what I'd find delightful were I to gain greater experience. As a result, they help attune me to the "beauties" of objects, features which can be rich and lasting sources of sensorily-based, cognitively-engaging delight. Hume puts it thus:

> Those finer emotions of the mind are of a very tender and delicate nature, and require the concurrence of many favourable circumstances to make them play with facility and exactness, according to their general and established principles. The least exterior hindrance to such small springs, or the least internal disorder . . . and we shall be unable to judge of the catholic and universal beauty. The relation, which nature has placed between the form and the sentiment, will at least be more obscure; and it will require greater accuracy to trace and discern it. [ST 232–233]

A similar sort of authority, also related to an authority we already accord ourselves, attaches to the "test of time". Hume writes, concerning the relation "nature has put between form and sentiment" which underlies beauty:

> We shall be able to ascertain its influence not so much from the operation of each particular beauty, as from the durable admiration, which attends those works, that have survived all the caprices of mode and fashion, all the mistakes of ignorance and envy.
>
> The same HOMER, who pleased at ATHENS and ROME two thousand years ago, is still admired at PARIS and at LONDON. All the changes of climate, government, religion, and language, have not been able to obscure his glory. [ST 233]

Long exposure, developed sensibilities, the authority of countless experiences on the part of different individuals – how far we are from my inexperienced self overhearing a snatch of Middle

Eastern music at lunch and saying "Hmm, don't care much for that". It is natural to see this as a difference in attunement to musical value.

Over the course of a life, we participate in a complex critical and appreciative practice, attributing some authority to our own growing experience ("In the end, the proof of the pudding . . ."), making recommendations and seeking confirmation in the opinions of others ("Try it, you'll see for yourself"), and also showing some deference to various external sources of authority ("After what I've heard about it, I'm eager to try this place"). Situated within such a practice, which extends across societies and times and is held together both by our fundamental human sensory and cognitive similarities and by our reciprocal deferences, my judgments of beauty have at least a chance to be "normed by" the sources of aesthetic value, and words like 'beautiful' in my mouth have a chance of expressing genuinely aesthetic evaluations, even when I get things wrong.[11] We manage, that is, to have a domain of real distinctions concerning beauty, a domain of genuine taste, even though "subjective feelings" play an essential role in its shape.

Kant was also concerned to underwrite the possibility of objectivity in the domain of taste. Like Hume, he worried about various ways in which appreciation might be attuned or disattuned to genuine value. Kant writes:

> . . . everyone says: Hunger is the best sauce; [but] to people with a healthy appetite anything is tasty provided it is edible. Hence if people have a liking of this sort, that does not prove that they are selecting by taste. Only when their need has been satisfied can we tell who in a multitude of people has taste and who does not. [CJ 210]

Hunger makes our likings unreliable. But when, for Kant, could a subjective condition such as liking be a reliable guide to a purportedly objective matter, such as aesthetic value?

Kant could not pursue Hume's solution, of looking to the refinement and qualification of empirical faculties and sentiments. Hume's psychology attributes to "the internal frame and

[11] A common standard of time and shared conventions about when to arrive for (say) a noon engagement make it possible for me to be *on time*, but also *late*. In the case of good – and bad – taste, something more than this conventional infrastructure is required, e.g., Hume's account of *beauties* to be attuned to.

constitution of the mind" appetites and passions that are *directly* aimed at features of the world independent of the self, and are "antecedent" to self-interest or happiness (Inq 113–119). But in Kant's empirical psychology, by contrast, appetites and passions are always guided at base by one's own pleasure:

> All the inclinations together (which can be brought into a tolerable system and the satisfaction of which is then called one's happiness) constitute regard for oneself (*solipsismus*). [CPrRm 73][12]

Within such a psychology, to become ever more delicately attuned to nuance in one's empirical feelings would simply be to become ever more attentive to promoting personal pleasure, regardless of how the pleasure is produced, whether any appreciative or cognitive faculties are engaged, and whatever the nature of the cause of the pleasure. Pleasure and affect are in this sense "blind" for Kant (CJ 272), since "if our sole aim were enjoyment, it would be foolish to be scrupulous about the means of getting it" (CJ 208). An Oriental massage in which the joints and muscles are agreeably "squeezed and bent" would be lumped together with a stirring Greek tragedy (CJ 274).

In aesthetics, we must focus not on which phenomena produce the greatest or most intense pleasure, but rather on the "presentation" of objects to the senses: we must be able to see the object "as poets do", and "must base our judgment regarding it merely on how we see it" (CJ 270), that is, on the genuinely beauty-making characteristics. Self-oriented and pleasure-seeking, our empirical sentiments are careless as to modality. Kant thus foretold the fate that awaited aesthetics in the hands of that redoubtably thorough-going proponent of egoistic hedonism, Bentham: the only ground of discrimination would be *quantity*, the "mass of agreeable sensation" (CJ 266) – and pushpin (or Oriental massage) would indeed be deemed as good as poetry.

Moreover, Kant joined Hume in insisting that aesthetic judgments purport to be "non-personal" and communicable to others

[12] We can see an analogy with the case of theoretical reason. If we thought that all *inclination to believe* was essentially self-regarding (*solipsismus*), and attuned to gratification rather than objective conditions, truth, or evidence, then we would find genuine "epistemic worth" only in a dutiful capacity to resist epistemic inclination and regulate belief by epistemic principle alone. This would not make "epistemic dutifulness" into the "highest end" of epistemic activity – that would remain the marriage of justified belief with truth that constitutes knowledge – , but into an indispensable condition of it.

– in the sense not only of *informing* others concerning what we like, but of *recommending*, where each of us purports to have potential authority for others. "But," Kant argues,

> if we suppose that our liking for the object consists merely in the object's gratifying us through charm or emotion, then we also must not require anyone *else* to assent to an aesthetic judgment *we* make; for that sort of liking each person rightly consults only his private sense. [CJ 278]

For similar reasons, Kant insists that in order to ensure that our account is "concerned solely with aesthetic judgments", "we must not take for our examples such beautiful or sublime objects of nature as presuppose the concept of a purpose" (CJ 269–270). To the extent that the force of an example can be attributed to purpose (e.g., self-interest), the judgment will not be aesthetically attuned – we might substitute for the object of appreciation anything that would bring about the sought-after result equally well.

> . . . the purposiveness would be either teleological, and hence not aesthetic, or else be based on mere sensations of an object (gratification or pain) and hence not merely formal. [CJ 270]

Therefore:

> It seems, then, that we must not regard a judgment of taste as *egoistic* . . . we must acknowledge it to be a judgment that is entitled to a claim that everyone else ought also to agree with it. But if that is so, then it must be based on some a priori principle (whether objective or subjective) . . . [J]udgments of taste presuppose such a command, because they insist that our liking be connected *directly* with a presentation. [CJ 278]

If our judgment is to be attuned to the sources of aesthetic value by a "liking" that is "connected *directly* with a presentation", but empirical likings cannot do this, where then is taste's infrastructure, where to turn for regulation of our feeling of appreciation – for Kant insists that appreciation, even of the beautiful and the good, is a *liking*, a feeling (CJ 210)?

Kant looks to reason. The seeming peculiarity of Kant's aesthetic, that it sees aesthetic judgments as "demands of reason", can be understood in this light. But we must be careful, for such demands of reason are *not* demands based upon argument, rule, or conceptual demonstration:

> . . . the beautiful must not be estimated according to concepts, but by the final mode in which the imagination is attuned so as to accord with the faculty of concepts generally; and so rule and precept are incapable as serving as the requisite subjective standard for that aesthetic and unconditioned finality in fine art which has to make a warranted claim to being bound to please. Rather must such a standard be sought in the element of mere nature of the Subject, which cannot be comprehended under rules or concepts, that is to say, the supersensible substrate of all the Subject's faculties (unattainable by any concept of understanding) . . . . [CJm 344]

Here, then, we have Kant's version of the subjective attunement that affords reliable guidance concerning the beauty-making features of the world: the pleasure afforded by activity on the part of the self's supersensible substrate, when directly engaging the sensory "presentation" of the object. This substrate, shared as it is by all rational humanity, helps supply the needed infrastructure for a domain of objective taste. Now an invocation of a supersensible substrate may sound like hocus-pocus, but Kant deserves credit for refusing to be false to the "non-personal" compellingness of the experience of aesthetic appreciation, in order to satisfy an allegedly scientific egoistic, hedonist psychology. Not hiding its "unfathomableness", Kant gives the best explanation he can: only the rational self has the requisite formal, disinterested, "nonpersonal", and universal character to be the source of such a pleasure.

But Kant's rational self is not simply a *reasoning* self. Beauty is a "way of presenting" that requires concepts, yet Kant recognizes that aesthetic appreciation is not simply a matter of being "brought to concepts" (CJ 266). If we were nothing but "pure intelligences", "we would not present in this way" and could not see beauty (CJ 270). Nor is the rational self the whole infrastructure. According to Kant, beauty "holds" – presumably, is capable of "norming" judgment through feelings of appreciation and the practice of taste – only for "beings who are animal and yet rational, though it is not enough that they be rational" (CJ 210).[13]

Despite the indispensable role of reason, then, in attuning us to the beautiful, the normative force of judgments of beauty, even

[13] According to Kant, an appreciation of the *sublime* also depends upon a "way of presenting", and so is not available to a pure intelligence (CJ 270). However, he also believes that our capacity to appreciate the sublime does not depend upon our animal nature. More on the sublime, below.

for a rationalist aesthetic such as Kant's, is not the force of argument. We therefore cannot expect that we could *reconstruct* aesthetic justification in propositional terms. As in the case of *de se* attitudes, an attitude (in this case, aesthetic appreciation) may stand in a justified relationship to its proper object even though this relationship is not mirrored in an argumentative relationship among propositions.

In appreciation we find the right mix of force and freedom for normative guidance. On the one hand, "the liking involved in our taste for the beautiful is disinterested and *free*" (CJ 210). On the other hand, we all know the *compelling* character of aesthetic appreciation and good criticism: we find in our first-personal experience of the object, as informed by the contributions of the critic, something both likeable and convincing. "Ah, *now* I see it," we think, thereby feeling the force of aesthetic authority: a force of credible influence from the critic ("He helped me see it"), of convincing experience from our own case ("Now I get it"), of a compelling work ("There was a lot more in it than I thought"), and of a discovery of value that we can share with others ("You must try this" or "You must read his essay, it'll change how you look at Miró").

Wittgenstein, in his "Lectures on Aesthetics", gives as his model of aesthetic appreciation an example of this process, drawn from his own case:[14]

> Take the question: "How should poetry be read? What is the correct way of reading it?" . . . I had an experience with the 18th-century poet Klopstock. I found that the way to read him was to stress his metre abnormally. Klopstock put ⏑ – ⏑ (etc.) in front of his poems. When I read his poems in this new way, I said, "Ah-ha, now I know why he did this." What had happened? I had read this kind of stuff before and had been moderately bored, but when I read it in this particular way, intensely, I smiled, said "This is *grand*", etc. But I might not have said anything. The important fact is that I read it again and again . . . that I read the poems entirely differently, more intensely, and said to others: "Look! This is how they should be read." [LA, 4–5]

[14] Ludwig Wittgenstein, "Lectures on Aesthetics", in Cyrill Barrett (ed.), *L. Wittgenstein: Lectures and Conversations* (Berkeley: University of California Press, 1966). Hereinafter, LA.

Kant and Hume agree that, underlying aesthetic evaluation, there must be some form of "liking" or "enjoyment". Moreover, the liking in question must be sensorily-based, cognitively-engaging, discerning, disinterested, and communicable. If Hume is right, our essentially similar "internal fabric" – our empirical psychology and sentiments – can afford much of the ground for such a liking, since many of our appetites and passions take external conditions or sensory "forms or qualities" as their immediate objects and are disinterested in character, even though satisfying them will also yield pleasure. Thanks to additional qualification of feeling by the influence of reason, understanding, and the commerce of opinion, we can develop on this psychological "common ground" a domain of discernment and knowledge, where we can recognize and possess authority, and 'beauty' can have its true meaning – apart from fashionableness, novelty, endearing schlock, ponderous "importance", snobbish over-refinement, and so on. In Hume's account, as in Kant's, what possesses ultimate aesthetic authority is a *qualified appreciative attitude* and not a mere liking. In Hume's account, as in Kant's, much of the qualification of attitude is supplied by reason. And in Hume's account, as in Kant's, it seems we could not reconstruct aesthetic justification in terms of the force of argument.[15]

The normative authority of moral rules

Perhaps no one is really tempted by the idea that the normative force of aesthetic appreciation rests upon argument. But things might be different in the moral case, where the supremacy of reasoning and rules is often invoked. Perhaps in morality at least we will find it possible to account for normative force in terms of the force of argument.

Let us set aside for now a very general worry about this line of thought, briefly touched on in the introduction: any appeal to rules as a foundation for justification runs the risk of regress or circularity unless we can appeal to a super-rule of a mysterious kind. For now let us cheerfully assume that we don't mind mystery, as long as its name is *rationality*.

Kant's moral philosophy is often taken to be the *locus classicus* for

[15] For further discussion of Hume's aesthetic theory, see P. Railton, "Aesthetic Value, Moral Value, and the Ambitions of Naturalism", in Jerrold Levinson (ed.), *Aesthetics and Ethics* (Cambridge: Cambridge University Press, 1997).

the idea that normativity resides in rationality itself, and the moral law it prescribes. Perhaps this is indeed how we should understand his view: there is a super-rule, and it commands our obedience as a rational obligation. But is it obvious that this is how *he* understands his own most basic approach to normativity? We are told to have respect (*reverentia*) for the moral law, but Kant observes:

> Respect (*reverentia*) is, again, something subjective, a feeling of a special kind, not a judgment about an object that it would be a duty to bring about or promote. For, such a duty, regarded as a duty, could be represented to us only through the *respect* we have for it. A duty to have respect would thus amount to being put under obligation to duties . . . . [MM 402–403]

So it seems we must look for "a feeling of a special kind", not obligation, at the bottom of moral duty. What is this feeling like? Here is an example of the sort of reverential appreciative feeling Kant appears to have in mind:

> . . . to a humble, plain man, in whom I perceive righteousness in a higher degree than I am conscious of in myself, *my mind bows* whether I choose or not, however high I carry my head that he may not forget my superior position. . . . Respect is a tribute we cannot refuse to pay to merit whether we will or not; we can indeed outwardly withhold it, but we cannot help feeling it inwardly. [CPrR 76-77; compare G 454]

What we perceive in this individual is not simply more severe dutifulness than our own. We are all familiar with individuals who turn sensible everyday rules into severe duties that rise above all inclination, but our mind does not bow to that.[16] What we perceive, according to Kant, is greater *righteousness*, dutifulness that "includes" a good will (G 397).

In our appreciative encounter with it, we once again encounter the mixture of force and freedom characteristic of normative force. On the one hand, the respect is "freely paid" – for Kant, nothing in our experience suggests that any self-interested incentive or external coercion lies behind our appreciation. On the other hand, the respect is in a way compelled, it is something "we cannot help feeling", even when it comes in the face of interest. Kant writes:

[16] For a description of dutifulness of this kind, see David Schapiro, *Autonomy and Rigid Character* (New York: Basic Books, 1981), pp. 83-86.

> *Duty!* Thou sublime and mighty name that dost embrace nothing charming or insinuating but requirest submission and yet seekest not to move the will by threatening aught that would arouse natural aversion or terror which of itself finds entrance into the mind and yet gains reluctant reverence . . . . [CPrR 86]

Now this impressive paean might suggest an *intrinsic* evaluation of duty. But, as Paul Guyer reminds us,[17] Kant continues, still addressing "Duty":

> . . . what origin is there worthy of thee, and where is to be found the root of thy noble descent which proudly rejects all kinship with the inclinations and from which to be descended is the indispensable condition of the only worth which men can give themselves?
>
> It cannot be less than something which elevates man above himself as a part of the world of sense, something which connects him with an order of things which only the understanding can think and which has under it the whole system of all ends which alone is suitable to such unconditional practical laws as the moral. [CPrR 86–87]

Notice that the practical laws of morality, and even duty itself, are not self-subsistent sources of unconditional worth – their worth arises from their "descent", which does secure the noble standing of morality.[18]

At the bottom of morality's normative authority, then, Kant speaks not of an analytic demand of consistency nor a willful exercise of our capacity to govern ourselves by rules, but of an experienced synthetic demand and a free acknowledgement, the subjective expression of which is a feeling of a more aesthetic character, akin to the demand upon us that the appreciation for the sublime in nature involves:

> It is in fact difficult to think of a feeling for the sublime in nature without connecting it with a mental attunement similar to that for moral feeling. [CJ 128]

[17] See Paul Guyer, "Kant's Morality of Law and Morality of Freedom", in R. M. Dancy (ed.), *Kant and Critique* (Dordrecht: Kluwer, 1993), p. 70.

[18] Guyer emphasizes the consistency with which Kant, over the course of his philosophical career, recognized that all evaluation presupposes some values-in-their-own-right. The value Guyer identifies is the special *freedom* Kant attributes to human agents. See his "Kant's Morality of Law and Morality of Freedom".

For Kant, as we saw in the aesthetic case, human inclination and appetite cannot attune us to *this* sort of demand, because they are by nature self-interested ("*solipsismus*", CPrR 73) rather than non-personal and distinterested, and thus "human nature does not of itself harmonize with the good" (CJ 271). Kant therefore must find a faculty internal to us, capable of evincing or guiding a special sort of liking, a "moral feeling", that is attuned to the moral-value-making features of the world, the sources of moral *worth*. We can, he writes, be attuned to the good "only through the dominance that reason exerts over sensibility" (CJ 271). So, as in aesthetics, to underwrite a rational demand as grounded in the right sort of attunement, we must have recourse to a "supersensible substrate", a noumenal self. Moral judgments are akin to aesthetic judgments of sublimity – judgments of beauty draw in part upon our "animal" nature; for the moral and the sublime, reason alone, the "supersensible substrate", suffices.

Now for Hume, the "substrate" for moral and aesthetic judgment can be our empirical psychology, since it contains sentiments of a suitably "impersonal" and non-self-interested nature. For example,

> We are certain, that sympathy is a very powerful principle in human nature. We are also certain, that it has a great influence on the sense of beauty, when we regard external objects, as well as when we judge of morals. We find, that it has force sufficient to give us the strongest sentiments of approbation . . . . [T 618][19]

Thanks to sympathy, among other sentiments, our sentiment of direct approval can be attuned to the ends of others as such, and to the general interest, even when we have no personal interest at stake: reading ancient history, we wince at a tyrant's cruelty, and root for the hero to save the populace from him. And much aesthetic judgment, likewise, depends upon a capacity to feel the feelings of others. If well-developed, well-informed, and attentively listened to, such "impersonal" sentiments can attune us to – "harmonize" us with – the good and the beautiful.

> We may observe, that all the circumstances requisite for [sympathy's] operation are found in most of the virtues; which

[19] A more contemporary psychological account would notice that Hume's sympathy involves two elements: empathy (a direct internal simulation of the circumstances and mental states of others) and sympathy (a direct positive concern for their well-being).

> have, for the most part, a tendency to the good of society, or to that of the person possess'd of them. [T 618]

Sympathy can of course be misled, and may lead us astray. It may fail to be engaged in unfamiliar or misunderstood surroundings. Or it may immediately attune us to the evident pain of an animal undergoing an emergency veterinary procedure, making us wish fervently that the procedure would stop, even though this operation is necessary for the animal's survival. Sympathy – like aesthetic admiration – therefore must be assisted and qualified by knowledge, understanding of cause and effect, and reason, and by participation in a community in which our judgments may be challenged and improved if (as we tend to do) we launch our opinions into the public world and also to defer to some degree to the judgments of others and to social practices hammered into shape over the generations. Thus – once again, as in the aesthetic case – our feelings can develop greater freedom from prejudice, finer discrimination, and closer attunement to genuine moral distinctions.

By contrast Kant, as an egoistic hedonist in psychology but a universal humanist in morality, could no more entrust moral attunement to "solipsistic" empirical sentiment (cf. CPrR 73) than he could aesthetic attunement.[20] And thus we arrive at Kant's answer to the question why nature attached reason to will (which is, for Kant, also a *liking*; CJ 209): without the "substrate" of reason to ground impersonal feelings, we would arrive only at a personalistic willfulness, not a good (i.e., general) will. Hume gave us a story as to how the empirical, psychological "substrate" we share as humans generates likings that can be attuned to beauty and the general good. What mechanism does Kant give to explain how a "supersensible substrate" can function similarly?

[20] Contemporary empirical psychology on emotion, motivation, and moral development tends to favor a more Humean view. See for example, J. H. Barkow, L. Cosmides, and J. Tooby (eds.), *The Adapted Mind* (New York: Oxford University Press, 1992); Baron-Cohen, *Mindblindness*; Antonio Damasio, *Descartes' Error* (New York: Putnam, 1994); N. Eisenberg and J. Strayer (eds.), *Empathy and its Development* (Cambridge: Cambridge University Press, 1987); N. Eisenberg and P. Mussen (eds.), *The Roots of Prosocial Behavior in Children* (Cambridge: Cambridge University Press, 1989); Joseph LeDoux, *The Emotional Brain* (New York: Simon & Schuster, 1996); L. May, M. Friedman, and A. Clark (eds.), *Mind and Morals* (Cambridge: MIT Press, 1996); and David G. Myers, *The Pursuit of Happiness* (New York: William Morrow, 1992). Empathy has been credited in some historical cases with greater efficacy than principles in inhibiting compliance with cruelty commanded by authority. See Roy F. Baumeister, *Evil: Inside Human Violence and Cruelty* (New York: W.H. Freeman, 1997).

Here Kant is, as befits his penetration as a philosopher, entirely frank: he has no positive idea – the matter involves an "unfathomable depth of [a] supersensible power" (CJ 270; G).

Note, however, that Kant is also clear that reason cannot operate here by argument alone:

> . . . when in intuiting nature we expand our empirical power of presentation (mathematically or dynamically [a "might over the mind"]), then reason, the ability to [think] an independent and absolute totality, never fails to step in and arouse the mind to an effort, although a futile one . . . . . [W]e are compelled to subjectively *think* nature itself in its totality as the exhibition of something supersensible, without our being able to bring this exhibition about *objectively.*
>
> . . . We cannot determine this idea of the supersensible any further, and hence cannot *cognize* but can only *think* nature as an exhibition of it. . . . This judging strains the imagination because it is based on a feeling that the mind has a vocation that wholly transcends the domain of nature (namely, moral feeling), and it is with regard to this feeling that we judge the presentation of the object subjectively purposive. [CJ 268]

Our mind, in its "supersensible vocation", is here functioning in a way Hume would have recognized despite the heavily Kantian language: feeling and imagination are regulating judgment, beyond the scope of cognition and argument alone. Within this scheme, as within Hume's, we may *use* arguments to help us attain or correct a moral feeling or sentiment. For Kant, the "contradiction in conception" and "contradiction in will" tests of our practical maxims can place a purportedly good will face-to-face with its potential own limitations, deflating or affirming its self-representation as perfectly general. For Hume, understanding and general rules help to extend or correct untutored sympathy.

If reason's functioning as a supersensible substrate for feeling remains for Kant something of which he cannot give a positive account, he nonetheless believes we can convince ourselves of its possibility: we know from first-hand experience the "striking down" of our pretenses and humiliating acknowledgement of our own limitations, and we also know that reason alone among our faculties possesses the qualities necessary for such experience – it alone can furnish guidance that is impersonal. There is no mystery about this when we confront the sublime in nature or moral-

ity. The peculiar awe we experience when we come upon "a mountain whose snow-covered peak rises above the clouds" (OBS 47) or when we observe an act of genuine duty performed in spite of conditions of extreme "subjective limitation", has extraordinary power to move us, yet cannot be attributed to empirical sentiment. We find our own self-conceit "humiliated" or "struck down" (CPrR 73) in the presence of the sublime. Fortunately, we are not merely flattened. Instead, we are awakened to a value "beyond price", carried beyond ourselves for the moment to sense a "direct liking", a liking even of that which strikes at the very heart of our own prideful self-interest. Thus it recruits our fundamental allegiance, despite any personal interest to the contrary.

> To behold virtue in her proper shape is nothing other than to show morality stripped of all admixture with the sensuous and of all the spurious adornments of reward or self-love. How much she then casts into the shade all else that appears attractive to the inclinations can be readily perceived by every man if he will exert his reason in the slightest . . . . [G 61–62n]

No wonder such a "presentation" moves us, and yields not the "cold and lifeless approval, without any moving force or emotion" (CJ 273, 274) that we would otherwise expect from any merely un-self-interested presentation. Confronted with the sublime, we are not tempted to think, "Yeah, but what's it to *me*?" No wonder such a "presentation" is regulative for our wills when we are rational, i.e., attuned via our "supersensible substrate".

This has an important implication for our normative life together: since it owes nothing to personal interest, our sense of the sublime in nature and in conduct should be "subjectively" confirmable by other rational beings in their own experience. Others, too, Kant is confident, will stand in awe before the Alps during a storm or find that their mind bows when observing a humble person doing his duty in the face of great temptation. Our moral understanding, like our aesthetic understanding, will be communicable to others in the form of a recommendation, and it will afford a compelling ground for life together that conflicting individual interests do not. The compulsion here is not at bottom that of will, or law, or rule, or consistency. Instead, it is a kind of liking that is free but not simply chosen, and that is regulative for action. It is, then, *our* attitude when we are "mentally attuned" by reason, and no mere submission – even though we

precisely recognize that it is not simply up to us what we make of it. This is the experience of normative authority.

The rule-breaking considerations

Duty belongs to a family of rule- or consistency-based notions. And indeed we typically assume that morally good conduct will follow rules and exhibit consistency. But if Kant is right, then behind these rules – exceptionless, in his system – lies something quite different: a kind of direct liking akin to the experience of the sublime. We do not have rules "all the way down", but must instead encounter a substantive appreciation of value and associated feelings.

Hume was acutely aware of the potential this affords for conflict. If following "the rules of reason" led always to conclusions that substantive evaluation and feeling also embraced, we'd have no difficulty. But at least in epistemology, Hume finds that following the strictest epistemic duties, to accord epistemic respect ("rational credence", we might say) only to conclusions justifiable by reason alone, leads him to an epistemic condition that he cannot find stably credible or genuinely compelling in the guidance of his overall epistemic life. Might the same be true in the moral case?

Consider Kant's discussion of obedience to a tyrannical ruler.

> . . . a people has a duty to put up with even what is held to be an unbearable abuse of supreme authority [since] its resistance to the highest legislation can never be regarded as other than contrary to law . . . . For a people to be authorized to resist, there would have to be a public law permitting it to resist, that is, the highest legislation would have to contain a provision that is not the highest and that makes the people, as subject, by one and the same judgment sovereign over him to whom it is subject. This is self-contradictory . . . . [MM 320][21]

Here Kant appeals to a consideration of consistency to ground a claim of duty. And he has an excellent point, emphasized earlier and in a characteristically different way by Hobbes: a sovereign can benefit us by solving the problem of potentially unending

[21] I am grateful to Tamar Schapiro for bringing this passage to my attention.

social conflict only if our agreement to obey does not contain a clause reserving to each the right to decide on his own authority when to obey.

Hume, likewise, is aware that "the *advantage* we reap from government" will be imperilled if each allows himself to regulate his own obedience in accord with his own ideas of what is just or beneficial. The result could only be "endless confusion, and render all government, in a great measure, ineffectual" (T 555). "We must, therefore, proceed by general rules and regulate ourselves by general interests" (T 555). But how is it possible for advantage-based duty to take on a life of its own?

> . . . there is a principle of human nature, which we have frequently taken notice of, that men are mightily addicted to *general rules*, and that we often carry our maxims beyond those reasons, which first adduc'd us to establish them. . . . It may, therefore, be thought, that in the case of allegiance our moral obligation of duty will not cease, even tho' the natural obligation of interest, which is its cause, has ceas'd . . . . [T 551]

Hume, political conservative that he was, has here a golden opportunity to embrace a Kant-like principle of passive obedience, and even continues "It may be thought that . . . men may be bound by *conscience* to submit to a tyrannical government" (T 551). But he shrinks from this conclusion:

> Those who took up arms against *Dionysus* or *Nero*, or *Philip the second*, have the favour of every reader in the perusal of their history; and nothing but the most violent perversion of common sense can ever lead us to condemn them. 'Tis certain, therefore, that in all our notions of morals we never entertain such an absurdity as that of passive obedience, but make allowances for resistance in the more flagrant instances of tyranny and oppression. [T 552]

How, then, does Hume block the unwanted conclusion of passive obedience? What general rule or practical maxim does he formulate for the citizen to follow to replace the rule of passive obedience? He offers none, only a general suggestion that "the obligation to obedience must cease" when it sufficiently loses its point, that is, "whenever the [common] interest ceases, in any great degree, and in a considerable number of instances" (T 553).

How, then, is this to work? "The common rule requires submission", but "grievous tyranny and oppression" allows individuals to

make "exceptions" (T 554). Here we have a discontinuous change, a departure from own conscientious dispositions to obey which "bind us down", as we rise up in active resistance to government. It looks as if the chief mechanism that awakens us from our "addiction" to general rules is a sympathetic sense of the violation of the general interest. Indeed, sympathy is strong enough that, however much we dislike mayhem and disorder, our approval is excited by rebellions against tyranny of which we hear only in histories or fiction. A morality that would put a people at the mercy of its rulers will not win our wholehearted admiration or esteem. Here we follow no maxim or rule, but a developed sentiment.

It is important to see, however, that the sentiment *is* developed. Self-love and sympathy alone do not yield any comprehension of when a complex political system is abusive or when such abuses have become too considerable. Justly and unjustly inflicted punishment alike look and feel painful; just and unjust war alike are costly and terrifying. An attunement to the general interest calls for complex awareness of cause and effect, and of long- vs. short-term, as well as sympathy for victims. Nonetheless, Hume's account is, in the Kantian sense, heteronomous, since it gives sentiments an essential role, and moreover it yields no strict maxim that individuals could legislate for themselves.[22]

But, stepping back from a model of autonomy as maxim-based self-legislation, if we reflect upon Hume's position on passive obedience vs. Kant's, which of the two, in fact, seems to provide greater practical or political autonomy? Which affords us, as citizens or as moral agents, greater scope to deploy and act on the full range of our human critical faculties?

Suppose Kant were to abandon his egoistic hedonism about human psychology and accept instead the Humean view that sentiments can help us to be attuned to legitimate grounds for moral, aesthetic, or epistemic evaluation. Would he still insist that our only hope for genuinely moral, aesthetic, or epistemic con-

[22] The difficulty of formulating a decision rule to be used by individuals here may be a difficulty *in principle*. Whether it makes sense for you to disobey a tyrant, for example, depends upon whether others will disobey, and their reasoning has a similar dependence upon yours. Problems such as this may admit of general criteria for evaluation (such as a standard of the general interest), but no decision rule or maxim that individuals can self-legislate that would satisfy those criteria. For discussion, see Donald Regan, *Utilitarianism and Cooperation* (Oxford: Clarendon, 1980). More generally, significant limitations of decidability and computability arise for any attempt to give individuals non-self-defeating maxims to guide their conduct in collective settings requiring coordination "autonomously" (in the literal sense – each following his or her own rule).

duct – or autonomy – lies in imposing over sentiment a regime of exceptionless rules?

Of course, I cannot answer on Kant's behalf, but I can attempt this: apply Kant's own test of *fundamental* normative authority, and see where it might lead. How is this possible? Kant's test, recall, involves a special sort of first-personal confirmation: when (for example) we confront the humble man who insists on being honest despite personal costs that we realize would likely overwhelm us, "the mind bows"; when we attend perceptually to sublime scenes in nature, we cannot help but be awed.

Return now to the tyrannical ruler and the obedient citizenry, who accept without resistance all forms of abuse and humiliation. Does "ordinary reason" (G 394) find passive obedience to tyranny sublime – does the mind indeed bow?

I'm willing to bet with Hume that in this case it does not. Impressive as the spectacle may be of passive obedience in the face of great abuse, and powerful as the will must be to restrain an individual feeling the tugs of inclination to strike back at the tyrant, does our mind really bow before this sight? Suppose that the peculiar abuse by government is an order to inform on our friends, to reveal their location to an authority whose plan is to eliminate or torture dissidents or religious minorities. It seems, perhaps, that we know Kant's answer: obey authority; never lie, even to conceal a friend (cf. SRL). And this is the sort of example that has often enough been used by critics of Kant as a *reductio* of his conception of the ground of morality.

But Kant deserves better treatment. Those of us who find in Kant's writings a deep insight into the authority of moral experience should not betray this insight by allowing critics to focus instead on his attempts to apply a multi-layered theory in practice, mediated by a defective empirical psychology. His application may go wrong in cases like "passive obedience", but the fundamentals may yet be sound.

At the fundamental level, I suspect, our mind simply does not bow at the spectacle of the citizen who, despite strong ties of family and friendship, reveals their location to a tyrannical authority. Such an act of will may be monumental, but it is not majestic, and even seems to us peculiarly self-contained or blind. Can we attribute this response on our part to self-interest? No, the response seems to be the same even when we consider a case from history or fiction. Is it then merely an unconsidered reflex? No, Hume is right that our initial reaction to disobedience is usually discomfort. But we reflect

further. The deep normative distress we feel when Germany's greatest moral philosopher defends the unalterable necessity of obedience to the state, and the exceptionless duty never to lie to conceal the location of a friend, is an impersonal and historical shudder. It arises from the full range of Humean faculties, developed through experience: reason, imagination, sense, sympathy, memory, and a feeling for one's place in history.

How different our reaction when we learn that Kant failed on one notable occasion to keep to his habit of regular afternoon walks – the afternoon he received Rousseau's *Emile*, and would not put it down. We might be less impressed by the iron will of Kant upon hearing this story, but we are more impressed by the man and his mind.

Let us conclude with a thought experiment using Kant's own division of the "three different relations that presentations have to the feeling of pleasure", namely, the *agreeable*, the *beautiful*, and the *good*, to understand our reactions and their normative force (CJ 210).

Suppose we had learned that Kant missed his afternoon walk only once, but not to read *Emile* – rather, to avoid a pesky visitor to town whom he knew to be lurking in wait for him with an embarrassing question he preferred not to answer. As a result we might like Kant better – he would be more amiable for showing this human tendency to indulge a desire to avoid an uncomfortable truth. But our self-conceit would not be struck down by this realization – instead, we would find it gratifying to our sense of ourselves that even Kant could be self-indulgent when it comes to allowing oneself to side-step an awkward truth. This we would find *agreeable*, but not in an altogether admiring way. Especially, the critic who finds Kantian moral rigorism excessive would smile inwardly, with perhaps a touch of condescension.

Suppose instead we had learned that he missed his afternoon walk on that one occasion in order to avoid spoiling the end of lovely afternoon tea with a visitor whom Kant rarely saw but personally admired. Then we would like the act, and also Kant, yet better. Moreover, we would like him and his act impersonally as well as personally – for someone to break from routine or personal resolution for such a reason shows a kind of gracefulness or *beauty* of gesture. Even those Kantian critics who find it gratifying to view him as a cold, "clockwork" Prussian would be taken a bit aback, and find a bit of appreciation of Kant creeping in.

But when we learn that in fact Kant missed his afternoon walk but once, in order to continue reading Rousseau's *Emile* – Rousseau! whose unruly mind, scandalous conduct, and colorfully inconsistent prose contrast so sharply with Kant's, but whose insights we know nonetheless reached to the core of Kant's thinking – we like this because it possesses something of the sublime. And we like Kant better, impersonally as well as personally, for showing in a concrete but dramatically appropriate way just how attuned he was to the insights that awaited him in Rousseau, how capable he was of being displaced from the ruts the mind is wont to settle into. We here find in both Kant and his mind something *good*, something estimable in its own right. That afternoon's display of "mental attunement" is much more impressive than would be the strength of will, consistency, or resistance to inclination that Kant would have exhibited had he instead overcome the desire to continue reading *Emile* and maintained above all a resolve to take an afternoon walk each day, exactly at the same time. Thus does Kant's omission strike a bit at the self-conceit of critics who might attempt to look upon him with intellectual condescension as hermetic, narrowly moralistic, trapped within his own technical language and scheme of categories. For when we appreciate this story, we cannot help but feel, freely, a kind of admiration for Kant as an intellect. And thus does the experience of normativity combine force and freedom.[23]

[23] Many colleagues and friends have helped me in developing ideas contained in this essay. Special thanks are due to Elizabeth Anderson, Paul Boghossian, Nomy Arpaly, Stephen Darwall, Allan Gibbard, David Hills, Mark Johnston, David Lewis, Donald Regan, Gideon Rosen, Michael Smith, David Velleman, and Kendall Walton, all whom have tried hard on a number of occasions to straighten out my thinking about normativity. I owe a particular debt to writings on normativity of Allan Gibbard and Christine Korsgaard, who have set out, from their own perspectives, much of the terrain I wander here. A long time ago, Nicholas Sturgeon made me realize I had to rethink Hume. And David Hills and Stephen Darwall deserve special thanks for patience in helping me to engage (insofar as I have!) with Kant's thought. Paul Guyer's writings and correspondence helped me find relevant passages in Kant. Jonathan Dancy gave me very useful comments on an earlier draft, and he and John Cottingham have been exceptionally considerate editors.

2

EXPLAINING NORMATIVITY: ON RATIONALITY AND THE JUSTIFICATION OF REASON[1]

Joseph Raz

Aspects of the world are normative in as much as they or their existence constitute reasons for persons, i.e. grounds which make certain beliefs, moods, emotions, intentions or actions appropriate or inappropriate. Our capacities to perceive and understand how things are, and what response is appropriate to them, and our ability to respond appropriately, make us into persons, i.e. creatures with the ability to direct their own life in accordance with their appreciation of themselves and their environment, and of the reasons with which, given how they are, the world presents them.

An explanation of normativity would explain the various puzzling aspects of this complex phenomenon. In particular it would explain how it is that aspects of the world can constitute reasons for cognitive, emotive, and volitional responses; how it is that we can come to realise that certain cognitive, emotional or volitional responses are appropriate in various circumstances, and inappropriate in others; and how it is that we can respond appropriately. This paper explores an aspect of the last of these questions.

1. Normativity and rationality

The normativity of all that is normative consists in the way it is, or provides, or is otherwise related to reasons. The normativity of rules, or of authority, or of morality, for example, consists in the fact that rules are reasons of a special kind, the fact that directives issued by legitimate authorities are reasons, and in the

[1] The first version of section one was included in papers presented at the Philosophy Colloquium at Berkeley, at the conference in memory of Jean Hampton at Tucson, and at a conference on practical reason in the Humbolt University, Berlin. I learnt from questions and comments of many who participated in those occasions. I owe a special debt to David Silver, who was the commentator on my paper at Tucson, and to Jonathan Dancy for most helpful comments on a later version of the paper.

fact that moral considerations are valid reasons. So ultimately the explanation of normativity is the explanation of what it is to be a reason, and of related puzzles about reasons.

Reason[2] is inherently normative. That is its central characteristic. Therefore, the accounts of normativity and of reason and rationality, though not identical, are inter-related. An account of rationality is an account of the capacity to perceive reasons and to conform to them, and of different forms of conforming to reasons, and their appropriateness in different contexts. To explain the capacity to conform to reason the account must explain the possibility of error, failure to perceive reasons correctly, and of failure to respond to them once perceived. An account of irrationality is an account of some of the ways of failing to conform to reason, those which render one, or one's behaviour, or emotions, etc. irrational. The core idea is that rationality is the ability to realise the normative significance of the normative features of the world, and the ability to respond accordingly.

In one sense of "rational", we, or anything else, are rational beings to the extent that we possess that ability, which I will call "capacity-rationality". The absence of capacity-rationality does not mean that a creature is irrational. It means that no judgements of rationality apply to that creature. In another sense, we, or anything else, are rational to the extent that in general we use that ability well. There is a further use of "rationality" in which it applies to specific human responses, or their absence. Our actions, intentions, beliefs, emotions, etc. can be severally rational, non-rational, or irrational.

Rationality makes us into persons. To yield an explanation of rationality the core idea has, of course, to be hedged and refined. I will confine myself to four observations which help locate the relations between capacity-rationality and normativity.

First, to be people who are rational in the first sense, that is to be people with the ability to perceive reasons and respond to them, we need a range of capacities which do not directly contribute to our rationality. They include some perceptual ability, and the capacity to control our movements at will. An impairment of our perceptual ability does not diminish our rationality. Nor does lack of muscular control, or other neurological or

[2] I will be using "reason" when it sounds natural to use it. The expression refers variously to all or some of the following: people's reasoning capacity, people's use of that capacity when referred to in a general way, the reasons which apply to people on the occasion(s) discussed. I will sometimes leave it to context to identify the meaning.

physical impairments of our ability to move at will. Possession of at least some perceptual ability, and of some ability to control one's movements at will, are presupposed by capacity-rationality.[3] But they are not themselves constituents of rationality.

There are other preconditions of capacity-rationality. Among them are psychological capacities, such as the possession of memory, the capacity for conceptual thought, and the capacities to form beliefs, and to reach decisions. Some of the preconditions relate to capacities possession of which is a precondition of being either rational or irrational, but where their successful exercise does not render one (or one's judgement, action, etc.) rational, nor does their failure render one irrational. Perceptual failures, muscular failures, failures of memory are examples of preconditions of this kind. The abilities to form beliefs and to reach decisions belong to the other kind. Success and failure in their exercise contribute to the evaluation of the rationality of one's beliefs, actions, etc., as well as to the evaluation of the rationality (in the second sense) of the person concerned.

Not surprisingly the boundary between those who do not meet the preconditions of capacity-rationality, and are neither rational nor irrational, and those who meet the preconditions, and are irrational, is not a sharp one. Someone who lacks a minimal capacity to make up his mind about anything fails the preconditions and is neither rational nor irrational. He lacks capacity-rationality. But those who have a minimal ability to make up their minds, and constantly vacillate, finding it always difficult to make up their minds, and almost impossible not to change almost every decision they take soon after taking it, are irrational.

What is capacity-rationality in itself? It is a capacity to see the normative significance of the way things are, to comprehend what reasons they constitute, and the significance of that fact for oneself.[4] As indicated at the outset, part of the explanation of

[3] My claim is that their possession to some degree, during some part of the creature's life, is necessary for that creature to have capacity rationality at all, not merely for his ability to display that capacity or exercise it. It is true, though, that beyond that minimum lacking the ancillary capacities may impede the exercise of rationality, or make it temporarily impossible.

[4] Needless to say possession of this capacity itself is a matter of degree. Some humans and other creatures can recognise (the normative aspect of) some types of reasons but not of others. Some humans and other creatures can reflect about the fact that the existence of certain facts constitutes reasons, and form general views about the nature of reasons and rationality. Others can do so to a limited degree, or cannot do so at all. There is little point in trying to fix a test of personhood which will endow the concept with relatively sharp boundaries, which it does not possess.

normativity consists in the explanation of this capacity. Here again we encounter the overlap between the explanations of normativity and of rationality.

These remarks lead to my *second* point. Some accounts of rationality identify it with a reasoning ability. That ability consists, at least in part, in the ability to recognise inferential relations. That is it consists of, or includes, an ability to figure out what conclusions follow from given premises (the active side of the capacity), and an ability to recognise that inferences are valid or invalid[5] when this is pointed out to one (the passive side of the capacity). While at least a minimal reasoning ability is among the constituents of capacity-rationality, it does not exhaust it, at least not if the capacity to reason is the ability to construct and understand inferences.

That a certain proposition follows from certain premises is, other things being equal,[6] a reason for not believing the premises without believing the conclusion.[7] Given that rationality is the capacity to realise the normative significance of facts, that is to realise whether they constitute reasons, and which reasons they constitute, and to respond appropriately, the capacity to see the normative significance of inferences is high among the constituents of capacity-rationality. But what of reasoning power in general? One possibility is that the capacity to realise that C follows from P is like the capacity to see that the house is on fire. They are ways of realising the existence of facts which constitute reasons, and their possession is a precondition of being able to become aware of the normative significance of those facts, and to respond appropriately. On this account neither perceptual capacity nor reasoning ability are themselves constituents of

[5] I am using these terms to designate the success and failure of any inference, not merely of deductive ones.

[6] The existence of defeasible valid inferences requires the qualification "other things being equal". Typically when the inference is defeated it yields no reason at all.

[7] It is not, however, not even when the inference is a deductive one, a conclusive reason. Mere knowledge that a set of propositions is self-contradictory is a reason, if one does believe in at least some of them, to refrain from believing in one or more of them so that one's beliefs will not be self-contradictory. When one has no further information about the location of the contradiction and the reasons for it, the only way to know that one conformed with this reason is to refrain from believing any of the propositions in the set. When the contradictory set is large the price of doing so can be very great. Sometimes, it may be impossible to conform with this reason (e.g. if the set includes all one's beliefs). But even when it is possible it may be unjustified, given the price. After all, knowledge that the set is contradictory is no more than knowledge that one of its propositions is false. We know on inductive grounds, that even if our beliefs are consistent at least one of them is false. That is no (adequate) reason to suspend all our beliefs. Why should the fact that our beliefs are contradictory be such a reason?

rationality. They are merely preconditions of its exercise. Alternatively, it is arguable that reasoning is unlike perception in that it is involved in almost any recognition of the normative significance of anything. Even recognising that since C follows from P one has, other thing being equal, reason not to believe P and reject C involves reasoning. The close involvement of reasoning in capacity-rationality would justify, on this view, regarding the ability to reason as a constituent of rationality.

There is no doubt that reasoning ability is closely involved with rationality, even though the pervasiveness of its involvement may be debated. Yet it is doubtful that this involvement would in itself justify the identification of reasoning ability with capacity-rationality. There is, however, a better argument for that conclusion: realising that C follows from P is not merely realising the existence of a fact which is a reason (as in seeing that the house is on fire). It is realising that there is a reason. Understanding that C follows from P is the same as, or at least involves, understanding that, other things being equal, one has reason not to believe P and reject C. That is part of what it means that the one follows from the other.

This argument notwithstanding, this seems to be a case where our concepts are not as neat and tidy as philosophers may wish them to be. While possession of a minimal reasoning capacity is a constituent of capacity-rationality, that ability may be modest. In general, failures of our reasoning powers do not warrant a judgement of irrationality. Failure to see that a conclusion follows from certain premises most commonly merely shows that one is not very bright, or just not very good at reasoning. In most cases it does not show that one is irrational, as failure to realise that if a conclusion follows from premises then one has reason not to believe the premises while rejecting the conclusion does. On the other hand, failure of *elementary* reasoning does establish irrationality: Failure to realise, in normal circumstances, that it follows from the fact that one's destination is not far away that it will not take long to get there, is a failure of rationality.

It follows that there is a certain distance between capacity-rationality and reasoning ability. Good reasoners can be habitually irrational, and, more commonly, perfectly rational people can be bad reasoners. They often make mistakes, but that does not impugn their rationality.[8]

[8] The point discussed in the text helps in dissolving an apparent asymmetry between

The third point was anticipated in the previous remarks, and helps illustrate them. Our rationality, I claimed, consists in the ability to recognise the normativity of features of the world. That ability expresses itself in the proper functioning, in relevant respects, of our faculties. The point I am striving to make is that our rationality expresses itself not only in our deliberation and reasoning, nor in any other specific act or activity, but more widely in the way we function, in so far as that functioning is, or should be, responsive to reasons. Take a simple example. I mentioned above that habitual failure to take decisions, or to form judgements where they are called for, is a form of irrationality. Imagine that whenever Sylvia leaves her home she locks her front door behind her, walks to the gate, turns back, goes to the door and, to check that it is really locked, unlocks and relocks it, goes out of the gate and immediately turns back, goes back to the door and repeats the action, and so on several times. This form of indecisiveness amounts to irrationality. At some level Sylvia probably recognises that her actions are irrational. But at the same time she is seized by anxiety and doubt. Maybe the key did not turn full circle and did not lock the door, maybe she imagined feeling the bolt move, and in fact it did not, maybe when she unlocked the door to test what she did before she forgot to relock it.

We are all familiar with mild forms of such anxiety. Barristers are supposed to capitalise on it by undermining witnesses' confidence that they really saw what they saw, that they really remember what they remember. At the same time we recognise that the capacity for such self-doubt is itself an aspect of our rationality. It manifests our ability to monitor (mostly below the level of awareness) our reactions and their appropriateness to the circumstances, that is to monitor our responsiveness to reason. But to act rationally we need to preserve a proper balance between resoluteness and openness to doubt. It is not a balance we can decide upon. While Sylvia can grit her teeth and decide not to turn back again for the fifth time, such conscious decisions only minimise her irrationality. To be rational she must act 'automatically' in a way which allows for a

theoretical and practical reasons. How is it, one may ask, that practical (and aesthetic) reasons are normative on their face, being facts like: 'Doing A will give you pleasure', or 'you promised to do A', whereas anything can be a reason for belief? That there are clouds and high winds is a reason to believe that it will rain. But there is nothing inherently normative in the fact that there are clouds and high winds. The explanation is that the fact that there are clouds is no reason to believe anything, though that it follows from the fact that there are clouds that rain is likely is a reason, and is normative on its face.

proper openness to doubt without relapsing into indecisiveness and anxiety. To be rational she, and we, must function properly, and that functioning must be automatic, rather than a product of deliberation and decision.

The fourth and last point to be made here about rationality is by now obvious. I am treating rationality as a unified concept, designating a unified capacity, which straddles the divides between practical and theoretical rationality, as well as between procedural and substantive rationality, and others. Some writers believe that there are two concepts of rationality in use, and presumably they designate two different abilities. Parfit sets the distinction thus: "To be substantively rational, we must care about certain things, such as our own well-being." "To be procedurally rational, we must deliberate in certain ways, but we are not required to have any particular desires or aims, such as concern for our own well-being."[9] He does not elaborate. What could he mean? It is possible for a person[10] irrationally to fail to have desires or intentions or goals which he has adequate reason to have. It is also possible for someone who generally has the goals which rationally he should have to display irrationality often when deliberating. There is no reason to think that failures of rationality are randomly distributed among the different occasions on which rationality is called for. For example, motivated irrationality, such as self-deception, would manifest itself selectively on occasions where its existence would serve its underlying motive. Parfit's distinction may, therefore, be understood as a distinction between success in being rational in different aspects of one's life. But the context makes clear that he does not mean it in this way. He seems to think that there are two different notions of rationality, each designating a different capacity. His discussion raises the possibility that possession of the rationality identified by one of these notions is independent of possession of the rationality marked by the other. Whether or not this is Parfit's meaning it is a common view, and a mistaken one.

The division between substantive and procedural rationality (and between substantive and instrumental rationality) took hold among philosophers who doubted that reason is directly involved in the choice of ends, but believed that reason has a role to play

[9] D. Parfit, "Reason and Motivation" *Aristotelian Society*, supp. Vol. 71 (1997) 99, at 101.

[10] By their nature persons are rational in the sense of having capacity-rationality.

in practical thought, which the notion of procedural rationality captures. Some allowed that we talk of rationality in a more extensive sense, and claimed that that is merely an ambiguity in the meaning of 'reason' and 'rationality', perhaps resulting from those primitive times when people believed that rationality affects the choice of ends as well. In fact reason affects our choice of ends and the desires we have just as much as it affects our deliberations and our beliefs. We cannot have a desire except for a reason.[11] Once that is allowed the motivation for the division of rationality into two distinct capacities disappears. There is no reason for thinking that the capacities which enable us to discern and respond to reasons for desires are different from those which enable us to discern and respond to reasons for belief. One may well need some special abilities (discriminating eyesight, or palate, a good ear, or a capacity for empathy etc.) to be able to discern various non-instrumental reasons. But these, while presupposed by capacity rationality, or by the ability to use it, are not identical with it, nor are they constituent elements of it. They do not show that there is more than one concept of rationality.

It is not clear what Parfit means to include in "to deliberate in a certain way". The power of reasoning in and of itself does not establish any degree of rationality. Imagine a person who as a hobby picks on arbitrary collections of propositions, perhaps chosen randomly from newspapers and magazines, and works out various other propositions which follow from them. Putting on one side the suspicion that the very activity displays irrationality, does he show himself to be good at procedural rationality? I do not think so. Our person is good at recognising the existence of inferential relations. He is not changing his beliefs in any way, not even conditionally (i.e. the story does not assume that he acquires beliefs of the sort: were I to believe these premises I would have a reason to accept these conclusions). For all we know he does not understand the normative significance of inferences. Perhaps procedural rationality is meant to include more than such reasoning. There may be no obstacle to enriching the notion to include steadfastness of resolution and other aspects of proper functioning of the kind alluded to above. The question is whether once enriched the capacities included under the head-

[11] Though occasionally people have urges which are unreasoned. I have argued to this conclusion in "Incommensurability and Agency" in *Incommensurability, Incomparability, and Practical Reason* ed. R. Chang (Cambridge, Mass.: Harvard UP, 1998).

ing of "procedural rationality" will be different from the capacities which make one so-called substantively rational.

To repeat: to those who believe that no rational capacities are involved in persons having goals the answer is obvious. But since we desire only what we think of as worth desiring, our desires are among our responses to perceived reasons. This is true of a desire to drink when thirsty as much as of the desire to become a good teacher. In conceiving desires and in adopting and maintaining goals we deploy all the capacities which are involved in so-called procedural rationality, and there seems to be no other capacity involved, at least none which can relate to rationality. The reasoning ability and other capacities which make people rational in forming beliefs about scientific matters, or about the weather, or anything else which can be said not to be in itself normative, are the same abilities which make people rational in the way they adopt and maintain goals. Therefore, there is only one kind of rationality.

As was allowed above there may be factors which may interfere with the display of rationality in one area but not in another. But these will not be enough to establish complete independence of one's success in being rational in the different domains, and they will not track the supposed distinction between procedural and substantive rationality. Typical examples of such selective failures of rationality are those occasioned by motivated irrationality. They affect one's thoughts, feelings and goals only when the motive leading to the irrationality comes into play. But they can affect one's choice of goals as well as one's 'theoretical' beliefs.

These remarks on the nature of rationality and its relation to normativity form the backdrop to the discussion that follows. They concern capacity-rationality, and though they inevitably have implications regarding the meaning of judgements that this or that action, desire, emotion, attitude, belief, etc. is rational or irrational, these implications are not straightforward. As they do not affect the rest of the argument they need not be explored here, beyond one point. Obviously to judge a belief, desire, emotion etc. as rational is to note that having them is at the very least consistent with a successful deployment of our capacity for rationality.[12] The standard by which success is to be measured is far from

[12] "This was a rational belief for X to have" may mean no more than that. "He rationally came to the belief that ..." indicates much more. It presupposes actual use of one's rational capacity, and asserts that that has been successful. There are many variations and nuances in attributions of rationality and irrationality to beliefs, etc.

clear. It is doubtful that there is only one standard employed on all occasions. It seems likely that we recognise a range of standards, and on each occasion we implicitly employ the one we find fitting in the circumstances. The most demanding standard regards as rational only those beliefs, etc. which are consistent with complete identification of all the reasons which apply to the situation[13] and a perfectly appropriate response to them. Everything else is judged irrational.[14] This standard seems to be used only rarely. A very lenient standard regards any belief to be rational unless holding it results from a failure (successfully) to employ one's rational capacity which involves gross mistakes, etc. Some standards relativise to one's age (what would be rational for a child to believe may be irrational for an adult), educational background (given that you have a degree in physics it is irrational for you to believe that), position in life, or to some other social variable. Common are standards which tie irrationality to blame. One's beliefs are irrational when one is blameworthy for having them. The best of those employ something like the legal test of negligence: a belief is irrational if and only if holding it displays lack of care and diligence in one's epistemic conduct.

2. The centrality of the ability to reason

When studying reasons we study normative aspects of the world. When discussing rationality we discuss our perceptions of, and responses to, reasons. Our ability to reason is central to our rationality in all its manifestations, that is regarding reasons for belief, action, emotion, or anything else.

One way to bring out the point is by reminding ourselves that emotions, attitudes, desires, and intentions have a cognitive content and cognitive presuppositions. Their rationality depends, in part, on the rationality of the beliefs which contribute to them, and which are presupposed by them. This dependence is asymmetric. Beliefs do not depend on feelings, desires or intentions in the same manner. The rationality of beliefs enjoys a certain primacy for being involved with other forms of rationality, which it does not presuppose.

[13] Even this demanding standard allows that epistemic reasons vary with context. Therefore, even according to it one's beliefs may be rational and false.

[14] There is no reason for the standard for irrationality to be the contradictory of the standard for being rational. The two can be logically independent, allowing for beliefs which are neither rational nor irrational.

More specifically, at least a rudimentary reasoning ability is involved in all rational responses to reason, simply because they are responses under the control of the agent, though admittedly I am using "responses" loosely here, to indicate that rationality depends on appropriateness or intelligibility in the circumstances. Emotions, desires, intentions, or beliefs are rational depending on (a) whether they belong to a rational agent; (b) whether their occurrence is under the control of the agent; and (c) whether they are appropriate or intelligible given the reasons for and against them, as these reasons are, or as they are reasonably perceived by the agent.[15]

It is worth stopping to comment on the second of these conditions. It embodies several of the points belaboured in the previous section. I will address my remarks to the case of rationality in the endorsement of beliefs, but they apply with minor modifications to rationality in one's intentions, emotions, decisions, actions, etc. Control consists in the proper functioning of a person's rational faculty, the proper functioning of the person's ability to recognise and respond to reasons, rather than in any particular performance, such as an action, or a deliberate decision. Control is manifested when a belief is adopted, or endorsed in a process in which the ability to recognise reasons and respond to them (proper appreciation of and response to perception, or to testimony, for example) is active. It can be active even when beliefs are formed without deliberation or awareness, but when the agent's critical faculties would have stopped their formation had they been rationally suspect.

Is it not a distortion to claim, as I just did, that the proper functioning of a faculty, even of rationality, depends on being in control of our emotions, actions, beliefs, and the will? After all the central use of "control" relates to exercise of the agent's will: we control what we do, and how we do it, and we control our emotions, to the extent that we intentionally hold our emotions in check by an effort of will, etc. But there is a wider, more basic, use of control. People's being out of control means that their will itself is not under their control, and our will and beliefs are out of control when they are systematically irrational.

Capacity rationality is a more fundamental capacity than the will, which is the capacity for intentional action, for forming

[15] This point is subject to the baseline question discussed above.

intentions and taking decisions. Rationality is like dispositional abilities, that is abilities which are manifested when the circumstances are right. (E.g., a rope which can take a 100 kg weight is a rope which will take weights up to 100 kg. without snapping, unless . . .). Our fundamental psychological abilities are of this kind, except that they are subject to complex possible interferences, many of them due to psychological factors: memory can fail when people are depressed, etc. Rationality is like that: it is the ability to respond appropriately to (perceived) normative aspects of the world, and this means that rational beings respond appropriately to perceived normative aspects of the world, when no failure of attention, emotional upset, mood, memory, will, etc. interferes.

Like other dispositional capacities rationality is a capacity which displays itself when the occasion presents itself, so long as no distorting factors interfere. That is what the second condition for the rationality of beliefs, actions, etc. signifies. People's beliefs (and again these remarks apply with some changes to other objects of our rationality) are rational only if they are formed and maintained while the people involved are in control of their formation, and continued endorsement. That is, while their reason controls how they come to adopt or endorse their beliefs. Reason controls the formation and endorsement of beliefs when, whether or not their formation or endorsement involves deliberations, beliefs are formed in processes which stop people from having them when their formation or endorsement is not warranted by reasons, as the agents see them, given their understanding of the situation they are in. A different aspect of control (and all these are matters of degree) is exercised when reason makes people endorse propositions which they are aware of compelling or at least adequate reasons to endorse.

Some people will say that when reason is in control we engage in subconscious reasoning. Whether or not such claims can be vindicated, the exercise of our rationality must be represented in the same way that reasoning is represented. The exercise of reason which manifests its control over our beliefs, emotions, intentions, desires, etc. is subject to the same rules that govern explicit reasoning. To that extent capacity for reasoning is central to rationality and is involved in all its manifestations. This raises the question of the standing of principles of reasoning.

3. Can reason be justified?

In setting out (at the beginning of this essay) the contours of the problems of explaining the nature of normativity I did not once refer to the justification of normativity. Is not that the primary task of a theory of normativity or of reason?

It is not easy to make sense of the very quest for the justification of normativity. We can ask whether this fact or that is a cogent reason for action or belief, etc. We can raise more general questions about types of facts: For example: does the law (i.e. the fact that one is legally required to perform an action) constitute a binding reason for action? Do people have good reason to conform to the practices of their country? But what is it to justify reason as such? Presumably the question is whether we are ever justified in holding anything as constituting an "objective" reason? Or, whether it is possible for anything to be a reason? Or, whether there are any facts which are reasons?[16]

These explanations of the quest for the justification of normativity do not, however, dispose altogether of the difficulty in understanding it. Even once anxieties about the objectivity of reasons are put on one side questions remain. They are not dissipated by looking for the justification of statements that this or that is a reason for action or belief. Such justifications take many shapes and forms: burning the cat would be cruel (and that is a reason not to burn it) because it would inflict gratuitous suffering, and so on. These are the mundane arguments for this reason or that, which we are all familiar with. The quest for the justification of normativity cannot be whittled down to the normal arguments for the truth of a statement about this reason or that. It is a search for the vindication of the methods of reasoning employed in such mundane arguments, or for the discovery of a super principle which justifies confidence in the whole enterprise of reason, the whole enterprise of discerning reasons and responding to them. Is that a meaningful and a sensible quest?

One common reply is that every argument to debunk reason would be self-defeating, for it will have to use reason and thus its own validity depends on the assumption that it seeks to challenge,

[16] Alternatively, perhaps the question meant is whether it makes sense to talk of anything being a reason? But this question comes very close to the explanatory questions I mentioned in the previous section. They explore what is the sense of normative discourse. I for one find little reason to doubt that normative discourse is meaningful.

i.e. that reason is justified. Recently Tom Nagel has advanced a whole array of arguments in support of the objectivity, universality and reality of reason (these are his terms) including a version of the argument that challenges to reason are self-defeating.[17] Nagel's argument rests on two pillars. First that "one cannot criticise something with nothing".[18] The second is the fact that we cannot escape relying on reason.[19] Both pillars are suspect.

Reductio ad absurdum is a familiar form of argument which, at least prima facie, need not presuppose anything. It refutes a supposition by deriving a contradiction from it. As presented *reductio* arguments quite commonly presuppose premises and rules or methods of argumentation, which are accepted as uncontroversial. But need this be the case? Is there some reason why one must rely on some premises, rules or methods of argument other than those which are refuted by the *reductio*? I do not think so. In a way no *reductio* argument does. They all take the following form: Using rules (or methods) of argument R_1 to R_n a contradiction can be derived from premises P_1 to P_m. Therefore, at least one of the premises or at least one of the rules of inference is false or invalid. As I said, commonly in such arguments all but one or a few premises are accepted as being true, hence the conclusion is that at least one of the remaining ones is false. But strictly speaking such arguments do no more than impugn one of the premises and rules of inference. Of course, *reductio* arguments use rules of inference, not least in their final step: All these premises and rules yield a contradiction. So at least one of the premises is false or at least one of the rules invalid. That does not matter to my point, since the rules of inference relied upon are themselves put in doubt by the argument. Nothing is relied upon without being cast into doubt at the end.

To use *reductio* to challenge the validity of reason one would reduce the number of premises to the minimum, including only logical truths or indubitable other truths, and use only the most basic rules of inference. If *they* give rise to a contradiction then it

[17] In chapters 2 and 4 of *The Last Word* (New York: O.U.P. 1996)

[18] *Ibid.* p. 20.

[19] Nagel is, of course, aware of the fact that the inescapability of a belief in the objectivity of reason is no proof of its correctness: see. p. 33. He rests his case on the claim that no sceptical conclusion follows from the fact that explanation must come to an end, and that "the language, and the truth of some other form of subjectivism is not shown by the fact that justification comes to an end at certain points at which there is natural agreement in judgements. Nothing about the framework of thought is shown by these facts". (p. 34). I agree with these propositions. They are consistent with the less sanguine view expressed in the main text.

follows that at least one of them is false or invalid. Would not such a conclusion be tantamount to a sceptical argument against reason itself? For example suppose that using substitution and modus ponens only, one derives a contradiction from (x) x = x. Could not such a proof lay claim to being a refutation of the cogency of reason?

Moreover, if we reject Nagel's claim that "one cannot criticise something with nothing", then the fact that we cannot escape relying on reason may not be a very powerful point in its defence. It may turn out to be the case that while we cannot avoid relying on reason, we know that we are doomed to rely on an incoherent system of thought.

From the fact that if there is such a radical *reductio* argument it will refute the validity of reason and the value of rationality it does not follow that they stand refuted. One can perhaps conclude (a) that Nagel, and others, have failed to produce general reasons for the impossibility of a successful sceptical argument against reason, and yet hold (b) that no such argument exists, and therefore that we have no reason to doubt the cogency of reason. But that view does not seem to be quite adequate to the situation.

First, it is not entirely true to say that no challenges of the kind described exist. Paradoxes, some old some new, such as the liar paradox, or Zeno's paradoxes, or the sorites paradox, have puzzled thinkers. Such paradoxes seem to be of the required kind. If not singly then cumulatively they challenge the coherence of reason. They take assumptions that lie at the very foundation of our conceptual thought and reduce them to a contradiction. They challenge the coherence of the concept of truth, the concepts of change and of time, and of the concepts of identity, of objects and of possessing properties. It will take a brave man to say that they were all solved successfully. But we need not argue about that. What is of interest is that for long stretches of time people did not know how to solve them, and knew that they did not know how to solve them. Yet they carried on regardless. Rational thought did not stop, was not abandoned, in spite of awareness that, for all one knew, there were contradictions at its foundations. I do not know of any serious, let alone successful, argument that that was irrational, that it was irrational of people to carry on using reason, in spite of being aware of unresolved paradoxes concerning its basic features.

How are we to understand this reaction? I am less interested in its historical explanation than in the question whether it is a

rational reaction on general grounds, independent of the specific historical context. One possible explanation is that even if the use of reason is incoherent and self-contradictory, we are condemned to carry on using it. To be sure we can abandon reason, but we cannot reason our way into doing so. We abandon reason, or it abandons us, when we suffer brain damage, stupefy ourselves with hallucinatory and other drugs, etc. People can take action deliberately in order to be rational no more, but they cannot get there simply by reasoning their way into scepticism about reason. But that answer is unsatisfactory. If we can abandon reason, or cause it to abandon us, why is it not the case that exposing the contradictions at its foundations constitutes an argument for doing so? One cannot answer that such an argument relies on reason. It relies only on the fact that reason is self-defeating and self-contradictory.

One may point out that the step from the proposition that reason is self-defeating, and self-contradictory, or simply that it is unjustified, to "let's blow our brains out" relies on the rationality that has just been defeated. Perhaps the bankruptcy of reason cannot be a reason for abandoning reason. But at least it means that there is no reason not to do so. We cannot conclude that we may do this if we want to, if that means that we have reason to believe that abandoning reason is permissible. But we can hold that we may do so, meaning simply that there is nothing against doing so – no reason against doing so, since there are no reasons for anything. That conclusion is devastating enough. It does not leave things in equilibrium. It is not neutral as between abandoning reason and not doing so. For, if the debunking arguments are successful, then to choose to continue to rely on reason is either to choose to be self-deluding, or to choose arbitrarily with open eyes to follow a debunked mode of life. It is to act against the spirit of the action. It is to rely on a method when by the light of that method itself it should not be relied upon. That does not establish a reason for not relying on it, but it places reliance psychologically at a disadvantage.

4. Arguing against substantive principles of reason[20]

We cannot reason ourselves into abandoning faith in reason not because it is impossible for us to abandon faith in reason as a

[20] The considerations advanced below are presented informally. I assume an inter-

result of a successful argument that we should do so, but because no such argument can be sound. *Reductio* arguments of the kind I mentioned do not constitute a *reductio* of rationality. It is useful here to distinguish between the formal notion of reason, and substantive doctrines about the nature of reason. The formal notion of reason is fixed by the very abstract and essential characteristics that mark the kind of thinking which is governed by reason. We have no word for it other than "thinking". It is, however, thinking in a narrow sense, the sort of thinking that we refer to when saying: "wait a minute. I am trying to think", and not the sort of thinking that day dreaming, free association, fantasising, etc. are.[21] The formal notion of reason singles out a type of thinking which is marked by the discipline it is subjected to, a discipline which enables one to distinguish instances of successful thinking, from flawed thinking. Possibly thinking in the narrow sense can be adequately characterised by two central essential properties:

(a) It is thinking which is subject to evaluation as correct or incorrect.

(b) The standards by which success of episodes of thinking is judged depend on the reliability of the process of thinking which meets them in yielding justified intentions, decisions and beliefs, that is ones which are adequate, given the normative aspects of the world.[22]

In short reason is a discipline which governs thinking, or a type of thinking. Thinking in conformity with reason is successful thinking, and thinking which does not conform to reason is unsuccessful as an instance of thinking of that kind. The substantive doctrine of reason spells out the content of that discipline. Rules of inference (deductive and non-deductive alike) and the central concepts they depend on are a central part of the substantive doctrine of reason.

dependence of words, meanings, and conkcepts, but do not clarify it. Nor do I try to make more precise the notion of a principle of reasoning. Given a generous understanding of the notion not all principles of reasoning are constitutive of meanings or concepts. That does not affect the case put here which is sound as applied to those principles of reasoning which are constitutive of meanings and concepts. If anything, the contingency of other principles of reasoning is even easier to establish, but I will not consider them here.

[21] Though other forms of thinking may borrow parts of the discipline of reason which marks the narrow notion of thinking.

[22] Notice that not all thinking aims at justifying beliefs or intentions. It can be part of the telling of invented stories, fanciful imagining, etc. It is, however, thinking of the relevant kind if it is governed by standards the use of which can justify intentions and beliefs.

I will argue that the paradoxes of reason cast doubt on the substantive doctrine of reason, on the cogency of the concepts and rules of inference it employs, not on the possibility of reason in itself. That is they do not impugn formal reason. Do they cast doubt on the way we reason? Or on our understanding of the way we reason, on our theory of reason? Does the substantive doctrine of reason I referred to consist of the principles which govern our reason or our attempts at an explicit articulation of those principles?

They may do either. In any case the two are not entirely separate. Once we develop explicit accounts of the nature of reason, of its substantive doctrines, those accounts, while they never exhaust the forms of reasoning we engage in, do affect how those who are more or less aware of them and accept them, reason. Solutions to the liar paradox cannot be said simply to improve our understanding of how we reason all along. They change our reasoning practices, modify the rules used in them, or restrict their application. And the same goes for solutions to most other logical paradoxes. This is particularly likely to be true of the ancient paradoxes. Regarding them it is unlikely that the persistent failure of attempts to solve them is due merely to misunderstanding of how we actually reason, and therefore unlikely that solutions point merely to misunderstandings rather than to the need for a change in the concepts which generate the paradoxes.

If I am right then reason, i.e. the doctrines of reason, can be successfully challenged, and we respond to such challenges by modifying it, modifying reasoning practices and the principles which govern them. I am not suggesting that they are modified only in response to paradoxes and sceptical challenges. However, so long as they do not bring new paradoxes in their wake, modifications which free our practices from paradoxes constitute advances or improvement in our reasoning practices, and in the principles of reason we use.

On this view reasoning principles are social principles, evolving roughly in the ways in which social practices generally evolve. But they can also be challenged on grounds of incoherence, or unreliability. They can be improved in response to such challenges. In this picture, changes in logical and conceptual principles of reasoning parallel changes in inductive clues and scientific methods of experimentation. Even though changes in principles of reasoning involving conceptual shifts occur mostly in

informal ways and are harder to document, they respond to pressures similar to, though more diffuse than, those affecting scientific methods.

One reason for viewing with suspicion the claim that reasoning principles are based on practices which evolve in ways similar to that of other practices which govern our life is the familiar philosophical doctrine that regards changes in reasoning practices as nothing but corrections of mistakes. Such corrections occur when the changed practices comply more closely with the universal principles of reason. This objection admits that not only the explicit articulations of principles of reasoning can be faulted and can change, but that actual reasoning practices can be affected by contradictions and paradoxes, and can change to avoid them. But the objection denies that that process should be regarded as a process of change in *the substantive principles of reason.* They are universal and timeless. The changes we observe are in our imperfect attempts to conform to them.

The objection presupposes, however, that there is one and only one set of correct or valid principles of reason. This seems to me implausible, or at least in need of qualification. Principles of reasoning and of meaning and reference are, of course, intimately related (i.e. if we abandon a principle of reason – e.g., the excluded middle – we change the meaning or content of the propositions which are governed by these principles of reasoning). Therefore, if there can be systems of concepts such that (a) none of them is better than any of the others, and (b) there is no possible system of concepts and reasoning which is better than they are, then there is more than one ideal or correct system of concepts and reasoning, ideal in that they cannot be improved upon. Systems meeting these conditions also meet the further condition, namely that each includes concepts which are not part of the other. Therefore, such systems are incommensurate.[23]

The mere possibility of incommensurability among systems of reasoning (expressed in different languages or segments of lan-

[23] My invocation of meaning incommensurability carries no sceptical implications. I do not claim that those who understand a system of concepts and reasoning (or, for that matter, a theory) cannot understand others which are incommensurate with it. Nor do I believe that incommensurability implies the possibility of incompatible truths, each vindicated within its own system or language, and each refuting the other. The possibility of incommensurate systems of concepts presupposes that truths that can be stated in one are compatible with truths which can only be stated in the other, though the tests of compatibility may involve extending the range of concepts in one or both to make sense of the very notion of compatibility and incompatibility between incommensurate propositions.

guages) does not cast doubt on the universality and timelessness of the principles of reasoning. However, once that possibility is admitted it becomes difficult to resist the thought that there is indeed an indefinite number of incommensurate systems of reasoning, and incommensurate languages or segments of languages expressing them, each with its own ideal or correct principles of reason. The number and identity of the historically instantiated systems of concepts, and of the rules of inference and reasoning associated with them, are matters of historical contingency. Moreover, it is possible for more than one correct system to be instantiated. That makes it hard to deny not only that the historical instantiation of principle of reasoning is a matter of contingent fact, but that the principles themselves are historical products emerging at particular points in time. The alternative is to assume the existence of an infinite number of sets of principles of reason, most of which it is impossible for us to find out about until such time, if ever, when the related new language evolves.

Let it be granted that ideal or correct principles of reason are historical products, that is, practices which arise in time: must we also concede that practices riddled with paradoxes embody principles of reason, rather than that they embody mistaken principles which are accepted as if they were principles of reason, but which are not? Admittedly it is puzzling to think that there can be paradox-ridden principles of reason. Adding that they are imperfect or defective principles of reason, rather than removing the oddity of this view, adds to it. Is it not a contradiction in terms to think of imperfect or defective principles of reason?

Nevertheless, we must accept that this is so, for the alternative is unacceptable. The first and most radical way of understanding the situation of people whose practices of reasoning are infected by paradoxes has to be rejected. It would be wrong to say that the people whose practices of reasoning were riddled with paradoxes did not have any principles of reasoning enabling them to distinguish cogent from erroneous arguments, and rightly guiding them in their thinking. It is evident that they engaged in thinking guided by reason just as much as anyone else does.

A more moderate understanding of their situation will have it that their practices of reasoning were mistaken, but had limited validity in being imperfect approximations of the sound principles of reason. On this view we can say that they were, by and large, guided by the sound principles of reason, except that their practices failed to incorporate and follow them adequately, thus

leading them on occasion to incorrect applications of the right principles. Apart from other weaknesses, this way of understanding their situation presupposes that there is a unique set of sound principles of reason, whereas – as was suggested above – there seem to be indefinitely many incommensurate sound sets of principles of reason.

The most promising suggestion is a modification of the second proposal, to allow for a plurality of incommensurate sets of principles of reason. It says that practices of reasoning infected by paradox can rightly be regarded as mistaken approximations of some sound principle of reason or other. Being such approximations we can say, as in the second proposal, that the people who followed them could be in part vindicated. They can be said to have been imperfectly guided by the sound principles of reason which their practices approximate. This suggestion purchases the advantage of not allowing that principles of reason can themselves be imperfect at too easy a price. In the absence of further criteria, and it is not clear what they might be, any practice of thinking approximates some set of principles of reason or other. This may not be an objection to the suggestion we are examining. Possibly no practices which allow one to distinguish correct from mistaken thoughts, episodes of thinking, or transitions of thought, can be discounted. All of them are practices of reasoning in the minimal sense we are exploring, i.e. they approximate some sound principles of reason, and their practitioners can be said to be guided by the sound principles these practices approximate.

Nevertheless, I find this suggestion unacceptable for it seems that any imperfect (i.e. paradox infected) practices of reasoning approximate not one but many sound principles of reason, many of which are unknown to us or to the practitioners, and, given our time and place, many are unknowable by us or by them. In these circumstances, denying that people are governed and guided by the principles they have, and insisting that they are really guided by principles beyond these, principles which they do not know, and perhaps cannot know, is an unhelpful verbal trick to avoid saying that principles of reason can be imperfect. The suggestion serves no other purpose. The "more honest" course is to keep the connection between principles people follow and the practices they engage in. This connection allows us to explain how people can be guided by principles which in some sense they do not know, i.e. of which they are not (fully) aware, and the con-

tent of which they cannot articulate. The price of allowing that principles of reason can be imperfect seems no price at all.

The preceding argument for the historical character of principles of reason presupposes meaning incommensurability. Is the presupposition justified? The issue is complex and it is easy to argue on the one hand for, and on the other hand against meaning incommensurability. Not everything which can be expressed in one language can be expressed in all others without extending their resources, either by adding new words, or phrases, or enriching their grammar. Once, however, enrichment is allowed what is there to stop a language from being enriched to whatever degree may be necessary to enable it to express whatever the other does, and to include all the concepts of the other? Can it be enriched by simply absorbing the other language as a part? Is not that the way the "languages of science" came to enrich the standard natural languages which, because of the existence of such specialised segments, can now express what they could not express without them?

If this is where the argument resides then I suspect that there is no 'principled' solution to this problem. That is, it cannot be solved by arguments which disregard the contingent and historical nature of languages (both natural languages, and specialised segments of them like the languages of science or of law). The issue of incommensurability of meaning turns out to be the problem of the identity of languages, and that is a historical matter. Whether something is a segment of another language or a separate language can depend not on relations of meanings and of rules of grammar, but on how the two developed and who uses them and when. But if the issue of meaning incommensurability is one of historical development then meaning incommensurability is possible and therefore principles of reasoning are themselves historical products.

I will, therefore, proceed on the assumption that the substantive principles of reason are historical products which can be challenged on grounds of self-contradiction, incoherence, and unreliability, and which can change to avoid such challenges. This allows scepticism more scope than Nagel allows it. It does not, however, allow for an attack on reason or rationality itself. Such an attack would have to be addressed to the formal concept of reason, and show that it is incoherent.

Before we consider this problem, one last word regarding critiques of substantive doctrines of reason. The preceding discus-

sion aimed to show that such doctrines are capable of being criticised. When paradoxes are discovered principles are revised to avoid them. This in itself does not constitute a justification of those principles which are not infected by paradox. Does one not need to provide them with some justification other than the timid response "so far no paradoxes affecting them have come to light"? Yes, and no. It is certainly possible to explain why such principles are valid. But the explanation is not of a kind normally thought of as justification. Rather the explanation will relate to the constitutive role of such principles, reconciling the fact that they are constitutive of a mode of thinking with the possibility of alternative, incommensurate, sets of principles, and noting that the validity of the principles does not guarantee that whatever thoughts they are manifested in are about "an independent, objective reality."[24]

5. The standing of formal reason

Is it possible for a sceptical argument challenging the coherence of the formal notion of reason, or the case for its use, to be successful? Many of the attacks on reason and rationality which we witnessed in the course of history are meaningful and arguably justified. Such, for example, are attacks on rationalism claiming either that people have come to exaggerate what can be achieved simply by conformity with principles of reason, or that the imagination and fantasy, or feelings and emotions, are of value, as well as thought in the narrow sense. None of these nor other similar claims need be disputed. Nor need we engage in an argument about the instrumental importance of rational thought, e.g., whether people or the human race generally can survive for long without it. The question is whether it is possible to prove that rational thought is inherently bankrupt. If it has instrumental value, the debunking argument goes, it has it in spite of its failure to live up to its aspiration to be a reliable guide to truth. Can one show that rational thought as such, not merely this or that substantive doctrine of rational thought, is incoherent?

Two possible routes towards this conclusion suggest themselves. Sceptical argument following the first route would prove that no account of the content of reason can escape a *reductio* argument showing it to be incoherent. While the refutation of

[24] See my "Notes on Value and Objectivity ".

any specific account of reason has to proceed via a *reductio* argument, an argument that any account of reason is subject to a successful refutation need not itself employ a *reductio* method of argumentation. Any form of proof would do.[25] The second line of attack would consist of arguments showing that the very formal notion of reason is incoherent. I do not believe that we have been confronted by any such arguments. But it may be useful to speculate on the effect they may, if successful, have.

Think first of the impact of the paradoxes affecting the substantive doctrines of reasoning. Their solutions did not lead to wholesale abandonment of rational concepts or principles of reasoning. Rather, they led to modifications of existing principles and practices of reasoning. This is what one would expect if one believes that all adjustments forced by arguments would tend to be the minimal necessary to satisfy the reasons forcing them. In terms of familiar metaphors: we start where we are and we proceed to adjust our position from there. We move not towards some unique ideal but towards a system of concepts and of reasoning which is readily reachable from our starting points and which seems to avoid the difficulties which force us to move from our initial position. It seems reasonable to assume that more radical paradoxes affecting the cogency of the formal concept of reason will, if they ever materialise, lead to similar partial adjustments.

It is true that the strategies of challenging formal reason seem to allow a more far-reaching conclusion. They seem to allow the conclusion that no principles of reason whatsoever can be valid. But appearances are misleading. To entertain that thought is to entertain the thought that propositions and concepts are discrete with no logical or conceptual relations between them. That would make mastering any of them impossible. Hence the most that a challenge to formal reason can succeed in doing is forcing adjustments to the concept of formal reason.

There is, however, a difficulty in conceiving of such an adjustment as anything other than a correction of a mistake regarding the one and only notion of reason. When considering the principles of reason my suggestion was that the pressure of paradoxes leads one to move from one's initial system of concepts and of principles of reasoning, through the smallest adjustment which

[25] Pyrrhonist sceptical arguments are of this kind, as they aim to show that if a proposition is supported by reason so is its negation.

happens to suggest itself, to what is strictly speaking an incommensurable alternative system. Both systems are instantiations of the formal notion of reason, though the second is (so far as we know) superior to the first in not being infected by paradoxes. When we abandon, through some adjustment, our formal notion of reason such a picture is impossible to sustain, for there is no genus which covers both the earlier and the later formal notions of reason. It seems as if strictly speaking we must regard the previous notion as a mistaken (because affected by paradoxes) version of the second. That is, we must assume that there is but one, though possibly not yet correctly understood, notion of reason.

The imagined radical paradoxes do not lead to a change in the concept of reason, but to the correction of mistaken beliefs about its nature. We cannot distance ourselves from the formal concept of reason, and adopt another in its stead. This enterprise does not make sense. This conclusion should not be misinterpreted. It is not that we are captives of "our" concept of reason and cannot abandon it even if it is mistaken. We can so to speak "abandon it", though what we abandon is not the concept of reason but the views we happen to hold about its nature.[26] These views are what sceptical arguments can hope to refute. The very possibility of formal reason cannot be refuted for the notion of such a refutation does not make sense. It assumes either that concepts are discrete with no conceptual relations between them, or that there can be an alternative to formal reason, and that is nonsense.[27]

6. Conclusion

Let me take stock. We saw (in section 3) that the thought that normativity as such should be defended and justified often amounts to a demand for the justification of reason. It encompasses questions I did not touch on; primarily doubts about the objectivity of reason, and about the possibility that people may be motivated by reason.[28] Putting such doubts to one side we saw that it is possible to advance sceptical arguments against any of the principles of reason. Such principles are historical products in

[26] We can of course "abandon" it in the trivial sense of ceasing to reflect about it.

[27] We can of course become creatures which cannot think (in the narrow sense of that word). That, however, is a fact about us not about the concept of reason.

[28] See Joseph Raz, *Engaging Reason* (Oxford: OUP, 1999) chs. 5 and 6.

the same way that languages and other systems of concepts are. They can be replaced by others which are, hopefully, free of paradox. Such changes are normally achieved not through wholesale rejection of principles of reasoning and of the systems of concepts which gave rise to them, but through their adjustment and modification. There cannot, however, be sceptical arguments against reason itself. For while it is possible for human beings to stop engaging in thinking, in the narrow sense of the word, and even to lose the ability to do so[29], it makes no sense to think that the concept of thought or of reason can be rejected or be found defective. Our understanding of it can be found wanting. It can change, but what can change is the understanding of a concept which remains the same.[30] It follows that the validity of specific normative principles can be called into question. And that we – human beings – can become creatures incapable of being guided by normative considerations. But so long as we are capable of rational thought we are capable of being normatively guided, and, while we can explain the nature of reason and normativity, there is no such enterprise as justifying normativity.

[29] It is plausible to suppose that we can stop thinking altogether only if we lose the ability to think.

[30] To avoid doubt let me add that throughout history the methods of reasoning employed by people have changed, hopefully improved, in many ways which have nothing to do with paradoxes, and sceptical arguments of any kind. Such improvements can happen as people's knowledge and understanding of the world they live in increases. The discussion above does not bear on such developments.

3

EPISTEMIC NORMS AND THEORETICAL DELIBERATION

Christopher Hookway

1. Introduction: two systems of epistemic norms

Some of our beliefs are justified, and some are not; some of them count as knowledge and some do not. When we describe a belief as justified or identify a state as one of knowledge, we make an evaluation of it. Most recent epistemology has been concerned with describing the standards we employ in making such evaluations and with establishing the right with which we endorse and apply those standards. This has involved the search for a theory of justification and an adequate explication of the concept of knowledge. Since beliefs are not actions, and since it is common to deny that belief formation can itself be subject to the will, it is natural to conclude that if such evaluations are the primary focus of epistemic evaluation, the systems of norms that guide them are rather different from those which guide our practical reasoning and action.

There is another way of thinking about the role of normative standards in our search for truth, one which has received less attention in general epistemology. We engage in theoretical deliberation, attempting to think and reason in a reflective and disciplined way with a view to arriving at a settled belief on some matter. This kind of deliberation seems to be an activity: we set ourselves cognitive goals; we consider the best means of achieving these goals; we reflect on the best strategies to employ in our deliberations; and so on. Deliberation can be carried out well or poorly; and in general, we hope that if our deliberations are carried out well, they will lead to true justified beliefs. In this respect, individual deliberation is analogous to public inquiry in which we attempt to reach the truth through discussion, experimentation and other investigative techniques. The latter is evidently a form of goal directed activity, and it is reasonable to accept that the former is a kind of activity too. Thus there is a second focus for the study of epistemic norms. We can examine the normative standards that guide us when we try

to carry out theoretical deliberations and conduct inquiries. These norms will guide us in formulating cognitive goals and selecting methods of deliberation or inquiry to employ in pursuit of them. Unlike the first focus, this one should reveal connections and parallels between norms of practical and of theoretical rationality.

How should we think about the relations between these two approaches to our normative practice? If we are rational, there should be a harmony or fit between them: beliefs that result from well conducted inquiries and deliberations should be justified or meet the standards required for knowledge. The most common approach implicitly assumes that the first set of standards is primary: there are not two distinct systems of norms here. Perhaps the normative judgments that we make in the course of our deliberations and inquiries use the vocabulary of 'knowledge', 'justification' and the like: the first set of norms suffices for all our epistemic needs. This seems plausible if we accept that a) the first set of norms will trace dependencies of our beliefs upon one another through structures of logical argument, and also that b) deliberation is simply a sequence of such arguments. I shall argue below that this conception of deliberation is a distortion, and that the first set of norms is not sufficient for our needs. We may think about whether we ought to repeat an experiment or observation, explore how far prejudice has interfered with our judgment, consider whether we abandoned our views too readily in the face of the disagreement of a figure of authority and so on. If the concepts of knowledge and justification have a fundamental role in these reflections, this is not phenomenologically evident. We have strategies and heuristics of self-questioning and doubt which do not seem to use these concepts. A second, more promising way to defend the primacy of the first set of concepts is by arguing that they have a role in evaluating the normative standards and vocabularies we do use when assessing our inquiries: norms of deliberation are defended by showing that their adoption will lead us to acquire justified beliefs or knowledge. How readily we should defer to experts should be settled by reference to whether doing so is a means to securing justified beliefs.

Alternatively, we might treat the second set of norms as primary. We might define *justification*, for example, in terms of the norms which govern deliberation and inquiry. A belief is justified when it is the product of good inquiry or deliberation – or where the relevant norms decree that no further inquiry or deliberation

is required. Since most epistemologists agree that there are different concepts of justification which answer to different cognitive needs, there may be problems in the way of working this out in detail. Someone's beliefs may be 'subjectively justified' but 'objectively unjustified' if they are the product of inquiry or deliberation whose flaws were not recognisable by the inquirer. Developing these needed distinctions requires an account of what makes an inquiry good, of when the norms someone reasonably accepts are not, in fact, the norms of good inquiry. We can treat the second set of norms as primary only if these judgments need not rest on claims about whether the deliberative norms will lead to knowledge or justified belief. We could do this by arguing that they yield beliefs that are true or are likely to be true. Or we could argue that they yield satisfactory and fruitful answers to our questions and problems. Justification would not then be a fundamental item in our vocabulary of epistemic evaluation.

There is also room for a complex view which denies either primacy claim: we have distinct systems of evaluative concepts which serve different purposes, and these may or may not be in harmony. We face deliberative questions about how to inquire or deliberate; and we make evaluations of states of apparent knowledge, perhaps when we are assessing potential informants. My sympathies lie with the view that the second set of epistemic norms is primary, and that they can be used to explain concepts such as justified belief. However defending this view is not the concern of the present chapter, which is an attempt to explore some features of the second set of norms, those that govern deliberation and inquiry.

We begin by exploring some features of practical deliberation, with the aim of exploiting some analogies between theoretical and practical inquiry. We exercise reflective self control over our inquiries and deliberations by raising questions about the methods we employ and the routes our thought has taken. Our epistemic values are manifested as much in the questions we do not raise as in those that we do. Indeed, possession of epistemic virtue consists in part in acquiring a body of capacities that ensure that questions arise only when they should. Our need is to be sufficiently reflective but not excessively reflective: we must learn to respect what I have elsewhere called the limits of responsible reflection. Taking this seriously encourages an approach to epistemic evaluation which is analogous to accounts of ethical evaluation which attach importance to states of character such as virtues.

2. Some features of practical deliberation

I shall use the term 'practical deliberation' to refer to an activity of conscious reflection that can both lead up to, and accompany, action. It is important to distinguish these two roles. When deliberation leads up to action, it attempts to answer a question concerning what to do: *Shall I do A?*; *How shall I achieve E?*; and so on. This can lead to a decision which determines the will, a decision in accordance with which we then act. But deliberation does not end when we begin to act. We raise and try to answer further questions about how to adapt our conduct to circumstances, about how to carry out different stages of the activity, and about whether and when we should revise our goals and our ideas about means of achieving them. As well as guiding decisions about what to do, practical deliberation enables us to monitor our behaviour, and if we could not do that, we would often not be able to act successfully.

Whether we can achieve our goal may depend upon our ability to carry out further deliberations that monitor our pursuit of it. From observing others, I may decide that playing squash is fun and is a good thing to do. Playing successfully depends upon having good hand/eye coordination and being able to respond rapidly to the play of the opponent. It requires a collection of skills and habits which it may take extensive practice to acquire. Exercise of these skills and capacities will be largely unreflective. Success depends upon not reflecting too much – that slows us down – but also upon sometimes reflecting: thinking about weaknesses in the opponent, wondering rapidly about where a lob will land and so on. I need to understand *when* to reflect, knowing both its limits and where it is appropriate. And it is important that I do not have to think about (reflect upon) where those limits lie during the course of play.

What does this example show us? First it illustrates how deliberation may lead us to adopt a goal. Second it shows that sometimes we need to deliberate in the course of carrying out the activity: we raise subordinate questions and think actively about how we should answer them. But, third, much of the knowledge and information that guides us in performing the activity effectively does so without being represented in deliberation. This can be for several reasons. Reflection takes time and there will often not be time to carry out the required reflection in time to act upon its results. And we only have to adopt a moderate sort of holism about reasons to see that the range of questions that *could*

be deliberated upon is enormous, if not infinite. Moreover our processing capacities are limited; ball skills, for example, may function very effectively so long as I trust them and do not try to intervene in their operation. Whether my activities are successful will depend upon my possessing skills whose operations I cannot control: if I lack them, I may simply miss the ball or see it go in an unintended direction. Once I do begin to reflect, my reflection will not engage with the skills and routines that I have learned. We can easily see that many of the questions that *could* be raised either need not, or should not, be considered within the context of a particular activity. Sometimes we may reflect about whether some question need or should be considered. But often this second order question will itself be one of those that is best not reflectively considered.

If we try to describe the phenomenology of such deliberation, we might say that some questions occur to us as relevant and important, while other wholly legitimate ones do not intrude upon our attention at all. If we are rational, the questions we need to consider are normally among those that do occur to us; in general, questions and issues emerge as requiring attention only if pursuing them will contribute to the success of our activity. Indeed we can be confident in our ability to reason effectively only if we are also confident that, by and large, issues enter our conscious deliberations if and only if their doing so is important for the success of our activities. The standards that are reflected in the fact that certain questions and facts do not occur to us are an important part of the normative structure of deliberation. Deliberation is a process of conscious reflection. It occurs against a background of what seem to be habits of thought and reflection. Unless we possess the appropriate body of habits, we will be incapable of the kinds of reflection required for effective deliberation.

(To avoid misunderstanding, we should mention a matter of terminology. Use of the word 'habit' can suggest that we are concerned with rigid, inflexible, repetitive patterns in behaviour. This goes beyond the claim that, perhaps through a process of habituation, we acquire abilities to make these evaluations in a largely automatic and unreflective way; making them becomes second nature. I do not intend to go beyond this latter claim; it is no part of my view that we have to rely upon such rigid patterns. Since there is no obvious choice of word to use to describe such capacities, I shall continue to use 'habit', sometimes speaking instead of

capacities and dispositions. I suspect that this usage accords with some of the later pragmatists, but I am aware that it is not fully satisfactory.)

This claim about the requirements for effective deliberation is important for understanding the role of states such as virtues in shaping our deliberations and behaviour. Consider the example of *courage*. Suppose I am walking in the hills. The weather closes in and it becomes very windy. I have to cross a narrow ridge – and it will take considerable courage to do so. We might imagine someone who was aware of this, who desired to show the appropriate courage, but was simply frozen to the spot. The case that interests me here, though, is slightly different. Courage requires a sort of confidence (compare William James on chasm jumping).[1] Suppose I start to cross, and thoughts keep occurring to me about: the danger that some of the rocks are loose; the danger that a sudden gust will hit me on an exposed part of the walk and I wonder whether I will be able to withstand it; whether I might slip when I jump from one rock to another; etc. Now I may be aware that most of these dangers can be discounted; and I will also be aware that if I spend too much time thinking about them, this will weaken my resolve. They clutter my deliberations, upset the rhythm of my striding pattern, encourage counter-productive hesitations and so on.

In similar vein, deliberation that promotes benevolent behaviour must have a distinctive character. The welfare of others must be perceived as providing reasons for action. And thoughts of the minor personal costs of benevolence must not clutter and impede deliberation. Such thoughts must have little weight; the question whether there are such costs should not prompt inquiry; ideally such thoughts have no deliberative salience at all. They intrude only when it would be appropriate for them to do so. There is a further complication here which is connected with Kant's claim that benevolence is an imperfect duty. We are not required to seize every opportunity to act benevolently; but a good life will embrace at least some such possibilities. Again we might suppose that it would clutter our deliberations if we were required to think about

[1] In 'The Sentiment of Rationality', James imagines climbing in the Alps, and being forced into a position from which the only escape is by means of a 'terrible leap'. Although he has no evidence of his ability to make such leaps he suggests that 'hope and confidence in myself make me sure I will not miss my aim, and nerve my feet to execute what without these subjective emotions would perhaps have been impossible'. *Selected Writings*, ed. G. Bird. (London: Dent, 1995) p. 43.

each occasion to determine whether it was one on which benevolence was desirable. But it would then be easy to fail to seize on *any* such occasion. Ideally what is required is that sometimes – in appropriate and sufficient cases - the possibility of benevolent action occurs to us. But that this occurs should not be something which we consciously monitor and control.

So: an adequate account of deliberation will have to take seriously the idea that its success will depend upon possession of dispositions to raise questions, find considerations salient or relevant and so on. They reflect our deliberative values, shaping the routes that our deliberations take and, hopefully, contributing to their success. Of course, they may provide obstacles to the success of our deliberations, embodying flawed normative standards. Indeed they can even fail to accord with standards that we would consciously endorse, impeding our attempts to reason in accord with these avowed standards. Some forms of irrationality arise, then, when our normative and other commitments are not in harmony with (or sustained by) the normative standards that are embodied in these habits or dispositions. We find things salient when we shouldn't, and we fail to find things salient when we should. We find things salient even when we know we shouldn't, and fail to do so even when we know we should. And our ability to deliberate well may be compromised even if we are aware of the position and try to resist these misleading saliences.

Theoretical deliberation and inquiry displays these features too. Monitoring inquiries and deliberations involves asking and trying to answer subordinate questions about the plausibility and coherence of solutions to problems, about the adequacy of methods of answering those questions and about the reliability of means of executing those methods and so on. Once again, these questions occur to us and we normally trust that the appropriate questions will do this and inappropriate ones will not. Facts, observations and memories will also occur to us, become salient, when, we hope, they serve as presuppositions of subordinate questions that should be asked or when their relevance to our concerns should be assessed or evaluated. Effective responsible inquiry depends upon the trustworthiness of our habits of observation, reflection and questioning.

These phenomena raise two sorts of issues. First, how should we describe and explain the normative dimension of these dispositions, habits and unreflective saliences? Under what circumstances are they *good*: what features should we want them to have?

We can say vaguely that questions should occur to us only when they are 'relevant' or 'salient' for our deliberative aims, but this cries out for further clarification. Second, it seems problematic that we appear to have no control over these habits and dispositions. Discussing the merits of intellectual 'wholeheartedness', John Dewey wrote that when an inquirer is 'absorbed':

> The subject carries him on. Questions occur to him spontaneously; a flood of suggestions pour in on him; further inquiries and readings are indicated and followed; instead of having to use his energy to hold his mind to the subject . . . the material holds and buoys his mind up and gives an onward impetus to thinking.[2]

Dewey's discussion is useful in identifying the two sided character of effective deliberation. Questions 'occur to him'; suggestions 'pour in on him': such idioms indicate that we are the passive recipients of these questions and suggestions. We find ourselves facing them, and we do not control – or even understand – the processes that produce them. Wholeheartedness is viewed by Dewey as a virtue because it allows this influx full rein: we don't have to rely upon reflection and the will to overcome distraction and to point us in the right direction. On the other hand, when Dewey describes this as the exercise of 'an intellectual force', and as an expression of an attitude, he presents the agent as confidently embracing his own capacities for responding to problems intelligently and rationally. This passive influx is a kind of sensitivity to the normative demands of reason and, in line with the suggestion of Michael Smith and Philip Pettit, an expression of our freedom of mind.[3] I shall now say something about each of these issues.

3. Salience, relevance, appropriateness

Relative to a particular stage in the conduct of a particular deliberation or inquiry, agents take propositions or questions to be

[2] See John Dewey, *How We Think* (Boston: D. C. Heath and Co, 1933), p. 31–2. This passage, and related ones, are also discussed by Linda Zagzebski in *Virtues of the Mind* (Cambridge: Cambridge University Press, 1996), p. 173f).

[3] Philip Pettit and Michael Smith, 'Freedom in thought and desire', *Journal of Philosophy*, 93, 1990. The importance of making sense of the idea that the fact that we do not have full reflective control over our cognitive activities need not compromise our freedom of mind is discussed in Hookway *Scepticism* (London: Routledge, 1990) passim.

salient if they treat them as significant, as something of which account should be taken in carrying out the deliberation or inquiry in question. There are at least two grades of psychological salience we might take note of. There are questions and propositions which occur to the deliberator, that flow in upon him or force themselves upon his attention. And there are questions and propositions whose importance he would acknowledge if they were proposed by a third party. In each case, it is a fact about someone's psychology that, within a particular context, they find themselves attaching salience to a given question or proposition. In fact there are more layers of complexity than so far allowed. A question or proposition can force itself upon my attention only for me to judge, often quite quickly, that it should not have done so: I instinctively reject the psychological salience that it displays. And this may, or may not, prevent the question or proposition from continuing to force itself upon my attention. When it does continue to make claims upon me, this may be a case where I am genuinely the passive recipient of its demands, seeing them as an unwelcome impediment to rational decision making. For the present, it is important that we have a complex array of psychological phenomena of deliberative salience.

It is important that someone's finding a question or proposition salient in the context of a deliberation or inquiry involves taking an evaluative stance towards it. Salient questions are those which, we take it, *ought* to be considered and answered. There is something passive about this salience – we *find ourselves* treating a question or proposition as salient, and salient questions and propositions will often just *occur to* us. But success in our inquiries and deliberations involves our trusting these habits of salience, having confidence in them. And this confidence involves the expectation that the questions that occur to us are those which we ought to consider. As we shall see in the next section, this becomes significant when we address the issue of how they relate to issues of activity, passivity and freedom of mind.

Since treating something as salient involves an evaluation, questions arise about the adequacy of the evaluations that they embody: are the patterns of evaluation which they impose upon our deliberations and inquiries as they ought to be? It is not immediately obvious what the appropriate vocabulary is for these evaluations. We might say that these questions and propositions occur to us only when they are *relevant* to our deliberative concerns. If this is to be more than just an empty place holder, we need to say something

more about what 'relevance' consists in. And any way of doing this is likely to draw the line in the wrong place. We noted that in certain circumstances, it might be wrong for the confident hill walker to raise questions about the security of the stone on which he is about to stand or about the dangers of being caught by a sudden gust as he leaps from one rock to another. But no one could deny that this information is *relevant* to his task. Had he already known that the rock was very insecure, he would surely have altered his route. Checking the state of the rock would provide information that would enable him to draw conclusions from his background knowledge that could not otherwise be obtained.[4] But this does not provide a sufficient condition for the subordinate inquiry being warranted. This relevant new information may not bring enough benefits; and the costs of obtaining it may be too high. A question should be salient, we may suppose, if the attempt to answer it will produce considerable benefits for relatively little cost. What sorts of things can counts as benefits and costs here?

Let us begin by mentioning some possible benefits. Suppose that failing to carry out the subordinate inquiry raises to an unacceptable level the chances of the overall deliberation or inquiry failing to achieve success. This could involve the risk of the overall deliberation or inquiry reaching no solution to its problem at all or the risk of its arriving at a solution that is incomplete or simply wrong. If considering the question raises the probability of arriving at ones goal, then this provides a good but defeasible reason for thinking that the question should be raised.

The costs are more varied. Even if possessing an answer to the question would increase the probability of our inquiry being successful, reaching the answer will take time and energy. If other questions are competing for the agent's attention, if time is short, if economic resources or energy are in short supply, then the overall probability of success may be reduced by pausing to address this question. And even if that is not so, a small increase in the chance of success may not compensate for the efforts and stresses of addressing this subordinate question. I shall not here explore different formal treatments of this interplay of costs and benefits.

The costs we have just considered show that the salience of a reflective question should be a function of the time and energy

[4] Thus considering the question will indeed provide 'relevant' information if 'relevance' is defined along the lines adopted by Dan Sperber and Dierdre Wilson in *Relevance* (Oxford: Blackwell, 1995) chapter three.

that addressing it will occupy. It can delay completion of the overall inquiry or prevent consideration of other more pressing questions. If we are impressed by the idea that questions of justification lead to an infinite regress, then there might be an argument for refusing to pursue some relevant questions which is based upon less contingent considerations. If indefinitely many questions are candidates for consideration, then it would be disastrous if we were to find all of them salient. More importantly, some of our knowledge takes the form of habits of thought, skills in the use of concepts and argument forms, ways of exercising judgment in weighing considerations and so on. It is evident that the bases on which we do this are not available to introspection: complex cognitive achievements can have a sort of phenomenological immediacy.[5] The more reflective we are about how our deliberations and inquiries are conducted, the less this practical knowledge can guide them. This may be the most serious cost of raising such reflective questions. Reflectively seeking relevant information can reduce our chances of success in inquiry by blocking access to all that knowledge which is not open to introspective monitoring.

4. Passivity, activity and reflection

We now turn to the second of the issues introduced at the end of section two. Talk of questions and propositions occurring to us, or flowing in upon us, makes the agent seem like a rather passive participant in his own inquiries and deliberations. We do not decide to admit these questions and considerations, they just arrive. On the other hand, I have suggested that this influx reflects *our* epistemic values, and we can surely be held responsible for at least some failings in their operation. There are cases where what occurs is properly described as passive. For example the question of the reliability of an informant may arise, to be reasonably rejected, but then to constantly reassert itself, occupying time and energy as an intrusive consideration whose irrelevance is evident from the moment of each arrival. This form of passive occurrence of questions and considerations is plainly irrational or pathological and is a feature of one's cognitive functioning from which one can reasonably

[5] This claim is defended further in 'Mimicking Foundationalism:on Sentiment and Self-control', *European Journal of Philosophy*, I, pp. 156–74, and 'Doubt: Affective States and the Regulation of Inquiry', *Canadian Journal of Philosophy*, Supp. vol. 26, pp. 203–26.

feel alienated or distanced: the 'values' which prompt the question are not one's own. Indeed the 'salience' of the question is something *merely* psychological. One believes, correctly, that these questions ought not to be salient; and it is impossible to view their constant reappearance as a kind of sensitivity to something with normative authority. Our concern is with those common and distinctive cases where salience is seen as an expression of one's own normative standards. Some idioms used by Dewey and others are suggestive: allowing ourselves to be guided by such saliences expresses one of our attitudes; it is a reflection of our intelligence or rationality; it is not an impediment to our freedom of mind; it may be a manifestation of something of the nature of a virtue or character trait; it is reflected in the confidence with which we carry out our inquiries.

Let us return to some of our earlier analogies with moral virtues. Somebody benevolent, we suggested, does not directly interrogate each situation she faces, wondering whether it provides an opportunity for benevolent action, and investigating the matter before proceeding. One of Kant's arguments for the necessity of virtues is that they are required to alert us to suitable occasions for fulfilling our imperfect duties: we must perform an appropriate number of benevolent acts, and the benevolent person will be appropriately moved by the occasions for benevolence that occur to them. Although he is not actively on the lookout for occasions for exercising benevolence, he is 'open to them', confident that such occasions will grasp his attention and call upon him to act. This 'openness' has a dimension which makes it inappropriate to think of it as a merely passive matter. He takes himself to possess a sensitivity to the suffering of others, a sensitivity which is manifested in the sympathy and concern that such circumstances produce in him. It is a response which is 'automatic', which is not subject to critical self-control. But it is a manifestation of his values. Were he to encounter someone whose suffering did not produce this response in him, this would be viewed as a cognitive failing that would merit a range of reactive emotional responses. He may also react in a similar way to other people when they do not so respond to suffering.

A similar cognitive example is provided by the state of being *observant*. Someone with this cognitive virtue is presumably not expected to be constantly on the look out for interesting snippets of information. She does not incessantly scan the room considering the

interest and importance of each thing she sees. Rather she is, once again, *open* to her surroundings, taking notice of things that are interesting and important: relatively few interesting or important matters escape her attention, just as few trivialities occupy it. There is a further respect in which these sensitivities are 'active'. The observant person has a disposition to take note of things, attending to them and examining them further; and the benevolent person has a tendency to respond to suffering by taking an immediate active interest in its degree and the ways in which it might be alleviated. The responses already involve active attention and investigation.

Some similar phenomena are identified in a passage of Quine's about our ability to assess the force of evidence.

> The sifting of evidence [is]... a strangely passive affair, apart from the effort to intercept helpful stimuli: we just try to be as sensitively responsive as possible to the ensuing interplay of chain stimulations. What conscious policy does one follow, then, when not simply passive toward this interanimation of sentences? Consciously the quest seems to be for the simplest story. Yet this quality of simplicity is more easily sensed than described. Perhaps our vaunted sense of simplicity, or of likeliest explanation, is in many cases just a feeling of conviction attaching to the blind resultant of the interplay of chain stimulations in their various strengths.[6]

Both aspects of the phenomena we are concerned with are present here. The weighing of evidence is explicitly described as 'passive': we do not control the processes by which it occurs and may not be able to formulate the standards we follow. The feeling of conviction or awareness of the 'simplicity' of the view that results signals our active endorsement of it: we often confidently embrace the results of our passive sifting, employing them in guiding conduct and pursuing further inquiries. There is no role for reflective control over the process of sifting: in an earlier chapter I described Quine's naturalised epistemology as resting upon his acceptance of the 'shallowness of reflection'.[7] But, we may add, our cognitive successes are based upon our legitimate confidence in the capacities which are involved in this process.[8]

[6] W. V. Quine, *Word and Object* (Cambridge MA: MIT Press, 1960), p. 19.
[7] 'Naturalized Epistemology and Epistemic Evaluation', *Inquiry*, 37, p. 476ff.
[8] According to 'Doubt: affective states and the regulation of inquiry', these auto-

In this section we have emphasised the interplay of idioms of passivity and of active endorsement in the ways in which we describe out relations to the questions and propositions that we find salient. But we have so far said little about what this active endorsement, this confidence in or embracing of our sense of salience, consists in. The remainder of the chapter addresses this issue.

5. Virtues and confidence

We have been taking note of some of the ingredients of our practice of epistemic evaluation, noting how deliberation is guided by a kind of sensitivity to the salience of questions and considerations which is largely automatic or even habitual. At several points, I have remarked that the phenomena we have noted encourage an approach to epistemic evaluation which appeals to states of epistemic character or virtues. We now turn to a slightly fuller discussion of this claim.

If someone understands the concepts of benevolence and courage, then she must be aware of the sorts of considerations that are relevant to whether an action or person falls under the concept in question. She may also be skilled at establishing whether some act is courageous – whether performed by somebody else or by herself. She may even know that it is good to be courageous and be disposed to regret when she acts in ways that are not courageous. In spite of this grasp of the concepts, she may still lack the corresponding virtues. The habitual or automatic evaluations reflected in the agent's deliberative practice may still reflect selfishness or timidity. Whether this is correctly described as a form of irrationality is a controversial matter and need not be addressed here. The important issue is that someone's deliberative practice can display a pattern of evaluation which is at odds with evaluations that are sincerely endorsed by the agent and that can reflect an ethical commitment on the agent's part.

Consider an epistemic example. An agent asks: am I being

matic, unreflective evaluations are generally tinged with sentiment or with affective qualities. My active engagement with them can be manifested in the fact that, as well as apparently endorsing the inquiries which they govern, I endorse further inquiries that develop from them. The evaluations spread through my cognitive practice and this does not lead to any sense of alienation from them: they are suitably integrated into my evaluative practice. A related point, suggested in our discussion of Quine is that our participation in inquiry depends crucially upon the exercise of *judgment*. This notion, too, is intended to capture the crucial combination of the passive and the active in the evaluations that are made.

properly open minded or free of prejudice? And, given that she can answer that question correctly and feel appropriate reactive emotions, a negative response may be an occasion for regret and self-criticism. Such reflective self-criticism need not produce self-mastery; the opinions produced by the flawed deliberation or inquiry may well survive. Furthermore, however firmly the agent is committed to avoiding such prejudice, the required reflective questions may not actually arise. In that case, a sincerely held epistemic value may, without irrationality, fail to shape our theoretical deliberations.

We can draw a distinction here. A deliberative project is an attempt to solve a problem or answer a question through reflective reasoning and inquiry: it has a definite goal. Monitoring for failures of open-mindedness can occur internally to that project, as something one does as part of carrying it out. It can also occur as something external to the project, a third person perspective reflection upon how one has performed. These examples suggest that a concern with virtue, with open-mindedness or freedom from prejudice for example, can preoccupy the agent from this third person perspective, as part of a deliberative project which is directed upon, but is external to, the project of primary concern. And it can do this even if such evaluations are not operative *within* the project that is of primary concern, indeed even if they cannot engage with the execution of that project. The agent we are concerned with can use virtue *concepts* in reflection and evaluation but cannot display the corresponding virtues in the conduct or regulation of inquiries and deliberations. The agent knows that courage or open mindedness are good things, but (somehow) is unable actually to *be* courageous or open minded. And this can be the case in spite of the fact that she is fully aware of the situation.

I am interested in one particular sort of case, where we want to say that the agent does not *know how* to act courageously or benevolently. A number of different things are involved here, and thinking about them will help us to see some important features of rational deliberation. In the normal case, if one is benevolent or courageous, this is manifested (primarily?) in the shape taken by one's deliberations (not just in the acts one performs). The thought is that the courageous person's continuing deliberations have a distinctive character, and it is possible both to know what courage is, and to desire to perform courageous acts, while one's deliberations cannot take the required form. Indeed earlier sections of this chapter attempted to say something about just what this involves. If one is properly courageous, this is reflected in the

considerations and questions that emerge as salient in the course of practical deliberations that lead up to action or that monitor it while it is in progress. This, I think, is an Aristotelian thought: it is manifested in a state similar to akrasia which is often present as an intermediate stage in acquiring a virtue. Until the habits of inference and deliberation that are partly constitutive of courage are present, then one may well be incapable of the sorts of deliberations that are required if one is to act as one judges one should. We have to learn to think as a courageous person would think.[9]

And this learning has to involve the acquisition of skills and capacities. It cannot be a matter of mastering rules which are consciously applied in planning and evaluating deliberations. This is because what has to be learned has a sort of negative character; it is manifested in the fact that distinctive thoughts and questions do *not* occur to you in the course of your deliberations. The courageous person is not someone who is ready to ignore the minor dangers that he faces. Rather he is someone for whom those dangers do not present reasons for deliberation at all. He has to learn not to find certain kinds of considerations salient. Until he has acquired this negative deliberative capacity, he is likely to fall into a state similar to akrasia, an inability to control onself which represents a stage in the acquisition of an Aristotelian virtue. He cannot perform actions or carry out inquiries, which he knows to be called for by values that he endorses, because he lacks the capacities for deliberation which are required for the successful exercise of those virtues.

Can we find any epistemic cases which share these features? First I have been making use of the idea of normative commitments which have a sort of negative character: they are manifested in the fact that various questions and considerations are not salient, do not occur to us as relevant to our deliberations. And through examining examples, we have seen how time considerations, and limitations to our processing capacities, mean that unless we have mastered such normative standards in the form of habits and skills, we will be unable to exercise the deliberative capacities that are required for effective actions. Too many questions arise – possibly indefinitely many – and deliberation could

[9] The analogies between these phenomena and akrasia are evident from Myles Burnyeat's paper 'Aristotle on learning to be good', in A. Rorty (ed) *Essays on Aristotle's Ethics* (Berkeley and Los Angeles, University of California Press, 1980). Linda Zagzebski also alludes to such cases in *Virtues of the Mind*, p. 150ff.

never come to a close. Or questions may arise which inhibit the effectiveness of habits of inference and deliberation whose unfettered operation is necessary for the success of our projects. What we require is that questions and considerations become *salient* only when they contribute to the success of our deliberation.

It seems quite plausible that someone who acknowledges the merits of avoiding undue deference to authorities should be fated regularly to experience the regret that follows from their failure to do so. Or that someone's readiness to admire open-mindedness should ensure that they are rarely able to admire their own inquiries and deliberations. Training and experience are required before the appropriate habits are required and the 'right' questions emerge in the course of inquiry. That someone might lack the ability to display the cognitive virtues which she endorses seems a regular feature of our experience. We could also envisage the case of someone impressed by the truth of philosophical scepticism who lamented his inability to raise appropriate challenges to his everyday beliefs. Possession of epistemic virtue depends upon the possession of skills and habits whose possession is largely independent of the recognition that some state is, in fact, such a virtue. And possession of these capacities seems to be what is required for confidence in one's deliberative skills, for example in one's sense of salience, not to be an impediment to one's freedom of mind.

Conclusion

This chapter has drawn attention to some features of the norms governing deliberation, both practical and theoretical, which lend support to the suggestion that a virtue based approach can offer insights into our practice of epistemic evaluation. Deliberation involves raising and trying to answer questions, including reflective questions about the conduct of the inquiry or deliberation itself. I have emphasised that our ability to deliberate well depends upon our possessing evaluative knowledge with a distinctively negative and apparently passive character: this knowledge is manifested in the fact that various questions and considerations do *not* occur to us as salient. The second half of the chapter has attempted to raise some questions about how this notion of salience should be understood and about how this apparent passivity can be reconciled with the fact that we can take a form of active responsibility for the success of our deliberations.

A crucial question has concerned how we can embrace, or possess confidence in, our ways of finding considerations salient to our deliberations: what is the basis of our assurance that we are rational in relying upon our 'automatic' evaluations? We rely upon relatively stable character traits. We must feel confident that they enable us to be sensitive to the demands of reason: they guide us in taking questions seriously only when they are genuinely salient; they enable us to respond to the complexities and idiosyncrasies of particular cases; and they can develop flexibly in the light of experience and inquiry. As our examples have suggested, such traits are similar to the *virtues* that have been discussed by philosophers of morality. It is natural to conclude that attention to virtues will contribute to our understanding of epistemic evaluation too.[10]

[10] I am grateful for the comments I received when drafts of this paper were delivered at the University of Birmingham, at the Moral Science Club in Cambridge and at the University of Valencia. I am also pleased to have benefited from Jonathan Dancy's comments which have led to many improvements.

4

NORMATIVE REQUIREMENTS[1]

John Broome

1. Introduction

Our most familiar normative concepts are 'ought' and 'a reason'. Most of the philosophical discussion of normativity revolves around these two. But there is another that is equally fundamental. I call it 'normative requirement'. It is not so familiar, and is often confused with the other two. This chapter describes normative requirements, and shows what an important feature of normativity they are.

A normative requirement is a relation: one thing normatively requires another. Sections 2 and 3 distinguish various normative relations in a formal way, in order to separate the relation of normative requirement from others. Sections 4, 5 and 6 set out some examples of normative requirements. The one in section 6 comes from the context of practical reasoning, and practical reasoning forms the main context for the rest of the chapter. Section 7 explains how normative requirements are easily confused with reasons, but the rest of the chapter demonstrates how important it is to distinguish the two. Section 8 shows how very differently they behave when there are conflicts. Sections 9, 10 and 11 describe three mistakes that have been made by philosophers as a result of ignoring the existence of normative requirements. They all constitute misunderstandings of the nature of practical reasoning. Recognizing each of them helps to illuminate the working of normative requirements, and to show their importance.

[1] This paper was written while I was a Visiting Fellow at the Swedish Collegium for Advanced Study in the Social Sciences. I thank the Collegium for its very generous hospitality. I have benefited greatly from discussions with Simon Blackburn, Jonathan Dancy, Sven Danielsson, Stephen Darwall, Susan Hurley, Donald Hubin, Karsten Klint Jensen, Jan Odelstad, Derek Parfit, Philip Percival, Christian Piller and John Skorupski on the subject of this paper.

2. Types of normative relation: detaching relations

No doubt there are some things you ought to do, ought to believe, ought to want, and so on. To accommodate all these cases, I shall take 'you ought' to govern a proposition. Grammatically, 'ought' takes an infinitive rather than a noun clause, but an infinitive, like a noun clause, may denote a proposition. The subject is implicit; it is normally the subject of the governing verb. 'I hope to see the Pacific' means the same as 'I hope that I see the Pacific'. Similarly, 'You ought to relax' would mean the same as 'You ought that you relax', if only grammar would permit this latter sentence. 'You ought' governs the proposition that you relax. I shall formally represent 'you ought' by a propositional operator 'O'; I shall write

$$Oq,$$

where q is a proposition.

'O' cannot be comfortably translated into English. Whenever I can, I shall render 'Oq' as 'You ought to q'. This expression sounds strange because we are used to denoting propositions by sentences, which become noun clauses when they occupy a subordinate position. But as I say, an infinitive in a subordinate position can also denote a proposition. So 'You ought to q' is arguably grammatical. However, it fails as a rendering when 'O' governs a complex proposition, or when I need to quantify over propositions. Then, in desperation, I shall have to adopt the unsatisfactory 'You ought to see to it that q'. I use 'to see to it' as mere grammatical padding, which allows a noun clause to plug into an 'ought'. In common usage, 'You ought to believe in God' differs in meaning from 'You ought to see to it that you believe in God'. But as I use 'to see to it', these sentences have the same meaning. So do the sentences 'You ought not to see to it that q' and 'You ought to see to it that not q'.

What you ought to see to is often supposed to be determined by a balance of considerations, which may conflict. For example, there may be some considerations in favour of believing the suspect is guilty, and some against. To cover cases like this, we say you *have a reason* to believe she is guilty. This is consistent with your also having a reason not to believe it. A reason, in this sense, is *pro tanto*. In this chapter I shall only use 'a reason' in this sense. I shall write:

$$Rq$$

and translate it 'You have a reason to q'. If you have a reason to q and no reason not to q, then you ought to q.

If you have a reason to q, some fact or other constitutes your reason. This fact is a reason for you to q, and makes it the case that you have a reason to q. Let p be the proposition that this fact obtains. Slightly inaccurately, I shall say that the proposition p (rather than the fact) is a reason for you to q, and that p gives you a reason to q. A plausible example is: that you are thirsty is a reason for you to drink water.

If p is a reason for you to q, a particular normative relation holds between the propositions p and q: the relation of being a reason to. I shall write

$$p \text{ reasons } q, \tag{1}$$

meaning 'p is a reason for you to q'. If this relation holds, one consequence is that

$$p \rightarrow \mathrm{R}q, \tag{2}$$

where $\rightarrow$ is the material conditional.

(2) is a consequence of (1), but not equivalent to (1). Because (2) is only a material conditional, it can be true even if p is not a reason for you to q; for instance, it is true whenever p is false. (1) says that, if p, you have a reason to q, and furthermore, you have a reason because of p. We could say that (1) is (2) with *determination* added, from left to right. I shall not try to analyse this idea of determination, but leave it intuitive. It is roughly analogous to causation. I shall call (1) a 'determining relation' and call (2) the 'logical factor' of (1).

An important feature of the reasons relation follows from (2). A material conditional allows its consequent to be detached by modus ponens. If p reasons q, (2) tells us that, if p is the case, you have a reason to q. If you are thirsty, you have a reason to drink water.

If you have a reason to q, there is some fact that makes this the case. Similarly, if you ought to q, there is some fact that makes this the case, too. Let p be the proposition that this fact obtains. Then a different normative relation holds between p and q. I shall write

$$p \text{ oughts } q, \tag{3}$$

meaning 'p makes it the case that you ought to q'. If this relation holds, one consequence is that

$$p \rightarrow \mathrm{O}q. \tag{4}$$

(4) is a consequence of (3), but not equivalent to (3). (3) is (4) with determination added, from left to right. (4) is the logical factor of the determining relation (3).

The consequent in (4) is detachable like the consequent in (2). So, given (3), if p is the case, you ought to q.

The difference between the oughts relation (3) and the reasons relation (1) might be put like this: the former makes a strict demand on you; the latter a slack one. Suppose p is true but q is not. Then if the oughts relation holds, you are definitely failing to see to something you ought to see to. You ought to see to it that q, and you do not. On the other hand, if only the reasons relation holds, you may be failing to see to nothing you ought to see to. You have a reason to see to it that q, but you may also have a better reason not to see to it that q, and in that case you are doing nothing wrong if you do not see to it.

The oughts relation and the reasons relation are two sorts of normative relation that may hold between two propositions p and q. Both permit a normative conclusion to be detached, if p is the case. One permits Rq to be detached, the other Oq.

3. Types of normative relations: non-detaching relations

Other normative relations may hold between propositions. In this chapter, I shall be particularly concerned with one I call 'normative requirement'.[2] I shall write

p requires q (5)

and translate it as 'p normatively requires you to q'. If this relation holds, one consequence is that

$O(p \rightarrow q)$ (6)

you ought to see to it that, if p is true, so is q.

(6) is a consequence of (5), but not equivalent to (5). Because (6) contains only a material conditional, it may be true even if p does not normatively require you to q. For example, suppose you ought to clean your teeth:

O(You clean your teeth).

[2] In his 'Practical reasoning and the logic of requirement' (in *Practical Reason*, edited by S. Körner, Blackwell, 1974, pp. 2–13, reprinted in *Practical Reasoning*, edited by Joseph Raz, Oxford University Press, 1978, pp. 118–127), Roderick Chisholm provides a logic for a relation he calls 'requirement'. Chisholm's 'requirement' is *not* the same as my 'normative requirement', because it implies (2). See his D7, on p. 126 of the reprinted version. Chisholm's 'requirement' is my 'reasons' relation (1).

Logical equivalents can surely be substituted within the scope of 'O'. So it follows that:

O(Grass is red or grass is not red → You clean your teeth).

But plainly the tautology that grass is red or grass is not red does not require you to clean your teeth; it is irrelevant to that. (5) says you ought to see to it that, if *p*, then *q*, and furthermore, it is *p* that requires you to *q*. Once again, (5) is (6) with determination added, from left to right. (6) is the logical factor of the determining relation (5).

From (6), neither of the material conditionals (2) or (4) follows. Consequently, neither follows from (5) either. Whatever the nature of the determination (5) adds to (6), it is plainly not a material conditional like (2) or (4). Consequently, a normative requirement does not permit a normative conclusion to be detached by modus ponens. To put it more graphically, in (5) normativity is attached to the relation between the propositions *p* and *q*, whereas in (1) and (3), which permit detachment, the normativity is attached to the consequent *q*.

For my purposes, the essential features of the normative requirement (5) are, first, that it implies (6) and, second, that it does not imply (2) or (4). Beyond those two features, I shall deliberately leave the concept of a normative requirement as open and intuitive as possible. (I shall say a little more about its logic in section 8.) In sections 4, 5 and 6, I shall give examples to demonstrate that normative requirements are important in practice. However we may try to systematize normativity, we shall always need to give a place to normative requirements. Consequently, I do not want to tie them down to any particular deontic system or account of normativity.[3]

For completeness, I need to mention a fourth normative relation, which I shall call 'normative recommending'. I shall write

p recommends *q* (7)

to mean that *p* normatively recommends you to *q*. If this relation holds, one consequence is that

[3] Deontic logic contains a notion of conditional obligation, written $O(q|p)$, which could serve as a model for normative requirement. (See the survey in Lennart Åqvist's 'Deontic logic', in *Handbook of Philosophical Logic, Volume II*, edited by D. Gabbay and F. Guenthner, Reidel, 1984, pp. 605–714.) But deontic logic will not give us much help because the analysis of conditional obligation remains unsettled.

$$R(p \rightarrow q); \quad (8)$$

you have a reason to see to it that, if p is true, so is q. (8) is the logical factor of the determining relation (7). (7) does not imply either (2) or (4); it does not permit a normative conclusion to be detached by modus ponens.

The difference between the relation of normative requirement (5) and the relation of normative recommending (7) is that the former makes a strict demand on you and the latter a slack one. Suppose p is true but q is not. Then if the requirement relation holds, you are definitely failing to see to something you ought to see to. (6) tells us you ought to see to it that if p is true so is q, and you do not see to it. On the other hand, if only the recommending relation holds, you may be failing to see to nothing you ought to see to. (8) tells us you have a reason to see to it that if p is true so is q, but you may have a better reason not to see to this, and in that case you are doing nothing wrong if you do not see to it.

To summarize, the normative relations I have mentioned can be classified by the two criteria of detachment and strictness:

	Strict	Slack
Detaching	p oughts q $p \rightarrow Oq$	p reasons q $p \rightarrow Rq$
Non-detaching	p requires q $O(p \rightarrow q)$	p recommends q $R(p \rightarrow q)$

In each box of this table, I have put a determining relation together with its logical factor.

4. First example: believing you ought

This section and the next two give examples of normative requirements. The first example is the relation that holds between believing you ought to see to something and seeing to it.[4] Clearly some sort of normative relation holds between these things, but what is this relation exactly?

[4] This example appears in Jonathan Dancy's 'The logical conscience', *Analysis*, 37 (1977), pp. 81–4.

Is it that your believing you ought to see to something makes it the case that you ought to see to it? I shall write 'B' for 'you believe that'. Should we say:

BO*r* oughts *r*?

Certainly not; your belief cannot make itself true.

So instead, is the relation:

BO*r* reasons *r*? (9)

This too is wrong. Notice first that it is not very plausible. Suppose you ought not to *r*, and you ought not to believe you ought to *r*, but you do in fact believe you ought to *r*. Then it is not very plausible that you have any reason to *r*, just because of a false belief you ought not to have. Besides this implausibility, I have two arguments against (9), which seem to me conclusive.

First, suppose (9) was true, and suppose you believe you ought to *r*. Then you would have a reason to *r*. So if you had no contrary reason not to *r*, it would be the case that you ought to *r*. Therefore, in the special case where you have no contrary reason, your belief would make it the case that you ought to *r*. But your belief cannot make itself true even in this special case. So (9) cannot be true.

Second, (9) expresses a slack relation, whereas the relation between believing you ought to see to something and seeing to it must actually be strict. If you believe you ought to see to it that *r*, but you do not see to it, you are definitely failing in one respect. You are definitely not entirely as you ought to be. But if the relation between believing you ought to see to it that *r* and seeing to it that *r* was (9), you would not necessarily be failing at all. You might be perfectly as you ought to be. It might be that, though you have a reason to see to it that *r*, you have a better reason not to see to it that *r*. Then you are right not to do so.

So the relation we are after is neither the oughts nor the reasons relation. It must be the relation of normative requirement:

BO*r* requires *r*. (10)

The second argument I gave exemplifies a very useful test for distinguishing a normative requirement from a reason. The reasons relation is slack, so if a normative relation is strict (and it is not the oughts relation) it must be normative requirement. I shall call this the 'strictness test'. (9) may seem attractive at first; you might think 'Surely my believing I ought to *r* gives me *some*

reason to *r*'. But once you see the connection is strict, you should be better satisfied by (10) than by (9).

5. Second example: theoretical reasoning

The examples of normative requirement I am most concerned with come from the context of reasoning. I shall start with theoretical reasoning. Suppose a proposition q follows from a proposition p by a valid inference. That is to say:

$$p \vdash q.$$

Now suppose you believe p. Then a process of correct reasoning will bring you to believe q. (Assume the inference is immediate; I am not concerned with cases where you might reasonably fail to make an inference because it is difficult.) However, it is not necessarily the case that you ought to believe q, nor that you have a reason to believe q. For example, suppose you ought not to believe p, though you do. Then it plainly may not be the case that you ought to believe q or that you have a reason to believe q. So we cannot say either that:

$$\begin{array}{ll} & \mathrm{B}p \text{ oughts } \mathrm{B}q \\ \text{or that} & \mathrm{B}p \text{ reasons } \mathrm{B}q. \end{array} \qquad (11)$$

To reinforce this point, remember that p itself is a consequence of p. A belief in p is plainly not self-justifying, so it cannot be that either:

$$\begin{array}{ll} & \mathrm{B}p \text{ oughts } \mathrm{B}p \\ \text{or} & \mathrm{B}p \text{ reasons } \mathrm{B}p. \end{array}$$

Furthermore, we can apply the strictness test to rule out the reasons relation (11). The relation between believing p and believing q is strict. If you believe p but not q, you are definitely not entirely as you ought to be. But (11) expresses a slack relation; according to (11) you might believe p and not believe q, and yet be entirely as you ought to be, because you may have a stronger reasons not to believe q. This is not possible. To be sure, you might have a good reason not to believe q, and an appropriate response might be to stop believing p. That way, you can escape from the requirement that is imposed on you by your believing p. But if you do believe p and yet do not believe q, you are not entirely as you ought to be. So the relation is strict, which means it cannot be the reasons relation.

The upshot is that the relation between believing something and believing its consequence is another normative requirement:

Bp requires Bq. (12)

It is tempting to say: surely your belief in p gives you *some* reason to believe q. By now we see the relation is strict, this temptation should be better satisfied by accepting (12).

To generalize, a process of theoretical reasoning sets out from existing beliefs of yours and concludes in a new belief. The contents of your beliefs are propositions, and the content of the reasoning is a sequence of propositions. I shall call the process of reasoning 'correct' if its content constitutes a valid inference. If reasoning is correct, the propositions that constitute its content stand in a particular relation to each other: the relation such that the conclusion is validly derivable from the premises. The relation of normative requirement that holds between the beliefs mirrors this relation of inference that holds between the belief's contents. If one proposition follows from others, then believing that proposition is normatively required by believing the others.

6. Third example: practical reasoning

What I have said about theoretical reasoning goes for practical reasoning too. However, since the nature of practical reasoning is contested, to explain this point I shall first have to outline the process of practical reasoning as I see it. I shall stick to instrumental reasoning only. My account is set out more fully in my 'Practical reasoning'.[5]

Here is an example of practical, instrumental reasoning:

I am going to open the wine (13a)
and In order to open the wine, I must fetch the corkscrew, (13b)
so I shall fetch the corkscrew. (13c)

I mean (13a) to express an intention of yours, rather than a belief. I mean (13b) to express a belief. I mean the conclusion (13c) also to express an intention.

[5] Unpublished typescript.

You might actually go through this process of reasoning. Suppose you intend to open the wine, and then someone tells you the corkscrew is in the kitchen. This information imparts to you the belief that in order to open the wine, you must fetch the corkscrew. By reasoning, you form the intention of fetching the corkscrew.

Forming an intention this way is making a decision. Making a decision is as close to acting as reasoning can possibly get you. Reasoning could not actually get you to act, because acting requires more than reasoning ability. So this example of reasoning is as practical as reasoning can be.

Like all reasoning, this reasoning takes you from existing states of mind to a new one. Specifically, it takes you from an intention and a belief to a new intention. To describe the process in more detail, I shall assume your intentions and beliefs are propositional attitudes. That is to say, they are states of mind that have contents, and the contents are propositions. I shall assume your name is 'Pat', and I shall assume the proposition that Pat will open the wine is the same as the proposition that you, Pat, would express by saying 'I am going to open the wine'. So the content of your intention expressed in (13a) is the proposition that Pat will open the wine. Using the third person, and writing 'I' for 'you intend that', we can describe your reasoning process explicitly as follows:

	I(Pat will open the wine)	(14a)
and	B(In order for Pat to open the wine, Pat must fetch the corkscrew)	(14b)
leads to	I(Pat will fetch the corkscrew).	(14c)

This describes your reasoning. It is not a derivation. If you intend to open the wine, and if you believe that in order to do so you must fetch the corkscrew, it does not follow that you intend to fetch the corkscrew. You might not have this intention if you are irrational, for instance.

On the other hand, (13) sets out the content of your reasoning. It has the form of a genuine derivation. Intuitively, it is correct reasoning; you are right to derive the intention of fetching the corkscrew from your existing intention and belief. Furthermore, it is indeed correct reasoning, as the following argument shows.

Compare this process of theoretical reasoning:

	B(Pat will open the wine)	(15a)

and B(In order for Pat to open the wine, Pat must fetch the corkscrew) (15b)
leads to B(Pat will fetch the corkscrew). (15c)

Again, this is a description of reasoning rather than a derivation. Its content is:

Pat will open the wine (16a)
and In order for Pat to open the wine, Pat must fetch the corkscrew, (16b)
so Pat will fetch the corkscrew. (16c)

If you were running through this piece of reasoning, you would doubtless express it to yourself in the first person:

I shall open the wine
and In order for me to open the wine, I must fetch the corkscrew,
so I shall fetch the corkscrew.

(To make it clear that this is theoretical reasoning concluding in a belief, rather than practical reasoning concluding in an intention, imagine you are predicting your movements tomorrow, when you expect to be in a stupor.) This content constitutes a valid inference. The modality in (16b) is not needed for validity, but it does not invalidate the inference. If the premises are true, the conclusion is true too. For this reason, the theoretical reasoning described in (15) is correct reasoning. It is correct because its content constitutes a valid derivation.

The content of the practical reasoning (14) is the same syllogism (16) as the content of theoretical reasoning (15). The difference between the theoretical and the practical reasoning is not in the propositions that constitute their content, but in the attitude you take towards these propositions. In (15) your attitude is to take both premises as true. Because the conclusion is true if the premises are true, you cannot rationally take the premises as true without taking the conclusion as true. So your attitude towards the premises normatively requires you to take the conclusion as true. It requires you to believe it. In (14) your attitude towards the first premise is to set yourself to make it true. Towards the second premise, your attitude is to take it as true. Because the conclusion is true if the premises are true, you cannot rationally set yourself to make the first premise true, and take the second as true, without setting yourself to make

the conclusion true. Your attitude towards the premises normatively requires you to set yourself to make the conclusion true. It requires you to intend it.[6]

Both (14) and (15) correctly track truth through the valid derivation (16). (15) tracks it in a truth-taking way; (14) in a truth-making way. Both therefore constitute correct reasoning. (15) is correct theoretical reasoning; (14) correct practical reasoning.

I shall use the term 'conclusion' sloppily for the mental state that results from a piece of reasoning, as well as for its content, and the term 'premise' for a mental state from which reasoning sets out, as well as for its content. So the premises of (14) are the intention of opening the wine and the belief that to open the wine you need to fetch the corkscrew. The conclusion is the intention of fetching the corkscrew.

The premises of practical reasoning normatively require the conclusion. For instance:

	I(Pat will open the wine)
and	B(In order for Pat to open the wine, Pat must fetch the corkscrew)
requires	I(Pat will fetch the corkscrew).

The relation between premises and conclusion is not the reasons relation. You might be tempted to think that, in some sense or other, you have a reason to intend to fetch the corkscrew, if you intend to open the wine and believe that to do so you must fetch the corkscrew. But to see this is wrong, we can once again apply the strictness test. If you intend to open the wine, and believe that to do so you must fetch the corkscrew, you are definitely not entirely as you ought to be unless you intend to fetch the corkscrew. So the relation is strict, and therefore not the reasons relation.

Instead, it is the requirement relation. Consequently, no normative conclusion can be detached. Even if you intend to open the wine, and believe that to do so you must fetch the corkscrew, it does not follow that you ought to intend to fetch the corkscrew or that you have a reason to. Suppose, say, that you ought not to intend opening the wine in the first place, even though you do

[6] I am oversimplifying when I identify intending with the attitude of being set to make true. Not all cases of being set to make true are cases of intending. So here I am skating over some complications. Details are in my 'Practical reasoning'.

intend to open it. Then it may not be the case that you ought to fetch the corkscrew or have a reason to do so.

In general, intending an end normatively requires you to intend what you believe to be a necessary means. It does not give you a reason to intend what you believe to be a necessary means.

7. Reasons versus normative requirements

The examples show we must recognize normative relations that do not permit a normative conclusion to be detached in the way (2) and (4) do. This is an elementary and widely recognized point,[7] but also one that is widely ignored. The relation of normative requirement is very often confused with the relation of being a reason for. For example, it is very commonly said that rationality consists in acting and believing for reasons. Indeed, most of the literature on rationality is about reasons: it asks what is a reason for what. But actually a large part of rationality consists in conforming to normative requirements, and is not concerned with reasons at all. For instance, one part of rationality is doing what you believe you ought to do, and this does not necessarily mean acting for reasons. Another part is reasoning correctly. Correct reasoning will lead you to have beliefs and intentions that you are normatively required to have by others of your beliefs and intentions. But it may not lead you to beliefs and intentions you have reason to have.

Why the confusion? I think one explanation is that the reasons relation and the requires relation are both, in a sense, weakenings of the oughts relation. The table in section 3 shows that. It is easy to confuse the two weakenings. For example, suppose you notice there must be a normative connection between believing something and believing one of its consequences. But suppose you also notice that believing something does not make it the case that you ought to believe its consequence; the oughts relation is too strong. You need something weaker, and the reasons relation may be the first weakening you think of. Surely, you say, your belief must give you *some* sort of a reason to believe its consequence. But once you recognize that the requires relation constitutes an alternative weakening, I hope you will see it is the appropriate one for this case.

[7] Compare Robert Nozick, *The Normative Theory of Individual Choice* (Garland, 1990), pp. 94–8, which was written in 1963. The point receives a correct emphasis in Harry Gensler's *Formal Ethics* (Routledge, 1996), Chapter 3.

Another source of confusion is an idiom of English. When a conditional proposition contains a modality, we tend to attach the modality to the consequent, even if that is logically not the correct place for it. We say, for instance, 'If it is raining, it must be thawing'. We do not mean, 'If it is raining, necessarily it is thawing', but, 'Necessarily, if it is raining, it is thawing'. Moreover, the antecedent may be implicit and not even stated: we notice the rain and say simply, 'It must be thawing'. Similarly, the proposition that believing *p* normatively requires you to believe *q* would be idiomatically expressed by: 'If you believe *p*, you should believe *q*'. Sometimes we might say simply, 'You should believe *q*', leaving the antecedent implicit – for instance, if your belief in *p* has already been established in our conversation. These expressions are good idiomatic English, but they misrepresent the logic of what is said.

Why does it matter? In the rest of this chapter I hope to demonstrate the importance of the distinction between reasons and normative requirements.

8. Conflicts

Conflicts between reasons are quite different from conflicts between normative requirements.

Take a conflict between reasons first. Suppose the true proposition *p* is a reason for you to *q*, and the true proposition *r* is a reason for you not to *q*. For example, that it is a sunny day is a reason for you to walk to the office; that you feet are sore is a reason for you not to walk but drive instead. These reasons need to be weighed against each other and any other relevant reasons you might have, in order to determine whether or not you ought to walk. Weighing is just what reasons are made for; this is what it means for them to be *pro tanto*. When reasons conflict, everything is in order. The conflict is settled by weighing.

But now suppose the true proposition *p* normatively requires you to *q* and the true proposition *r* normatively requires you not to *q*. Then something is wrong. This is because a normative requirement is strict. For example, suppose you believe that platypuses produce milk and all creatures that produce milk are mammals. These beliefs normatively require you to believe that platypuses are mammals. But suppose you also believe platypuses lay eggs and no mammals lay eggs. These beliefs normatively require you not to believe platypuses are mammals. Your original beliefs

place you under conflicting requirements. This shows something is wrong. Indeed, your beliefs are inconsistent with each other.

For another example, suppose you intend to open the wine, and you believe that to do so you must fetch the corkscrew from the kitchen. This intention and belief require you to intend to fetch the corkscrew from the kitchen. But suppose you intend to keep a careful eye on what is going on in the dining room, and believe this cannot be done if you go into the kitchen. This intention and belief require you not to intend to fetch the corkscrew from the kitchen. Your existing intentions and beliefs place you under conflicting requirements. This shows something is wrong. Indeed, your intentions are inconsistent with each other.

When I say something is wrong, I mean you ought not to be in the state you are in. I can put it more formally. Suppose p normatively requires you to q, and r normatively requires you not to q.

$$p \text{ requires } q \tag{17}$$

and $$r \text{ requires } \neg q. \tag{18}$$

Then you ought to see to it that p and r are not both true:

$$\mathrm{O}(\neg(p \,\&\, r)). \tag{19}$$

It must be a feature of the logic of normative requirement that (19) is derivable from (17) and (18). Indeed, given some axioms of deontic logic, (19) can be derived from $\mathrm{O}(p \rightarrow q)$ and $\mathrm{O}(r \rightarrow \neg q)$, which are respectively the logical factors of (17) and (18).

If you are under conflicting normative requirements, you are not as you ought to be; something is wrong with your condition. That is what (19) tells us. Normally, you ought to go back and sort things out. You should sort out your inconsistent beliefs or intentions, for example. This is quite different from the appropriate response to a conflict of reasons. Conflicting reasons require no sorting out, but simply weighing against each other.

What about conflicts between reasons and normative requirements? There is no contest. Reasons are concerned with what you ought to see to, and normative requirements are not. For example, suppose the balance of reasons is in favour of your seeing to it that q, but you believe you ought not to see to it that q. Then you ought to see to it that q, because the reasons together determine what you ought to see to. Your contrary belief normatively requires you not to see to it that q, but this does not count at all in determining what you ought to see to.

9. The practicality of practical reasoning

Now I come to the first of three examples of mistakes that have been made as a result of confusing normative requirements with reasons. It is a mistake of my own. I mention it only because it helps to illuminate the distinction between a reason and a normative requirement.

I used to be puzzled by a claim of G. E. Moore's.[8] Moore thought that, when you have to choose between alternative acts available to you, you ought to choose the one that will have the best results. However, since you can never know for certain what the results of any act will be, you can never know for certain which of the acts available to you will have the best results. Consequently, Moore claimed that you can never know for certain what you ought to do.

I found that puzzling. Our rationality must be practical, so it must be able to engage with the predicament we find ourselves in in practice. One feature of our predicament is that we cannot know for certain what the results of our acts will be. Our rationality must be able to cope with this uncertainty. It must be able to determine an appropriate way of coping with it. So surely there is something you ought to do, *given* the uncertainty you are faced with, and surely you might know what it is.

In arguing against Moore, I imagined you trying to decide whether or not to go sailing. The benefits of sailing depend on the weather. You assign probabilities to the various possible states of the weather. I assumed these probabilities are such that the expected benefit of sailing is greater than the expected benefit of staying at home. Given that, I claimed you ought to go sailing. But Moore would have denied this is necessarily so. I said:

> The conclusion of practical reasoning is a judgement of what ought to be done. And, also, it has to be a judgement one can act on. But the only judgement one can act on is relative to the probabilities available. Suppose, when wondering whether to go sailing, having consulted the sky and the weather forecast, you now consult Moore. You ask him what you should do. He replies that you cannot possibly know what you should do, but that probably you should go sailing. You, though, need to know what to do. Impatient with Moore's shilly-shallying, you

[8] G. E. Moore, *Principia Ethica*, Cambridge University Press, p. 149.

> ask him 'So what do you suggest I do, then?' Pressed like this, Moore will certainly tell you to go sailing. This is not simply an ungrounded whim on his part. He believes that practical reason, given the probabilities, requires you to go sailing; it would be irrational on your part not to. Another way of expressing this belief of Moore's is that you ought to go sailing. If it is irrational for you not to go sailing, then you ought to go sailing.[9]

This passage of mine is full of errors, whereas the attitudes I attributed to Moore are correct. Moore is right to believe that practical reasoning requires you to go sailing. (Strictly, it is your probability assignments and your beliefs about the benefits of sailing in various weathers that require you to go sailing. Practical reasoning is strictly only the instrument by which you discover this normative requirement, rather than the source of the requirement itself. But this is a minor slip.) Moore is also right to tell you to go sailing; this is the decision normatively required by your beliefs. But, contrary to what I said, it does not follow that you ought to go sailing. Contrary to what I said, the conclusion of practical reasoning is not a judgement of what ought to be done, but a decision to do something. My argument against Moore was mistaken throughout.

I now see Moore might be right when he says we can never know what we ought to do. I do not insist that he is right, and certainly not that he is right for the right reason, but I think he might be right. If he is, rationality will still be practical. Practical reasoning will still be possible, because the conclusion of practical reasoning is not a belief about what we ought to do. We will still be able to reason properly in the face of uncertainty, and arrive by reasoning at intentions that are normatively required.

In responding to a view like Moore's, people sometimes call on the idea of a 'subjective ought'. About my example, they would say that, whatever you ought objectively to do – and you do not know – subjectively you ought to go sailing. I think this amounts to just another way of saying that your beliefs and probability assignments normatively require you to go sailing. I think 'subjective ought' is really just an alternative term for 'normative requirement'. If so, it is an unsatisfactory term for two reasons. First, it conceals the logical structure of the situation, because it does not make the 'ought' govern a conditional. Second, it

[9] John Broome, *Weighing Goods*, Blackwell, 1991, p. 128.

implicitly makes the ought relative to the wrong thing. It makes it relative to the subject, whereas it should be relative to a fact: the fact that imposes the normative requirement. In this example, it should be relative to your state of belief.

This second inaccuracy will be particularly conspicuous if you have inconsistent beliefs or intentions. Then it may happen that some of your beliefs and intentions normatively require you to see to something, and others normatively require you not to see to it. This is a comprehensible feature of your inconsistent condition. But it is not comprehensible to say you subjectively ought to see to something and also you subjectively ought not to see to it; this looks like a contradiction. I think it is best to avoid the notion of subjective ought.

10. The tortoise's mistake

My next example of a mistake is made by the tortoise in Simon Blackburn's 'Practical tortoise raising'.[10] In a discussion with Achilles, this tortoise impugns the cogency of instrumental reasoning.[11]

The discussion starts with Achilles's offering the tortoise this inference as an example of instrumental reasoning:

	You want the lettuce	(20a)
and	If you want the lettuce, you must cross the road,	(20b)
so	You must cross the road	(20c)

The tortoise objects on the grounds that (20b) is 'one of those off-colour conditionals where musts and oughts make the conclusion non-detachable', so that (20c) does not follow.

His point is this. (20b) superficially seems to mean:

If you want the lettuce, necessarily you cross the road. (21)

(20c), which means 'Necessarily you cross the road', can be validly derived from (20a) and (21). But (21) is plainly false, and not what (20b) really means. It really means:

Necessarily, if you want the lettuce, you cross the road. (22)

(20c) cannot be validly derived from (20a) and (22). So the tortoise is undoubtedly right to object to the syllogism (20). I can

[10] *Mind*, 104 (1995), pp. 695–711.
[11] pp. 708–10.

add the further objection that (22) is false. It is perfectly possible for the tortoise to want the lettuce and yet not cross the road.

Achilles spoiled his case by giving the tortoise a bad example of instrumental reasoning. We need to start again. As it happens, the tortoise himself supplies the materials for correct instrumental reasoning later in the conversation. He says 'If I am to get the lettuce, I must cross the road'. That is to say, a necessary means of getting the lettuce is to cross the road. So if the tortoise intends to get the lettuce, he can reason:

	I am going to get the lettuce	(23a)
and	In order to get the lettuce, I must cross the road,	(23b)
so	I shall cross the road.	(23c)

Provided the first premise and the conclusion express intentions, and the second premise a belief, this is instrumental reasoning exactly on the model of (13). It is correct instrumental reasoning. If the tortoise intends to get the lettuce, he is normatively required to intend to cross the road.

Certainly, he cannot reach the detached conclusion that he ought to cross the road. Still less, that he must cross the road. But the tortoise needs neither of these conclusions, and neither would be true. He only needs 'I shall cross the road', expressing the intention of crossing the road. This intention may be correctly derived from the premises.

Instrumental reasoning does not lead to any detached normative conclusion for the tortoise, nor place him under any detached necessity. The tortoise seems to assume he is therefore not placed under any requirement of rationality. But he is: rationality requires him to intend whatever he believes to be a necessary means to an end he intends.

In speaking through the tortoise, Simon Blackburn's primary purpose is not to object to instrumental reasoning but to show that 'There is always something else, something that is not under the control of fact and reason, which has to be given as a brute extra, if deliberation is ever to end by determining the will.'[12] I have not disagreed with that. In my example (23) of instrumental reasoning, the tortoise's will is already determined to get the lettuce. Instrumental reasoning comes into play once the will is determined on a particular end. Its effect is to determine the will to take a means to the end it is already determined on. Since it

[12] p. 695.

comes into play only when the will is already determined, Blackburn has no real need to object to instrumental reasoning.

Blackburn's example (20) of putative instrumental reasoning suggests to me that he thinks instrumental reasoning should take you from wanting an end to intending (having your will determined on) a means. But this is to conflate two separate steps. There is first the step from wanting an end to intending the end, and second the step from intending the end to intending the means. The second step is negotiated by instrumental reasoning proper, and is entirely under the control of reason. I have nothing to say about the first step in this chapter.

11. Korsgaard's mistake

Another example of a mistake comes from Christine Korsgaard's 'The normativity of instrumental reason'.[13] Korsgaard is concerned with how instrumental reasoning gets its rational authority over our actions. How does the pursuit of an end require you to take a means to it? The conclusion she draws is that unless the end itself is invested with normative force, you cannot be normatively required to take the means. 'Unless there are normative principles directing us to the adoption of certain ends, there can be no requirement to take the means to our ends.'[14] Instrumental reasoning could transmit normativity from the end to the means, but it cannot itself give the means normativity.

However, Korsgaard is also opposed to the realist view that some ends have normativity in the nature of things – that it is in the nature of these ends that they should be pursued. So she thinks the normativity must arise from a decision to pursue the end. She says:

> For the instrumental principle ['that practical reason requires us to take the means to our ends'[15]] to provide you with a reason [to take the means to an end], you must think that the fact that you will an end *is a reason* for the end. It's not exactly that there has to be a *further* reason; it's just that you must take the act of your own will to be normative for you. . . . [This] means that your willing the end gives it a normative status for you,

[13] In *Ethics and Practical Reason*, edited by Garrett Cullity and Berys Gaut (Oxford University Press, 1997), pp. 215–54.

[14] p. 220.

[15] This definition of the instrumental principle is implicit on p. 215.

> that your willing the end in a sense makes it good. The instrumental principle can only be normative if we take ourselves to be capable of giving laws to ourselves – or, in Kant's own phrase, if we take our own wills to be *legislative*.[16]

No doubt Korsgaard is right that you must take yourself to have a reason for your end *if* instrumental reasoning is to provide you with a reason to take a means. And if instrumental reasoning is to do this for you, I dare say her other conclusions would follow. But instrumental reasoning does not provide you with a reason to take a means. That is not how it works. Willing (or intending) an end normatively requires you to will whatever you believe is a necessary means to the end. I explained in section 6 how this normative requirement arises. Willing the end does not give you a reason to take the means, and it does not need to. So actually Korsgaard's conclusions do not follow. Willing an end need not give the end a normative status for you. Moreover, you can will an end without taking it as a law for yourself. You can simply decide to pursue it on one occasion.

Korsgaard's mistake illustrates an important feature of normative requirements. Reasoning is possible even in conditions that are unfavourable in a particular way. In your reasoning, you can take as premises beliefs and intentions you have no reason to have, and even beliefs and intentions you ought not to have. The nature of your reasoning is unaffected by whether or not you ought to have the beliefs and intentions it is premised on. Instrumental reasoning brings you to take appropriate means to your ends, and it is not paralysed if your ends happen to be ones you should not have. Similarly, your theoretical reasoning works well and in the same way, whether or not it is premised on beliefs you should not have.

How is this possible? It would not be possible if reasoning had to generate a reason for its conclusion. But in fact reasoning simply determines a normative requirement: that the premises require the conclusion. That is how reasoning is possible in unfavourable conditions.

12. Summary

Rationality does not consist entirely in acting for good reasons, as is commonly supposed. To a large extent it consists in following

[16] pp. 245–6.

normative requirements. Consequently, rationality may bring you to do things you have no reason to do.

In particular, correct reasoning – both theoretical and practical – constitutes a major part of rationality, and correct reasoning is governed by normative requirements rather than by reasons. If it were otherwise, we could not reason in unfavourable conditions, on the basis of premises we have no reason to hold.

5

NON-COGNITIVISM, NORMATIVITY, BELIEF

Frank Jackson

Non-cognitivism in ethics and non-cognitivism about rationality go hand in hand. One way to argue the point is to insist that, at some level, what it is ethical to do is also what it is rational to do. A less controversial way to make it is to note that the famous arguments for non-cognitivism in ethics apply *mutatis mutandis* to rationality; they work for ethics if and only if they work for rationality. As the point is widely accepted, I will make it briefly.[1] First, the open question argument has equal putative force in both contexts. Just as no amount of information couched in purely descriptive terms seems to close the question as to what it is morally right to do, so no amount of information couched in purely descriptive terms seems to close the question of what it is rational to believe or do. It always seems to make sense to ask, even after all the descriptive information is in, what one ought to do, independently of whether the 'ought' is given a moral or rational slant. Secondly, the argument from the persistence of disagreement has equal apparent force in both cases. Two people can agree on *all* the facts there are to agree on and yet, it seems, still have room to disagree about what is rational or moral. Finally, there seems to be an internal connection between judgements of rationality and morality, on the one hand, and action, on the other, of a kind that Hume argued cannot hold between judgements properly called, i.e. beliefs, and action. The judgement that an action is rational and the judgement that it is right both seem to point, by their very nature, to doing that thing. They may not point hard enough to get the rational or right result, but they seem to have an inbuilt 'direction' that Hume argued beliefs *per se* cannot have. Of course, many, myself included

[1] See, e.g., Simon Blackburn, *Spreading the Word* (Oxford: Oxford University Press, 1984), and Allan Gibbard, *Wise Choices, Apt Feelings* (Cambridge, Mass.: Harvard University Press, 1990).

as it happens, think that these arguments fail, which is why we are cognitivists. But it remains the case that if they work in the one case, then they work in the other as well.

Thus, to the extent that we should take non-cognitivism in ethics seriously – to the considerable extent, many urge – we should take non-cognitivism about rationality and normativity seriously. In this chapter, I raise what seems to me a serious, though essentially simple, problem for non-cognitivism about rationality. But my reason for highlighting the connection between non-cognitivism in ethics and non-cognitivism about rationality is not simply to identify my target as one worth shooting at. I want there to be guilt by association. I want to strengthen the case for cognitivism in ethics. However, the discussion to follow will be framed for the case of rationality and normativity alone.

The problem in a nutshell

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*. More generally, people ought to believe the fairly obvious consequences of what they believe. Likewise, people ought not to have inconsistent sets of beliefs; internal consistency is a normative constraint on belief.

These normativity claims hold independently of whether or not we should accept the Lewis-Stalnaker account, framed in terms of possible worlds, of belief contents, although the best way to describe the claims will be affected.[2] For example, on their account, subjects who believe that P and that if P then Q, automatically believe that Q – because every world where P, and where if P then Q, is a world where Q – *provided* that they have a single system of belief. On their account, cases where we want to say that a subject believes that P, and believes that if P then Q, while failing to believe that Q (as opposed to failing to realise that 'Q' captures what they believe) are cases where the subject has two systems of belief, according to one of which it is the case that P, and, according to the other, it is the case that if P then Q. The normative constraint is then describable as that

[2] See, e.g., Robert Stalnaker, *Inquiry* (Cambridge, Mass.: MIT Press, 1984) and David Lewis, *On the Plurality of Worlds* (Oxford: Blackwell, 1986).

a subject ought to put together aright his or her different systems of belief.

Also, there are normative constraints on the evolution of belief under the impact of evidence. Someone who believes that a certain hypothesis has survived many varied tests ought to believe the hypothesis. Someone who believes that a certain hypothesis explains the observed data ought to believe that hypothesis, or anyway increase their degree of belief in it. Someone who believes that all observed ravens are black ought to believe that the next raven to be observed will be black. And so on. Of course, how precisely to describe the way in which belief ought to evolve is highly controversial, but it is relatively non-controversial that it is not Rafferty's rules. There are right ways and wrong ways for belief to evolve. This follows from the persuasive, familiar idea that belief is a state directed to fitting the world (by contrast with desire, which is directed towards getting the world to fit it). Part of what it is to be directed towards fitting the world is tending to evolve in the right way.

The problem for non-cognitivism about normativity, in a nutshell, is that if non-cognitivism about normativity is correct, there is no such thing as *satisfying* normative constraints, and therefore, it would seem, no such thing as belief. For to satisfy a constraint is to have the relevant property. It is a constraint on being President of the United States that one be born in the United States; it is not enough to be a citizen. This means that you cannot be President unless you are born in the United States. Likewise, if belief is subject to normative constraints, being a believer requires that one have the relevant normative properties. For example, if someone believes that P and believes that if P then Q, then they have the property of being such that they ought to believe that Q. Again, for any believer, there is a way their beliefs ought to evolve under the impact of putative information. But if non-cognitivism is true, there are no normative properties to have, or to fail to have. Non-cognitivism is precisely the view that, although the language of normativity is meaningful, there are no properties corresponding to normative predicates. Equivalently, non-cognitivism about normativity is the view that normative predicates when attached to subject terms do not serve to make claims about how the subjects are; in consequence, there is no such thing as being how things are claimed to be by normative predicates and normative language in general. But then there is no such thing as subjects being how normative

predicates say that they are, and so no satisfying the normative constraints on being a believer. By Modus Tollens, therefore, we reach the conclusion that there are no believers and no beliefs.

The point can be made in terms of the familiar 'boo-hurrah' model for understanding non-cognitivism in ethics. Just as there is no such thing as satisfying 'is boo' or 'is hurrah', so there is no such thing as satisfying 'ought to be such and such', according to non-cognitivism. But then there can be no believers, for there are no things that satisfy the constraints necessary for being a believer. I take this to be a *reductio* of non-cognitivism about normativity.

Before I turn to consider various responses that might be made to this argument against non-cognitivism, there are three things to note to avoid possible confusions. First, talk of satisfying normative constraints can be read two ways. One thing that would often be meant by satisfying the normative constraint that I ought to believe that Q would be that I *in fact* believe that Q. In this sense, to satisfy a normative constraint is to have, as a matter of fact, the property one ought to have, or, equivalently, to obey the constraint. This is not what I mean here. I mean by satisfying a normative constraint having the normative property in question – in the case given, having the property of being such that I ought to believe that Q – the 'property' that is not a property at all according to non-cognitivism. Secondly, we can here set aside the debate over whether it is analytic that believers are, by and large, rational. Many hold that it is analytic that believers are, by and large, rational. For example, Donald Davidson and Daniel Dennett have argued that it is constitutive of our notion of being a believer that believers are largely rational. Their argument is that it is essential to being a believer that one be interpretable as such, and interpretation requires presuming that what one is interpreting is, to a considerable extent, rational. For example, the project of moving from subjects' behaviour to what they believe and desire, necessarily presumes that their behaviour, by and large, satisfies their desires if their beliefs are true.[3] Our argument, however, is based on the less contentious claim that there are conditions that believers ought to satisfy, independently of whether it is analytic, or constitutive of being a believer, that they

[3] See, e.g., Donald Davidson, *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984) and Daniel Dennett, *Brainstorms* (Montgomery, VT: Bradford Books, 1978).

typically satisfy them. Finally, I will be taking normativity and rationality to be inter-definable in the obvious way: rationality is conformity to norms. In consequence, I will treat the issue of non-cognitivism about normativity as one and the same as the issue of non-cognitivism about rationality.

No non-cognitivist about normativity worth his or her salt will accept the quick *reductio* offered two paragraphs back. The remainder of this chapter is concerned with various replies that he or she might make, with the exception of the reply that embraces the conclusion of the Modus Tollens. I will assume that eliminativism about believers, though a highly interesting position, is a manifestly false one.

There seem to me to be four rather different responses a non-cognitivist might make to our objection to non-cognitivism. One is to embrace minimalism about what it is to satisfy a normative constraint. A second response is to give a non-cognitivist construal of belief or of what it is to satisfy a normative constraint. (This is perhaps the most obvious response, but, I will argue, it is an untenable one.) A third is to offer a reductive account of the normativity constraints on being a believer. The final response is to deny the datum; to deny that there are normative constraints on believing. I will argue against these four responses in turn.

Minimalism

We raised our problem by saying (a) that believers must satisfy normative constraints, but (b) that, according to non-cognitivism, there is no such thing as *satisfying* normative constraints (in the sense of being such that one ought to phi, not in the sense of being such that one phi's). The minimalist response is that we are wrongly presuming a substantive notion of what it is to satisfy a constraint. We are presuming a picture of the debate between cognitivism and non-cognitivism that thinks that, when we have a predicate in English that is meaningful and which forms sentences that have all the marks of truth aptness when it is appended to subject term – they can have the predicate 'is true' appended without grammatical impropriety, they are declarative, they can appear as the antecedents of conditionals etc. –there is a *further* question as to whether the predicate ascribes a property.

No doubt, continues the minimalist response, non-cognitivists themselves often presume the mistaken picture. They often frame their view as being that predicates like 'is rational' and 'is

right' differ from a predicate like 'is square' in that only the latter, when appended to a subject term, ascribes a property in the sense of making a claim about or report on how the subject of the sentence is. But this is a mistake, argues the minimalist. The best that can be meant by saying that 'is square' ascribes a property is that sentences of the form 'X is square' are meaningful and have all the marks of truth aptness listed above. But then 'is rational', 'ought to believe that Q' and 'is right' ascribe properties, for it is certainly true that 'X is rational', 'S ought to believe that Q' and 'X is right' are meaningful, are syntactically well-formed and have all the marks of truth aptness.

The idea behind this response will be familiar from recent discussions of redundancy theories of truth, and of minimalism about truth and truth aptness in general.[4] I have two replies to it. The first is that it makes it very hard to say what non-cognitivism, the doctrine, *is*. The second reply contests the minimalist picture that underpins the response.

Non-cognitivism, in the sense in which it is an interesting new position about rationality and ethics first drawn to our attention by A.J. Ayer in its emotivist guise, is to be sharply distinguished from various versions of subjectivism.[5] Ayer made much of the fact that his view was no kind of subjectivism, and in this he is followed by Blackburn (*Spreading the Word*, p. 169) and Gibbard (*Wise Choices, Apt Feelings*, p. 153), and for very good reason. The famous arguments for non-cognitivism mentioned at the very beginning of this chapter apply against subjectivism if they apply anywhere. To make the point with one example: if subjectivism is true, agreement about all the facts is not consistent with disagreement about what is rational or right, for all the facts will include all there is to know about subjects' desires and attitudes. Thus, the argument from the persistence of disagreement is as much an argument against subjectivism as one against cognitivism in general. However, in order to distinguish non-cognitivism from subjectivism, we need to have recourse to a claim about language, as non-cognitivists themselves emphasise (e.g. Ayer, *Language, Truth and Logic*, pp. 104–5 and Gibbard, *Wise Choices, Apt Feelings*, p. 153). Subjectivism holds that sentences

[4] See, e.g., Paul Boghossian, 'The Status of Content', *Philosophical Review*, 99 (1990): 157-84. I take the term 'truth apt' from Crispin Wright, *Truth and Objectivity* (Cambridge, Mass.: Harvard University Press, 1992).

[5] A.J. Ayer, *Language, Truth and Logic* (London: Gollancz, 1936).

like 'X is rational' and 'X ought to be done' are truth apt, being true just when the producers of the sentences have (or are disposed to have in ideal circumstances, or whatever more complex story may be told) the appropriate desires or, more generally, attitudes. Subjectivism holds, as it is often put, that ethical and normative sentences make claims about, are reports of, subjects' attitudes. By contrast, according to non-cognitivism, 'X is rational' and 'X ought to be done' are not truth apt, being expressions of, rather than reports or descriptions of, attitudes, as it is often put.[6] This is, of course, why non-cognitivism is often called expressivism. But, according to the kind of minimalism that founds the response we are now discussing, 'X is rational' and 'X is right' must be truth apt – this follows from their syntax and their possession of the marks of truth aptness. The upshot is that the minimalist response destroys the distinctiveness of the doctrine it was designed to protect. The distinction between non-cognitivism and subjectivism which is so important for non-cognitivists requires precisely what minimalism denies is possible, namely, that, despite ethical and normative sentences being declarative and possessing all the marks of truth aptness, they are not in fact truth apt.[7]

I must, though, say immediately that in making these remarks, I run counter to some things Blackburn says in his most recent exposition of his version of non-cognitivism (or expressivism, as he mainly calls it).[8] There he allows that ethical and normative sentences are truth apt, and in particular that 'there are moral truths' (*Ruling Passions*, p. 319). And his motivation seems to be a broad sympathy with minimalism. Is he, then, still a non-cognitivist? At times in *Ruling Passions* he seems to be, at bottom, someone opposed to any kind of reductive analysis of ethical and normative sentences (see esp. *Ruling Passions*, p. 49); as someone, that is, who might belong in G.E. Moore's (cognitivist) camp. However, at a number of places, he uses the terms

[6] Philip Pettit and I argue in 'A Problem for Expressivism', *Analysis*, 58 (1998): 239–51, that there are serious problems for this common way of putting why the sentences in question are not truth apt. We argue, on the basis of considerations in the philosophy of language quite separate from minimalism, that it is much harder than non-cognitivists (expressivists) appreciate for ethical sentences to fail to be truth apt (and the same would go for normative ones, in general). But I here go along, for the purposes of the argument and ease of exposition, with the more usual view.

[7] Which is why a minimalist like Paul Horwich rejects non-cognitivism; see his 'Gibbard's Theory of Norms', *Philosophy and Public Affairs*, 22 (1993): 67–78.

[8] Simon Blackburn, *Ruling Passions* (Oxford: Clarendon Press, 1998).

distinctive of traditional non-cognitivism: 'Valuing something...is not to be understood as *describing* it...' (p. 49); 'Expressivism denies that when we assert values, we talk about our own states of mind... It says that we *voice* our states of mind, but denies that we thereby describe them' (p. 50). In any case, I will mean by non-cognitivism the view that ethical and normative sentences are not truth apt, and, in my view, this is the key to understanding how non-cognitivism comes to be a distinctive new view as opposed to some form of subjectivism or of Moorean anti-reductionism.

I should also say, though, that I reject the minimalism about truth aptness (and property-ascribing predication) that founds the minimalist response. I take a representationalist, Locke-inspired approach to the truth aptness of language.[9] I take it to be obvious that language is most especially a conventional system of physical structures for the transfer of putative information from one language user to another; something travellers in a foreign country whose language they do not understand are forcibly reminded of when they get lost or try to buy something in a shop.

On this view, what makes truth apt sentences truth apt is their role as convention-governed conveyers of the putative information, as convention-generated vehicles of representation. But how a sentence represents things as being is an a posteriori, contingent matter. The putative information I in fact give by using a given sentence might have been given by using a quite different sentence. The great diversity in languages in different parts of the world testifies to how contingent our agreements are. The key, on this approach, to a sentence's being truth apt is the way it is used to convey what users of the sentence believe about how things are. This makes sense of the rationale for the common view that indicative conditional sentences are not truth apt – those who hold this view do so precisely because they hold that indicative conditionals are used in the wrong way to count as conveying how their users take things to be.[10] On the Lockean picture, then, the question as to whether a sentence is truth apt, and correspondingly whether a predicate ascribes a property, is not closed by noting that the sentence is meaningful and has the right syntactical properties. The question turns on whether the sentence is *used* in the right way for truth aptness. Here I am, of

[9] For a fuller account of what follows, see Frank Jackson, 'II – Naturalism and the Fate of the M-Worlds', *Proceedings of the Aristotelian Society*, supp. vol. 71 (1997): 269–82.

[10] See, e.g., Ernest Adams, *The Logic of Conditionals* (Dordrecht: Reidel, 1975).

course, agreeing with what underpins the traditional non-cognitivist position, for non-cognitivists characteristically argue for their view on the basis of various features of the way we use normative sentences and predicates, as we noted at the beginning.

Non-cognitivism about belief

Most people will grant that

> If X is highly rational, then X is rational.

Could we then refute non-cognitivism about rationality by saying that a constraint on being highly rational is being rational, but non-cognitivism implies that there is no such thing as satisfying the constraint of rationality, and therefore it implies that nothing is highly rational? The answer is no. That would be *far* too quick. Non-cognitivists read the antecedent *and* the consequent of the conditional non-cognitively. For them, to grant the conditional amounts to something, very roughly, like granting that if one takes the 'highly rational' attitude to X, then one must at least take the 'rational' attitude to X. The conditional expresses some kind of constraint on behaviour and attitudes, not a claim that if something is the 'highly rational way', then it is the 'rational way'. In the same way, a non-cognitivist about normativity might respond to our problem by holding a non-cognitivist view of belief. If sentences about belief do not make claims about how things are, but instead express the 'belief' attitude, then the normativity of belief can be thought of as a constraint on attitudes, broadly speaking, rather than a constraint on how things must be. The idea would be that instead of thinking of

> If X believes that P, and X believes that if P then Q, then X ought to believe that Q

as saying that anyone who satisfies the description 'believes that P, and believes that if P then Q', also satisfies the description 'ought to believe that Q', we should think of it as prescribing a constraint on the attitudes that can properly be taken. As it might be put, if one takes the believes-that-P and the believes-that-if P-then-Q attitudes to X, it is prescribed that one take the ought-to-believe-that-Q attitude to X.

The basic problem with this line of reply is that non-cognitivism about belief is so implausible. There are, on the market, many accounts of belief that might be described as anti-realist.

They are accounts of belief which make what a person believes, and sometimes whether or not something is a believer at all, depend on how they are interpreted by others, and on the reactions of their language communities to them. These views are anti-realist in the sense that they make belief – both its content and the fact of belief – in part a matter of the responses of others. But these accounts are not non-cognitivist in the sense at issue here. Although, in some versions, they imply that 'X believes that P' lacks a truth value, this is because it will be an incomplete expression. It will be like 'X is tall' in the absence of any indication of a reference class. When suitably completed, perhaps by a specification of a possible interpreter, 'X believes that P' will have a truth value. But I know of no (at all plausible) account of belief that would count as non-cognitivist in the sense relevant here – in the sense, that is, which would mean that 'X believes that P' is in the wrong semantic category to have a truth value. And there is good reason to hold that no such account will be forthcoming, or at least there is for anyone sympathetic to the earlier account of truth aptness that ties it to representational role. For, on that account, the key to sentences' truth aptness lies in their having the right relation to the beliefs of the users of the sentences. Very roughly, to be truth apt in L is to be appropriately used in L to express speakers' beliefs. And it would be *very* strange if the sentences which one used to explain how various sentences get to be truth apt, that is, certain sentences in part about belief, did not themselves have truth values. It would mean that saying that something was an assertion proper in the sense of being truth apt was not itself an assertion proper!

It might be thought that non-cognitivists about normativity could grant that we must offer a cognitivist account of belief – if belief is not to be thought of cognitively, what is? – and respond to our problem by offering a non-cognitivist account of the sense in which there are normative constraints on belief. Our non-cognitivists might, reasonably enough, point out that their position is non-cognitivism about normativity, not about belief, and argue on that basis that it makes perfect sense for them to restrict their claim to the sense in which we should understand what it is to meet a normative constraint. They would, presumably, allow cognitivist accounts of what, in general, it is to meet a constraint; the suggestion would be restricted to how we should understand what it is to meet a normative constraint, a constraint framed in

normative terms.[11] It is that alone which should be understood in terms of the taking up of an attitude, or something in that line of country, rather than in terms of the possession of a feature. I suspect the reason our problem for non-cognitivism about normativity has not been more widely appreciated is an implicit assumption that something like this is a reasonable position for non-cognitivists to adopt. I think, against this, that it has a fatal flaw.

The position accepts the normativity of belief, while, of course, giving a non-cognitivist construal of what this amounts to. This means that its advocates accept that the language of belief and the language of normativity go together; for example, if you use the words 'believes that P, and believes that if P then Q' of someone, you are required to use the words 'ought to believe that Q' of them. Indeed, advocates of the position may well like this way of phrasing what they hold, because it reduces the temptation (as they see matters) to think of the normativity claim as one about features possessed – it portrays the normativity of belief as a constraint on the use of language that does not correspond to a constraint in the nature of that which the language is about. The fatal flaw is that, *as non-cognitivists about normativity*, they must hold that such a position contradicts their cognitivism about belief. In this context, to be a cognitivist about belief is to hold that the language of belief is descriptive. But the inspiration for non-cognitivism about normativity is the conviction – be it based on the open question argument, the persistence of normative disagreement, or the view that there is a connection that obtains between action and normativity judgements that is incompatible with cognitivism – that no amount of description requires or compels a normative judgement. You can describe until the cows come home and it still be completely open to you to refrain from adopting one of the attitudes that it is the business of the normative vocabulary to express: this is the conviction that lies behind non-cognitivism about normativity. But this would mean that, by its own lights, cognitivism about belief implies that there is no tight connection between using the language of belief and normative language. The combination of cognitivism about belief and non-cognitivism about normativity would debar non-cognitivists from explaining how using the language of belief requires using, or being prepared to use, the language of norms.

[11] Where, as always in this chapter, meeting a normative constraint means being such that one ought to phi, not the phi-ing itself

Perhaps the simplest way to see the point is through the corresponding point in the ethical case. Non-cognitivists hold that whether or not X killed Y is a cognitive matter that can be captured in terms of purely descriptive language. And they think that describing some incident in terms of how many people were killed, and why, cannot, in and of itself, require one also to use terms like 'wrong' in connection with it. But the view under discussion would commit non-cognitivists to holding that whether or not X has such and such beliefs is a cognitive matter, and so one capturable in terms of purely descriptive language; but then how can they (*they*) allow that the use of such language in and of itself requires one also to use terms like 'ought to believe that Q'? The datum we started from is that normativity seems part of our very concept of belief: we cannot say what belief is without, in one way or another, using normative terms. This is not true of killing. Although we are convinced that some killings are very wrong, it is plausible, and accepted by non-cognitivists, that our concept of a killing does not have normativity built into it.

It might be suggested that non-cognitivists can think of the language of belief as 'mixed', as having a cognitive part (of course) and a non-cognitive part. In this regard, terms for belief would be like the term 'murder' according to non-cognitivists, namely, roughly analysable into a descriptive part – killing – and a non-cognitive part that expresses a negative attitude towards the killing. It is hard to see how this helps. For, first, on non-cognitivist views, it is never true that anyone murders anyone. It is perhaps sometimes false that someone murders someone, namely, in those cases where they do not kill them. (This would follow from taking certain super-valuational approaches to how non-cognitivists might evaluate 'X murdered Y'.) But it is never true that murder happens, and it is never a fact about the world that someone has committed, is committing or will commit a murder, any more than it is ever a fact about the world that something was, is or will be boo or hooray. Never a fact, that is, in the serious sense of 'fact', and we have seen that non-cognitivists cannot afford to embrace the kind of minimalism that would challenge the coherence of the distinction between the serious and the non-serious sense in which something can be a fact. I said earlier that it would be very strange if the sentences which one used to explain how various sentences get to be truth apt, that is, certain sentences in part about belief, did not themselves have truth values. It is hardly an improvement to say that the sentences we use to explain

how sentences can get to be true, namely, certain sentences containing the word 'belief', cannot themselves be true.

A second problem is a more subtle one. The idea under consideration is that non-cognitivists might analyse talk of belief into talk of a 'colourless' basic representational state and a 'coloured' attitude towards that state. Thus, they might analyse an assertion that X believes that P as both saying that X believes* that P, and as expressing an attitude, where belief* is a purely representational state with no normative dimension. But how could belief*, if it serves to represent how things are, lack a normative dimension? Are we supposed to say that there is nothing wrong with representing that P and that if P then Q, while refraining from representing that Q? Once we have states of subjects that seek to represent how things are, we have states that are constrained by the aim of truth, states that seek to conform to the world, and, therefore, ones such that subjects ought to avoid obvious inconsistency regarding them, and ones that ought to respond to new information in the right way. In other words, the original problem can as easily be raised for belief* as for belief.

Finally, we should note that such a position on belief would be an internally inconsistent one for many non-cognitivists. I have in mind those non-cognitivists who argue for their position by urging that normative (and ethical) judgements have a connection with action that no belief (in their words!) can have.

Reductionism about the normativity of belief

Some uses of the normative terms are easily translated away. 'If you want to be fresh in the morning, you ought to go home now', seems simply to be another way of saying that if you do not go home now, you will not be fresh in the morning. (Or something along these lines; the detail is not important for the point at hand.) More generally, it is often suggested that many normative terms are, in effect, hypothetical imperatives which can be translated out more or less mechanically. The hard cases are those which are, in effect, categorical imperatives. It might, therefore, be suggested that the uses of normative terms in connection with belief are all of the easy variety. Although we naturally express certain constraints on belief in normative terms, we could as easily express them in purely descriptive terms.

For example, the rule that if one believes that P, and believes that if P then Q, one ought to believe that Q, it might be suggested, is

really nothing more than the purely descriptive claim that if someone does not believe that Q in such a case, they fail to believe a fairly obvious consequence of things they believe. However, we only need to state the suggestion to see its error. It leaves out the fact that there is something *wrong* with failing to believe fairly obvious consequences of other things one believes, that one ought to believe these fairly obvious consequences.

I am not suggesting, I hasten to add, that the use of normative terms in expressing constraints on belief cannot be reductively analysed. I belong to the unpopular party that thinks that *all* normative terms can be reductively analysed.[12] The point is that the issue is the same for the uses of the normative terms in the case of belief as it is in the 'hard' cases in general. Belief is, that is, one of the hard cases. This means that non-cognitivists cannot respond to our problem for their view by espousing reductionism about the normativity of belief. For non-cognitivists belong to the party that is convinced that all attempts at reductively analysing the hard cases are a failure; this is a *major* reason why they are non-cognitivists.

Eliminativism about the normativity of belief

Finally, we come to the response that seeks to meet our problem by abandoning the datum from which we started, namely, that there are normative constraints on belief. This view denies that if you believe that P, and believe that if P then Q, you ought to believe that Q, and denies that belief ought to evolve in certain ways under the impact of new information. And so on. But think for a moment of the context in which the issue has come up. We noted at the beginning that there is really one big issue that concerns, equally, normativity, rationality and morality. As we saw there, the general considerations that make all the philosophical trouble can be expressed equally in terms of normativity, rationality or morality. For example, Gibbard's very clear presentation of the argument from disagreement for non-cognitivism in chapter one of *Wise Choices, Apt Feelings* is framed in terms of rationality. He describes some famous puzzle cases in the literature on decision theory where it has proved impossible to obtain a consensus on what it is rational to do. In all these cases, smart people typically agree about all the facts, presumably understand the cases

[12] Frank Jackson, *From Metaphysics to Ethics: A Defence of Conceptual Analysis* (Oxford: Clarendon Press, 1998), chs 5 and 6.

and are not somehow confused (because they are smart and are given plenty of time to debate and think the issues through), and yet find themselves disagreeing about what it is rational to do in them. Although Gibbard does not particularly focus on the Newcomb paradox, he well might have. It is notorious that some smart people are convinced that it is rational to one-box, and others are convinced that it is rational to two-box, and that no amount of careful discussion and noting of possible ambiguities and confusions leads to general convergence. At the end of it all, most of the one-boxers are still one-boxers and most of the two-boxers are still two-boxers. Moreover, neither party is happy to describe the situation as a 'talking past' each other; each party insists that there is a genuine disagreement about what is rational, genuine because the other party is wrong. It seems, runs the argument, that their disagreement cannot then be over the (agreed) facts. Ergo, it is over something other than the facts. Ergo, non-cognitivism is true. But this means that invoking eliminativism to handle our problem would only help in the final analysis if we could embrace eliminativism about rationality. We would, that is, have to abandon the idea that some systems of belief, some people, and some actions are more rational than any others. This is very strong beer indeed.

It is important to be clear about what eliminativism about the normativity of belief means here. There is *a* sense in which non-cognitivists are automatically eliminativists about normativity and rationality in the same way that they are automatically nihilists about morality in *a* sense. They do not think that it is a matter of fact that some actions are more rational or moral than others; they do not think that any actions have the property of being more rational or morally better than other actions. The eliminativism we are talking about here is a denial of the datum that belief is normative from which we started, however, precisely, that datum should be construed. For example, the denial, as the non-cognitivists would construe it, includes denying that we have the attitudes towards someone who believes that P, and that if P then Q, we would express using the words 'ought to believe that Q'. It is this denial that is such strong beer.

Coda

After writing this chapter, I read 'Appendix: Common Questions' in Blackburn's *Ruling Passions*. The following passage appears on pp. 319–20.

Q. 20. If all descriptions of human psychology, including descriptions of what people believe and desire, are implicitly normative, how can normativity be essentially a matter of attitude?

Ans. All such descriptions are plausibly regarded as normative because they seem to imply views about what it 'makes sense' for a person to do, if they are in the states attributed to them. And affirming that something did or did not make sense is a normative judgement.

If we accept this line of reasoning, all I then add is that the verdict that a person's behaviour does or does not make sense itself expresses an attitude. And where is the harm in that?...

One way of encapsulating the germ of this chapter is that I think that this is a much more difficult question for non-cognitivists to answer than Blackburn allows. We noted earlier Blackburn's claim that 'Expressivism denies that when we assert values, we talk about our own states of mind... It says that we *voice* our states of mind, but denies that we thereby describe them' (*Ruling Passions*, p. 50). But this, combined with the normativity of talk of belief that Blackburn and I agree on, would mean that when I say that I believe something, I am not describing my own state of mind and, more generally, that it is impossible for people to describe their own states of mind.[13]

[13] I am indebted to many discussions of non-cognitivism and expressivism over the years with Philip Pettit and Michael Smith. If I remember right, the key point in this chapter first emerged in a discussion with Smith but he should not be held responsible. He also drew my attention to the passage I quote from Blackburn's book.

6

IRREALIST COGNITIVISM[1]

John Skorupski

1. Cognitivism without realism

What role do normative claims play in our thought? I believe that the correct answer must incorporate the following three points.

(1) Normative claims are assertions of normative propositions – judgeable contents which may be true or false. Normative propositions either explicitly concern what there is reason to believe, to do or to feel, or they are propositions from which such propositions about reasons are analytically deducible.[2] So they are about *reason relations*: relations whose relata are facts, persons, times, degrees of strength – and belief-, action-, or feeling-types. A *purely* normative proposition is a normative proposition from which no non-analytic non-normative proposition is analytically deducible.

(2) The distinction between normative and non-normative claims is characterisable in a philosophically uniform way across all three domains – the epistemic, practical and evaluative. Normative claims in all these domains – of what there is reason to believe, to do and to feel – are assertions which can be assessed as true or false; the same epistemology and metaphysics of the normative applies in all three cases.

(3) Nothing in the world makes purely normative propositions, in any of the domains, true or false. There are no worldly facts in virtue of which pure normative propositions are true

[1] Comments on earlier versions from at least the following have made a difference: Lars Binderup, John Broome, Peter Clark, Garrett Cullity, Jonathan Dancy, Andre Gallois, Paul Guyer, Adam Morton, David Papineau, Hilary Putnam, Peter Railton, Tom Ricketts, Stewart Shapiro, Tim Williamson, Crispin Wright. I thank them all.

[2] In a full discussion of normativity it will not be trivial what meaning one should attach to 'analytically deducible'; in particular because questions will arise about the normativity of logic itself. However, assuming only that *some* notion of this kind can be defended I think these further issues can be side-stepped here.

when they are true. There is no special sector of reality which they describe, 'represent', or 'fit'; they have no 'truth-makers'; there is nothing to which their truth 'corresponds'.

A view of normative claims which propounds these three theses may be described as cognitivist but irrealist. This chapter will defend irrealist cognitivism. Obviously the notion of a 'worldly fact' is crucial to this view: specific attention will given to it in section 5. I shall show that one can distinguish the metaphysically robust notion of worldly fact from a broader, purely nominal or ontologically non-committal notion of fact. When people resist the idea that normative claims are factual assertions I believe they have the notion of a worldly fact rather than a merely nominal notion of factuality in mind. The sticking point, for them, is the realist's thesis that normative claims are true or false *according to whether some fact obtains 'in the world'*. They are unlikely to object, in non-philosophical contexts, to talk about the fact that one should not eat meat (if they believe that one should not eat meat); but they do *not* think this to be a fact in just the way that it is a fact that there's a computer on this table. Nor, however, do they think that it's a different kind of fact in the sense that it belongs to a non-natural sector of reality inaccessible to perception or scientific inquiry. Rather, they think that in a sense of 'fact' which is salient and dominant for them it's not a fact at all.

These are instincts I want to defend systematically. Taking talk of 'facts' *simpliciter* to be talk of worldly facts (in other words reading 'fact' in the sense which I have just suggested is salient and dominant), the position I want to defend can be put thus: pure normative judgement is inherent in all cognition, including cognition of facts, but it is not itself judgement about some domain of fact. Facts give us reason to believe, do or feel: that they do so is not itself a further fact in any sector of reality, natural or non-natural. No proposition which solely pictures a state of affairs is normative. All cognition involves a philosophically fundamental dualism of descriptive and normative judgement, or in terms which will be used in what follows, of receptivity and spontaneity.

Such a view seems at first sight attractive and sensible, yet it is surprisingly uncommon. While there may be various reasons for that, I think there is one very fundamental reason. A deep metaphysical obstacle seems to lie in its way: informative cognitive content just seems to *be* factual content; content is factual content.

The cognitive irrealist[3] will have to dispel the seeming force of this idea. For if all cognitive content is factual content in the intended more than merely nominal sense then, it seems, we must be either non-cognitivists or realists about normative claims. The apparently forced choice greatly strengthens the case for one or other of the alternatives, which otherwise seem implausible or even bizarre. However this chapter will not argue directly against either of them. Its aim is to open the path for cognitive irrealism by dissolving the obstacle that stands in its way: the apparently metaphysical thesis that all content is factual content. It does so not by denying that there is a reading of 'all content is factual content' in which that dictum is innocuously true, but by denying that in that sense it expresses a metaphysical thesis. This conclusion is approached as follows. Section 2 considers the connection between judgement and convergence of inquirers; arguing that all genuine judgement incurs a certain convergence commitment. Drawing on that connection, section 3 outlines the contrast between the epistemology of the normative and the epistemology of the factual in the three domains of feeling, action and belief. Section 4 briefly surveys the views of the normative which are open if one accepts (contrary to the argument of this chapter) the *metaphysical* claim that all content is factual, section 5 addresses that claim as such. I conclude by noting that the view developed here has no 'non-naturalistic' implication.

2. The convergence thesis and the universality of reasons[4]

When I judge that *p*, I enter a commitment that inquirers who scrutinised the relevant evidence and argument available to them would agree that *p* unless I could fault their rationality or their evidence. I will refer to this as the 'convergence thesis'.

Judging that *p* means coming or continuing to *believe* ('genuinely', 'fully') that *p*. It contrasts with guessing that *p*, acting on that assumption, delivering a verdict in a situation in which one has to (while nevertheless not being quite sure) and so on. As to *rationality*, I use that term here unusually broadly. I mean it to

[3] 'Cognitive irrealism' is just a stylistic alternative to 'irrealist cognitivism'.

[4] In this section I recapitulate (and in some minor respects correct) points made in John Skorupski, 'Reasons and Reason' in Garrett Cullity and Berys Gaut (eds.), *Morality and Practical Reasoning* (Oxford: Oxford University Press, 1997), section 8.

cover everything that's involved in the capacity, given a state of information, to assess reasons of any kind (whether they be reasons to believe, to act or to feel), to estimate their weight on the basis of that information state, and to respond appropriately.

Rationality in this very broad sense can be contrasted with *receptivity*, the capacity to receive information. Further, one's rationality in this broad sense can be diminished in two broad ways. First it may be diminished by absences or defects in what I will call *spontaneity*, that is, in one's capacity for spontaneously appropriate normative responses. Of course a person can have some capacities of spontaneity without having others, just as he can have some capacities of receptivity without having others. These notions of spontaneity and receptivity will be important in what follows.[5] But secondly, we should note that even where there is no intrinsic lack or defect of spontaneity, one's broad rationality may nevertheless be diminished because of *interference* produced by such things as special pleading, wishful thinking, partiality to particular persuaders, exhaustion, inattention and so on. Spontaneity is thus not the only contributor to a well-functioning rationality. Faulting an inquirer's rationality involves identifying an inadequacy which might be *either* a weakness in the relevant qualities of spontaneity *or* an interference, and which in either case relevantly affects his judging propensity on the subject in question – in such a way as to justify one in discounting his judgement. Faulting the inquirer's evidence, on the other hand, would involve showing that it is the *information input* into his judging process which is misleading or insufficient – whether that is because of the limited nature of the incoming message, or in virtue of some impairment of receptivity – in such a way as to justify one in discounting his judgement, even if one finds no fault in his rationality.

Finally, let me stipulate that one *enters* or *incurs* a commitment that q by judging that p just if judging that p rationally constrains or requires one to accept that reasons for judging that it's *not* the

[5] I have chosen them partly with Kant in mind of course. And, like Kant, I take rationality as such to be a matter of spontaneity, involving no faculty of receptivity. Moreover I would want to argue that Kant should be interpreted as a cognitive irrealist, not a constructivist, about reason (as against his transcendental-idealist constructivism about the empirical world). But the notions of rationality and spontaneity to which I appeal are crucially broader than his, in that I take them to be applicable to affective as well as cognitive and conative responses – and hence to evaluative as well as theoretical and practical judgement.

case that *q* are reasons for withdrawing the judgement that *p*. I have put this negatively. When we incur a convergence commitment by making a judgement we do not have to give grounds for expecting the convergence, over and above our grounds for the judgement itself. The point is rather that *if* we come, in one way or another, to have reason to doubt that our judgement would attract convergence among fault-free inquirers, we thereby come to have reason to withdraw it. When those grounds for doubt become strong enough they force withdrawal.

So now what case can be made for the convergence thesis? It follows from three plausible principles. The first of these is the principle of Rationality:

> (1) It is irrational to judge that: *p* but there is not sufficient reason (warrant) to judge that *p*.

I cannot rationally judge that *p* while also judging that there is insufficient reason to judge that *p*. In judging that *p* I incur the commitment that there is sufficient reason for me to judge that *p*.

This does not show that if I judge that *p* I must hold that *anyone* who holds that there is insufficient reason to judge that *p* is faulty either in rationality or in evidence. To reach this result we need a second principle, the principle of the Universality of Reasons.

> (2) Given a total state of relevant evidence, [E], *x* is warranted in holding that *x* has sufficient reason to judge that *p* <–> (*y*) (*y* is in [E] –> *y* is warranted in holding that *y* has sufficient reason to judge that *p*)

And we also need a third principle, the principle of Evidence:

> (3) If *p* then any evidence that warrants *x* in judging that *x* has insufficient reason to think that *p* is either
>
> (i) also insufficient to decide whether or not it is the case that *p*,
>
> or
>
> (ii) misleading inasmuch as it warrants *x* in denying that *p*.

Since *x*'s judgement that he has such-and-such evidence may be warranted but false, 'evidence' in (2) and (3) refers to the total set of factual propositions, relevant to the question whether *p*, which *x* can warrantedly *take* to specify *x*'s evidence on that ques-

tion. And where the question can be settled without evidence [E] will be null. (That will be the case for all fundamental normative propositions, so the Principle of Evidence is irrelevant to the argument in their case.)

Now suppose I judge that *p* and that another thinker, *y*, does not hold that he, *y*, has sufficient reason to judge that *p*. By (1) I am committed to judging that I have sufficient reason, and thus (in cases where evidence is required) sufficient evidence, to judge that *p*. Adding (2) and (3) I am committed to one of four conclusions. The first possible conclusion is simply that *y* has not considered (sufficiently) the question whether *p*. The second is that *y*'s evidence is faulty for reason (3i) or (3ii). That is, it is either insufficient or misleading – partial or distorted in some way. The third possibility is that *y* is faulty in respect of rationality. That is, he doesn't just fail to consider the question whether *p*; he refuses to accept that *p* although his [E], his total relevant evidence, warrants him in holding that he has sufficient reason to believe that *p*. (If his [E] is the same as mine the conclusion that he has not considered the question or is faulty in broad rationality follows by (2)). The fourth possibility is that it is my own evidence or rationality which is faulty. In that case, by (1), I must withdraw my judgement that *p*. So given (1), (2) and (3) so long as I continue to judge that *p* I am committed to judging that any thinker who refuses (after examining the question) to accept that *p* is faulty in evidence or rationality.

In making this point we must distinguish between justification and warrant. I shall use the term 'justified' in such a way that we may be *justified* in holding a belief even when we are not *warranted* in holding it. In such a case our commitment to thinking that fault-free inquirers would converge on it is mistaken too. That is, the fourth possibility, that it is our own rationality or evidence which is faulty, may apply even in cases where the judgement we make is *justified*. In the first place this occurs when our evidence is misleading even though we have no reason to think it is. But also some aspect of our rationality may be faulty. One's belief may not be *warranted* by one's [E] even though one is justified in thinking that it is, and hence justified in the belief. 'Justification' is distinct from 'warrant' in that the former pertains to epistemic appraisal of the judger while the latter refers to a relation between his [E] and the propositions he could judge. One's belief is *justified* if holding it, in the circumstances in which one does hold it, opens one's epistemic virtue to no criticism. It must

be possible for one to be in this sense *justified* in making a judgement even if one has not fully eliminated the epistemic possibility that some relevant aspect of one's rationality is at fault. Otherwise the standard for justification would be set too high. It would require that one had some way of definitely ruling out that epistemic possibility, and had applied it, before one could be justified in making the judgement. We need no such requirement: even granting (as I argue below) that a fault of rationality is always in principle detectable through sufficiently careful self-examination and discussion, it does not follow that there has to be an effective procedure for detecting it, still less that the procedure must have been applied before we are justified in proceeding to judgement.

Notice also that the argument for the thesis does not turn on any particular view of *truth*. The three principles which provide its premises – Rationality, Universality of Reasons, and Evidence – will hold irrespective of what philosophical conception of truth one propounds. In particular, then, the argument in no way rests on defining truth in terms of convergence. There are many true factual judgements on which convergence of judgement could not occur because sufficient relevant evidence could not be collected; belief in them could not be warranted, even though they are true.

Nor is the convergence thesis based on the idea of truth as representation or fit, together with the point that accurate representing devices, aimed at one and the same state of affairs, are going to have to produce the same representation. This indeed is a watershed issue. Since I shall appeal to the convergence thesis in trying to show how a discourse may be cognitive *without* being factual, it is important to be clear that the case I have made for the convergence thesis does not itself rest on an argument from representation. It does not appeal to the idea that *accurate representation devices yield the same representation*.[6] The crucial consider-

[6] Crispin Wright argues along these lines (*Truth and Objectivity*, Cambridge, Mass. and London: Harvard University Press, 1992; 'Truth in Ethics,' *Ratio*, IX, 1996). He also holds that the concept of truth applicable to ethical judgements does not involve the notion of representation. The concept of ethical truth, he suggests, allows that a person may, without irrationality, judge that an ethical proposition *p* is true while *also* judging that convergence of fault-free thinkers (thinkers who suffer from no 'cognitive shortcoming') cannot be expected to occur on that judgement. I certainly agree with Wright that ethical truth involves no notion of 'fit' to a domain of reality. But I do not accept that there is any notion of truth which allows a person to judge a proposition true without incurring the convergence commitment – for the convergence commitment arises quite generally from the *rationality* of judgement rather than any particular notion of *truth*.

ation underlying it is rather the Universality of Reasons. (The other two premisses seem uncontroversial.) I take this to be a fundamental, underived, presupposition of all cognition and discourse. Reasons are reasons for everyone; the very idea of a reason, which all discourse invokes, is the idea of something which makes a universal claim. The contrast indicated here, between an argument from representation and an argument from the universality of reasons, goes to the heart of one's conception of the normative and through that to much else.

I add one final elucidation. Suppose you and I are in [E] but you seem accept that *p* and I don't. Must one of us be faulty in respect of rationality, as the universality of reasons implies? Well, we may not really be disagreeing about what should be *judged*. It's just that you think it's worth going with the hypothesis for various purposes and I don't. That's more like a practical decision and there may be all sorts of considerations outside [E] that make our different practical decision equally reasonable. For example you think it's worth buying the stock because you think the market will rise tomorrow, whereas I'm agnostic and don't want to buy. That may be because you're backing a hunch, because you're richer and can afford to, etc. We very often have to adopt some conjecture or other to get on with practical life. In contrast, if the disagreement is genuinely about whether [E] warrants the *judgement* that *p* then (I maintain) at least one of us must be faulty in respect of rationality, as the universality of reasons implies. Moreover each of us is committed to thinking that the other is open to criticism. For example you think I'm being epistemically cautious beyond what reason requires; I think you're being less cautious than reason requires. Might we conclude after discussion that there's no answer in this specific case as to what reason requires, because rational requirement is a vague notion, and this is one of those cases affected by its vagueness? Indeed we might. The thesis that reasons are universal is not intended to deny that 'reason' is a vague concept. However, if we do come to that conclusion in this particular case then in one respect you are the one who has to give way. You don't have to hold that it's *false* that [E] warrants the *judgement* that *p* – but you have conceded that you're not warranted in holding it to be true. Equally, I don't have to hold that it's *true* that [E] warrants the *judgement* that *p*, though I have to concede that I'm not warranted in holding it to be false. The difference is that you *shouldn't* believe that *p*, because you're not warranted in believing that [E] warrants the

judgement that *p*, whereas I can maintain my refusal to believe that *p*.

3. The dialogical epistemology of normative propositions versus the ontic epistemology of factual propositions

This discussion of the convergence thesis has not required us to distinguish between factual and normative judgements. The thesis applies to all judgements, simply as judgements. But now we come to what I want to claim is the fundamental epistemological asymmetry between normative and factual judgements. When we are dealing with a factual proposition there is the simple possibility that there may not be enough evidence to pass a verdict on its truth-value, however much inquiring we do. We, the inquirers, are situated in the same world as the state of affairs which obtains if the proposition is true. To judge that that state of affairs obtains we, or some of us, must have evidence: that is, we must receive information from that state of affairs (or from states of affairs from which its existence can be inferred).

In the case of fundamental purely normative propositions, in contrast, no such point applies. There is no comparable basis for the idea that a fundamental purely normative proposition might be evidentially undecidable – that is, that we could never have enough *evidence* to be warranted in accepting it as true or false.[7] To this distinction between normative and factual propositions there corresponds a distinction in their epistemology. It is the distinction between what I will call a discursive or dialogical epistemology and an ontic epistemology. The difference between them is that knowledge of factual propositions requires appropriate receptivity, whereas knowledge of fundamental normative propositions requires only spontaneity.

Let us consider first the epistemology of evaluations, by which in this context I mean judgements about what there is reason to feel. The discursive or dialogical character of the epistemology stands out particularly clearly in their case. We distinguish between having an emotional response and judging it justified. To judge that something is amusing, irritating, frightening or

[7] 'fundamental': a purely normative proposition deduced from decidable normative premisses together with some undecidable factual proposition would of course also be undecidable. Also, if reason relations are irredeemably vague then even a fundamental purely normative proposition may be *judgementally*, as against *evidentially*, undecidable.

moving, contemptible, admirable, despicable, desirable and so on, is to judge that there is *reason* to respond to it in that particular affective way. In each of these cases, the affective response typically carries with it a normative impulse, that is – when relevant resources of concept-possession are in place – a disposition to make the corresponding evaluation. For example, to feel admiration is typically to experience the feeling as reasonable, to experience the object as admirable. Where defeating considerations are absent, one simply *acquiesces* in that affective response, *acknowledges* it – that is, given the relevant resources of concept-possession, one judges that the object in question *is* admirable.[8] Or rather, that is the elementary case. There can also be situations in which the response is experienced as an alien intrusion, even in the absence of defeaters. 'For some reason I can't help finding this terribly irritating, though I can't see why'. In these alienated cases the response and its felt normativity are divorced, or perhaps one is oneself divorced from both the response and its felt normativity, rather as though one was hearing alien voices. But this can't be the typical case, if one's spontaneous capacity for evaluation is well-ordered.

The element of defeasibility underpins a distinction between the object seeming admirable to me and its being admirable – generally, between the object seeming to be *f*-worthy and it's being *f*-worthy. Without that general distinction the force of normativity that comes with the spontaneous affective response (the experienced rightness or reasonableness of the response) would encounter no resistance and generate no objective evaluative thought. What puts in place the distinction? An evaluator is warranted in judging that an object is *f*-worthy if

> (i) the object makes him feel *f* (or, if he is imagining a case, he can see that it would), in the unalienated way, and
> (ii) his [E] gives him no warrant to think that other competent evaluators will not or would not feel *f* in these circumstances, and thus would not be inclined to judge the object *f*-worthy.

The evaluator does not just feel the emotion; feeling it as he feels it, without defeating information about himself or others, he

[8] This is not a voluntarist position. 'Acquiescing in', 'acknowledging', the aptness of a response is not at all the same as *deciding* to treat it as apt. For more on this difference between cognitive realism and currently influential forms of voluntarism, such as that developed by Christine Korsgaard, see my critical notice of her *Sources of Normativity* – 'Rescuing Moral Obligation', in *European Journal of Philosophy*, VI (1998).

judges it to be *reasonable*, and by the Universality of Reasons he is committed to holding that other evaluators of competent rationality would also judge it to be reasonable. (Remember that I am using the term 'rationality' in a very broad sense here – it covers all aspects of the ability to judge the reasonableness of a response.) If the evaluator feels that the object of evaluation is not one of which he is a good judge, or that his own state is wrong, he will disqualify himself. So he might say "Well I was rather bored – but I was feeling tired, or I probably didn't understand what the point of it was or I don't know much about econometrics". Similarly, he may disqualify others – as incompetent judges of the subject, or as being in the wrong state. And the content of their judgements, in cases in which he himself feels confident, will for him be an important test of their competence.

This is discursive or dialogical epistemology. The only materials it calls on are experience of *f*, imagining of circumstances in which one would be spontaneously disposed to *f*, reflection on it and inter-subjective comparison and discussion.

The primitive criterion for judgements about what one has reason to *f* is the feeling *f* itself. This is true even in a very poor – easily defeasible – [E] and even if one is, without knowing it, a very poor judge. So we should distinguish between whether a person is *warranted* in his judgements and whether he is *authoritative* in his judgements. Judgements can be warranted at any level of authority. For example my judgement that something is admirable may be warranted even if I am a very inexperienced judge and my information about the object is poor. But my warrant is then very easily defeasible. The warranted judgement I make might be corrected by more information about the object – that is, by improving my [E]. Or it might be corrected in ways that work on the spontaneity that produced it. Greater familiarity with that kind of object (e.g. accurate spitting, cool style, a wine, a novel, a film . . .) may simply make me feel that it is not as admirable as I initially thought it. Comparison with the judgements of others, together with discussion of our disagreements, may have the same result. This last process relies on my implicit acceptance of the universality of reasons; it may make me judge that the object is not really admirable even though I continue to admire it. ("I still can't help admiring really accurate spitting, though I realise nowadays that it's not one of the great human achievements.") I

detach my response from its default-normativity. That usually, though not always, deadens the response.[9]

By these processes I increase my competence or authority in a domain of evaluation – or I reveal that I am not a good judge. For example I'm not a good judge of what is admirable because I'm over-impressionable, or partial to the performer, or I have peculiar beliefs about what we're in this world to do, etc. In general what one feels after experience, self-examination and discussion with others provides a less easily defeasible – but still in principle defeasible – warrant. Spontaneity may be more or less educated or developed.[10] And judgements about what there is reason to *f* interact with judgements about who is a good judge of *f*-worthiness. From the third-person standpoint we are often more interested in the authoritativeness than in the warrantedness of a person's evaluations.

In summary, two general conditions of judgement, namely, spontaneous normative response and the convergence commitment, suffice to give evaluative discourse a hermeneutic discipline – a discipline grounded in what can intelligibly, understandably, produce a given emotional response. The standard they jointly put in place is an ideal of competent or authoritative emotional response; yet such authority still finally rests on nothing other than the spontaneity and universality of reasons. *It does not require us to provide reasons for thinking that our judgement is tracking some domain of facts.*

The evaluative case is the simplest and clearest case to consider. The case of practical normative propositions, propositions about what there is reason to do, is more complicated. For in that case a variety of new questions arises. One big question is how reasons to feel connect with reasons to act. What are the bridge principles which take one from the domain of the evaluative to the domain of the practical? Another big question is whether there are pure requirements of practical reason. If reasons to feel are grounded in the spontaneity of feeling – the

[9] I have chosen the concept of the admirable as an example for discussion because it is what I have elsewhere called a "purely affective" evaluative concept, determined solely by hermeneutic criteria. This makes it easier to highlight the essential epistemological points; however I am not suggesting that all evaluative concepts are purely affective. Notably, the concept of the blameworthy is not. (See 'Reasons and Reason'.)

[10] 'Spontaneous' contrasts both with (1) 'factitious' or 'spurious' and (2) with 'artificial', i. e. conventionally prescribed. Thus cultivation or education of spontaneous responses should be contrasted on the one hand with indoctrination and on the other with the introduction of conventional rules.

typically experienced normativity of an affective response – reasons for action must be grounded in the spontaneity of the will, that is, the typically experienced normativity of an impulse to choose, resolve, act. Such impulses might always be conditional on an evaluation – in other words they might always arise from a connection with a normative affective response. I judge myself to have reason to do something because I judge myself to have reason to feel something (including desire among feelings). For example I think it would be appropriate to give you some flowers because I have reason to feel grateful. On the other hand, it may also be that there are categorical impulses of the will – dispositions to choose, or felt constraints on one's choice, which arise from no particular prior evaluation. That would be the Kantian view, according to which there are pure requirements of practical reason. Either way, for present purposes the relevant point is that whatever fundamental practical norms[11] there are, the discursive epistemology we are considering will apply to them. Its two elements will be the same as in the case of evaluations: (i) spontaneous impulses, in this case to *will* an action in an actual or imagined circumstance (ii) experience, reflection and discussion with others constrained by the convergence commitment.

So much, then, for the evaluative and the practical case. Now for the most difficult and confusing case – the epistemic, the case of reasons to believe. Let us again follow the path marked out so far. Here too we will expect to find fundamental purely normative propositions, in this case about what there is reason to believe. Let us call true propositions of this kind 'epistemic norms'. Once again the two elements in their epistemology will be spontaneity and universality: spontaneous impulses, typically experienced as normative, to believe something in an actual or imagined circumstance and reflection and discussion with others constrained by the convergence commitment. However the vital feature of epistemic norms is that they lead us to factual beliefs. That means that in the epistemic case there is a fundamental task which has to be discharged by the approach we are following, namely, that of elucidating in detail the way in which epistemic norms interact with factual propositions.

The approach we are following says that fundamental norms

[11] I use 'norm' simply as short for 'true purely normative proposition'.

correspond to no fact. This must apply as much to epistemic as to evaluative and practical norms. But how can it be reconciled with the point that epistemic norms guide us in forming factual beliefs about the world? Let us suppose, for the sake of an example, that enumerative induction is an epistemic norm. (It does not matter too much whether this is right, in that the same points could be made about alternative candidates, such as inference to the best explanation; but enumerative induction is a relatively simple case to consider.) Then the fundamental normative proposition in question is that

(1) Enumerative induction on appropriate premises and in the absence of defeating information yields reason to believe general propositions about the world.

Or in short, induction defeasibly warrants belief. The reason to believe which it provides varies in strength: its strength increases, *ceteris paribus*, with the number of cases and the variety of circumstances in which they are observed; however, and importantly, it is also highly sensitive to collateral information. In some domains about which we have such information our reason to believe a generalisation increases rapidly with confirming instances, in others it increases only slowly. Why should this be?

In answering the question we must distinguish the epistemic norm of induction, *viz.* (1), from a factual proposition, or rather a potential variety of factual propositions, about induction:

(2) Enumerative inductions – in some specified or in all domains of inquiry – yield general propositions which are not refutable by counter-examples (or yield them increasingly as time goes on, or yield them with such and such long-run frequency, etc.)

Whereas the normative proposition, (1), says in short that induction warrants belief, the factual proposition, (2) says in short that the world is induction-friendly. But this is itself a general proposition, or a class of general propositions, about the world. Thus – granting (1) – we may acquire a warrant to believe (2) by a second-order enumerative induction, in some specified domain for which we have sufficient confirming instances of so-far successful inductions. In that case induction is self-confirming for the domain. But it's also possible – granting (1) again – that a second-order induction will give us reason to *disbelieve* (2) in some domains. This result will reasonably make us

cautious about attaching significance, in such domains, to degrees of observed correlation which would in other domains warrant generalisation.

In this way and others (for example involving hypothetical inferences) as we build up a corpus of general beliefs we come to place greater rational credence in the reliability of induction in some domains than in others. That is why, as our information improves, the warrant provided by an induction can vary so much across domains. But notice that this process presupposes the epistemic norm of induction, and that no possible result of the process shows that *norm* – as against factual propositions of the form summarised in (2) – to be false. At the limit, induction might turn out to be internally self-undermining overall. By induction we come to realise that induction is not to be relied upon. In other words induction puts us into a state of information in which every induction is defeated. But the norm of induction, (1), remains true, even though, in the state of information to which it has in fact led us, no observed correlation would warrant generalisation. The truth of the norm corresponds to no fact and is refuted by no factual outcome.

This is only an example. I am not suggesting that inductive norms are the only epistemic norms there are. On the contrary, I believe that all concepts, not just the concept of generality, are constituted by epistemic norms governing the introduction and elimination of those concepts in thought, including thought about what there is reason to feel or do. If an epistemic norm in general is a true proposition of the form *it's being the case that p would give x reason (of degree d at time t) to believe that q* the concept partially constituted by the norm can occur in the content *p* or the content *q*.

4. The Alternatives

We've arrived at this account by taking seriously the guiding principles I started with: that a uniform account of normative claims should be given in all domains, action, feeling and belief, and that that account should take these claims to be propositional but not factual. Of course it would also be desirable to argue to it by eliminating the options which are open on the alternative assumption that all propositional content is factual.

I take it that we have reasons to believe some of the things we believe – and that we can know ourselves to have such reasons. That is, we can audit our beliefs and *tell* whether or not they are supported by reasons: not in the mode of hyperbolic doubt but in the mode of rational stock-taking which is a standard part of everyday reflective discourse. An account which forces one to deny that that is so is an account which can be set aside as inadequate.

The responses available to someone who holds that all content is factual content should be considered in this light. In the epistemic case they are as follows.

A. Non-cognitivism. So-called 'epistemic norms' are not propositions at all but express non-doxastic states of mind, such as shared intentions, or conventions, perhaps enshrined in rules of language.
B. Non-reductive realism: epistemic norms are true propositions whose truth corresponds to some non-natural domain of normative fact.
C. Reductive realism: epistemic norms are true propositions whose truth corresponds to some non-normative fact. The currently influential versions of this are naturalistic (rather than theological).
D. Eliminative realism. There are no epistemic norms. There is only the hypothetical imperative, which is a practical not an epistemic norm: if your aim is to believe the true, then you should conform your beliefs to truth-preserving principles, principles which correspond to the facts. Thus you should reason inductively (if you should) for the same reason that you should reason in accordance with the general truths of physics or sociology. Epistemology is a branch of natural science.

The combination of realism and naturalism, either in reductive or eliminative form, has a leading position in current epistemology. (Purely 'externalist' accounts of the notion of justification are a form of C or D.) (A), logical positivism, had a period of great dominance but is no longer influential, and (B), non-reductive realism about epistemic norms, is not greatly influential either.[12] This situation in epistemology contrasts interestingly with meta-ethics, where versions of all the positions I have listed,

[12] Non-naturalistic realism about induction is presented by Russell in his *Problems of Philosophy*. See also Laurence BonJour, 'A reconsideration of the problem of induction', *Philosophical Topics*, XIV (1986).

including non-cognitivism and non-reductive realism, are strongly represented in the contemporary debate.[13]

Note the important epistemological difference between the logical positivist version of naturalism and its currently dominant realist forms. Logical positivists took what I'm claiming are epistemic norms, but what they regarded as rules of language, to be preconditions of discourse about experience which are themselves independent of experience. They had to take these conditions as non-propositional *rules*, because they also held that all genuine *propositions* are factual – and, as good empiricists – that all factual propositions are empirical. In contrast, the currently dominant realist forms of naturalism deny the idea that there are preconditions of discourse about the world (an idea which belongs to the Kantian tradition, though it is not exclusive to it).[14]

The position I am proposing shares with logical positivism the view that factual discourse has non-factual preconditions. In our example of induction it prioritises (1) over (2) in the epistemological order, whereas realist forms of naturalism prioritise (2) over (1). Against logical positivism, however, I would argue that Wittgenstein's discussion of rule-following undermines the attempt to eliminate normative propositions and treat the distinction between factual propositions and expressions of 'rules' as exhaustive. In following a rule, I do not apply further rules; rather I grasp normative truths about how I should apply the rule to the facts, truths which do not themselves record further facts, either platonic facts or counterfactuals about how the community would behave.[15] As to naturalistic realism (in all its forms), it

[13] Eliminative naturalistic realism, the analogue of (D) in epistemology, seems to be the least popular, but was put forward explicitly by J. L. Mackie and is (I suspect) more or less influential among decision-theorists and others in social science. Note that in epistemology this view appeals to a hypothetical imperative, that is, to a principle of *practical* reason. Obviously that raises the question whether one can simultaneously take an eliminativist view of the hypothetical imperative.

[14] The affinities in this regard between logical positivism and Kant are brought out by Alberto Coffa, *The Semantic Tradition from Kant to Carnap: The Road to the Vienna Station* (Cambridge: Cambridge University Press, 1991). Philip Kitcher calls an approach of this kind 'transcendentalism'. See his 'Mill, mathematics, and the naturalist tradition,' in John Skorupski, ed., *Cambridge Companions to Philosophy: John Stuart Mill* (Cambridge: Cambridge University Press, 1998); for further discussion see my 'Logical Grammar, Transcendentalism and Normativity', *Philosophical Topics*, XXV (1997).

[15] See Skorupski, 'Why did Language matter to Analytical Philosophy?', *Ratio* IX (1996) and Skorupski, 'Logical Grammar, Transcendentalism and Normativity'. (This argument assumes that normative claims about how to apply a rule have truth value and

seems to me to remain as vulnerable as ever to the classical Kantian critique, which starts from the requirement mentioned above: that any account of reasons must be consistent with holding that we *have* reasons for some of our beliefs and can *tell* that we have.

Such large claims obviously call for extended vindication. However, as I said at the beginning, my concern here is not to criticise alternative positions; I have mentioned them just to indicate what line I would hope to take if I was doing that. And in mentioning them I have so far left out one of the alternatives: non-naturalistic realism about the normative. Is *this* the right option to adopt if the lines of criticism I have just mentioned are sound?

Those same lines of criticism, I would want to argue, can be extended to it. If *natural* facts cannot literally be identical with reasons no special non-natural magic can make *non*-natural facts literally identical with reasons. (Supposing there were such facts – they would just be further facts.) However the next section restricts itself to a necessary preliminary: showing how the account of the normative which I do want to defend, across all domains, *differs* from this realism. If my account is genuinely distinct and defensible that itself places the onus of proof on the side of the non-reductive realist. For if it is genuinely distinct and defensible then why be a realist about the normative, of any kind? Nothing in our normative practice calls for such realism (if I am right). A positive case would have to be made.

5. Nominal versus worldly facts; non-receptive versus receptive awareness

We have seen that the basic case against cognitivist irrealism comes from the thesis that content is always factual content. So what can be said in favour of the thesis?

In English the expression 'the fact that ...' is a nominalising operator on assertoric sentences. Prefixing it to such a sentence produces a noun-phrase which has a semantic value just when the nominalised sentence expresses a true proposition. That applies to *any* true sentence, normative or otherwise. We can refer to the fact that eating meat is wrong or that a piece of reasoning is unsound just as we can refer to the fact that there is a computer

cannot be interpreted non-cognitively. In this case as in others non-cognitivism faces familiar and, as it seems to me, insurmountable problems.)

on the table or that there are mountains in Switzerland. Applying the general principle that if a singular term '*a*' has a semantic value the semantic value is *a*, we conclude that if the expression 'the fact that eating meat is wrong' has a semantic value it's the fact that eating meat is wrong. And, since we've granted that such nominalisations ('the fact that ...') have a semantic value just when the nominalised sentence expresses a true proposition, if we accept that 'eating meat is wrong' expresses a truth then we must also accept that there are normative facts. In general this argument seems to show that all sentences which express truths correspond to facts.

Furthermore, the grammar of ordinary English suggests that facts, that is, the semantic values of these noun-phrases, are not identifiable with true propositions. For example the expression 'the fact' goes into any of these contexts:

> (un)aware of ... that p, (un)conscious of ... that p, remembers ... that p, has forgotten ... that p, recognises ... that p, acknowledges ... that p, notices ... that p, notes ... that p, registers ... that p,

But it doesn't go into 'believes ... that *p*', whereas the expression 'the proposition' or 'the true proposition' does. (This isn't just because 'believes' is not a factive verb; 'the fact' doesn't go into 'knows ... that *p*', or 'realises ... that *p*' either, even though you can't know or realise that *p* unless it is indeed the case – a fact – that *p*.[16])

This suggests that facts, including normative facts, are possible objects of awareness, but not possible objects of belief, whereas propositions are possible objects of belief but not possible objects of awareness. One can be aware of (the occurrence, obtaining of) an event, a state of affairs – or a fact, for a fact just is the occurrence or obtaining of an event or state of affairs. And all of this goes with a correspondence conception of truth, according to which a true proposition is one which describes an event which does occur, or a state of affairs which does obtain. Facts are what true propositions correspond to. Moreover we know that a proposition is true just if we are aware of the fact which it depicts, either inferentially or non-inferentially.

Do these points show that cognitivism about a domain of dis-

[16] 'The proposition' doesn't fit into 'realises ... that *p*' but does fit into 'knows ... that p'. But in the latter case it signifies knowledge by acquaintance. Likewise in the contexts '(un)aware of ... that *p*' etc.

course forces realism about it? I don't think so. We should distinguish between a purely nominal and an ontic or worldly conception of facts. In a purely nominal sense of 'fact' and 'correspond' all true propositions correspond to facts. But in the ontic sense of those words no fundamental normative proposition corresponds to facts.

Of course it must be shown that this distinction between the nominal and the ontic is not just a verbal manoeuvre. Let us return, then, to the contrast between spontaneity and receptivity. Consider the following two claims:

The Content Principle: it is epistemically possible that a thinker should think that *p* without it being the case that *p*, and that it should be the case that *p* even though the thinker does not think that *p*.

The Ontic Principle: only if there is good enough reason to believe that there exists an appropriate link between the fact that *p* – i.e. the occurrence of the event or obtaining of the state of affairs depicted by the proposition that *p* – and the thought that *p* is the thought that *p* warranted. To be warranted in thinking that *p* one must have good enough reason to think that one's disposition to think that *p* is well-grounded: linked in that appropriate way to the fact that *p*.[17]

The Content Principle is putatively a commitment of cognitivism as such, in that any judgeable content that *p* must satisfy it. 'Putatively', inasmuch as it gives rise to familiar problems about such apperceptive judgements as 'I exist', 'I am thinking', or perhaps even more generally, about the cognitive status of first-persons claims about one's own present experience. However whether the Content Principle requires restriction in the light of such cases does not affect the present point, which is that we have shown in sections 2 and 3 that fundamental normative propositions can satisfy it even though they depict no worldly state of affairs.

In contrast the Ontic Principle cannot on the view taken here be a general truth about all judgeable contents. It will apply only to contents which depict worldly states of affairs, and whose truth corresponds to a worldly fact. The very idea of a worldly fact is the

[17] This can be a default reason. Suppose I seem to see something red there and have no defeating collateral information. Then I am warranted in thinking that there's something red there; and I also have good enough reason to think my disposition to judge that there's something red there is grounded on there being something red there – it would be reasonable for me think that, whether or not I think it.

idea of a fact which can only be known to obtain by *receptive* awareness. To have that knowledge the thinker must be equipped with an appropriate receiver: a faculty which receives information from something else (outside or inside him). That faculty will consist in a capacity for immediate receptive awareness of some class of facts. And for any worldly fact that *p*, knowledge that *p* will require some transmission from the fact that *p* (or from facts in virtue of which the fact obtains, or from which it is inferable inductively or via testimony) to the knower, via a receptive capacity of awareness.

Irrealism about the normative says that normative knowledge of fundamental normative propositions *rests on no receptive awareness.* The only capacity it requires is the non-receptive cognitive capacity of rationality, a capacity which involves spontaneity and regulation by the universality of reasons, not receptivity. And here 'receptivity' covers *any* capacity of receptive awareness – be it the uncontroversial ones based on the human senses, the more controversial ones involved in self-awareness of one's own state of mind, or some other ones such as Platonic intuition or Martian telepathy or whatever. What is at stake between the irrealist and the non-reductive realist is whether normative 'facts' are nominal or worldly facts – and that turns on whether, contrary to the account given of our normative practice in sections 2 and 3, normative knowledge does after all involve some *sui generis* form of receptive awareness, awareness of facts in a normative sector of reality.

Ordinary English, as we've noticed, does allow us to refer to the fact that *p* wherever there is a true proposition that *p*; and it allows us to say of any person who knows a true fundamental normative proposition that that person is aware of the corresponding fact.[18] But for the irrealist, awareness in this normative case is non-receptive, requiring only the capacity of rationality (in the broad sense in which the word is being used here). The appropriate epistemology is dialogical, not ontic. Purely dialogical knowledge is knowledge that *p* which need not be based on (receptive) awareness of any (worldly) facts.[19]

One might envisage a regimentation of ordinary English

[18] Though it's also true that in the case of 'normative facts', say the fact that torture is wrong, native speakers in reflective moments typically show queasiness, and tend to responses like 'Well, in a sense it's a fact', etc. I don't myself think this is just the hangover of positivism.

[19] This may be too strong. Obviously, to discuss a normative proposition with others

which allowed no talk of purely nominal, non-worldly facts and non-receptive awareness. It would not allow the factuality operator to nominalise sentences which did not correspond to worldly facts but would nominalise them in some differently marked way. 'Aware' in this language would always refer to receptive awareness and 'fact' would always refer to worldly facts. That would be philosophically more elegant, if tiresome or even practically impossible, and it would lose none of ordinary English's expressive power.[20]

Yet even such regimented English, if it were possible, would still have expressions whose semantic values are reason-relations, and would allow for higher-order quantification over reason-relations. (E. g.: 'This fact has a relation to the proposition under consideration which that fact does not – it's a reason to believe it.') Does *this* mean that we must 'countenance reason-relations in our ontology'?

If we consider carefully, we see that the sole task of semantics is to specify what topic of discourse a speaker who uses a particular term in discourse has in mind. A semantics for a given language will specify the semantic values of its terms. That means: it will elucidate what topics of discourse are introduced by the use of those terms. When the topic of discourse assigned to a primitive term is presented as an existent, something which can satisfy predicates which only existents can satisfy, the term introducing it has a semantic value just if such an entity does exist. In understanding the term I must grasp what topic it introduces into discourse and I will have the topic of discourse in mind only if I have an *existent* in mind. But the topic of discourse assigned to a term may be a reason relation: not an existent but in Brentano's term an intentionally inexistent object of cognition. To understand such a term is still to grasp its semantic value, that is, to grasp that the

I do need to be receptively aware of some worldly facts (about what they're saying, and so on). I have not considered here what is required for normative knowledge, as against warranted normative belief; it may be that normative knowledge requires discussion with others – in which case it *would* – in that way – require receptive awareness of some worldly facts about others.

[20] A much-discussed problem for such linguistic revision is posed by 'thick' terms of appraisal, such as 'sentimental' or 'insolent', whose factual and normative elements are practically inextricable and perhaps extricable in principle only by introducing new vocabulary. We have very good reasons (and they are not just reasons of economy) to have such terms in our language, and thus not to indulge in wholesale regimentation – but we also have effective ways of extricating the elements in specific contexts in which we need to do so.

topic of discourse it introduces is a reason relation. Once again, one has that topic of discourse in mind only if one has the appropriate reason relation in mind. But that is not having any existent in mind, which is why it requires no receptive awareness of any existent. Nothing in the dialogical epistemology of the normative forces realism about reason relations on us – and what else could? As far as I can see, the only argument the realist can now reach for is the hoary thought that one cannot think of what does not exist.

A similar point holds for the thesis that understanding a sentence is knowing its truth-condition. It's true, precisely in the sense that in understanding a sentence I must know what *condition* must hold for the sentence to be true. But knowing what condition must hold doesn't have to mean: knowing what worldly fact must obtain for the sentence to be true – though it does mean that where the sentence is a factual sentence. The rationale of truth-conditional semantics is that declarative sentences have content, not that they depict worldly facts. To move from the one point to the other is to beg the question of whether content is always factual content, in a metaphysically inflationary way.

And what about the word 'true'? Anyone who thinks 'true' is by its meaning linked with the notion of correspondence, representation or description may prefer to think of normative propositions as, say, 'valid' ('valid', like 'true', will be a non-epistemic notion; but not a 'correspondence' notion) and perhaps reserve the term 'correct' to mean 'true or valid'. That may be a healthy preference. If we want to speak in this way we will talk about correctness conditions instead of truth conditions and instead of saying that one is justified in asserting that *p* is true whenever one is justified in asserting that *p* we will say that one is justified in asserting that *p* is correct whenever one is justified in asserting that *p*.[21]

6. Conclusion

Irrealist cognitivism says that thinking can be *about something*, namely about rational constraints on thinking, without being about something which occupies some sector of 'reality', the

[21] Of course the point about inextricability mentioned in note 22 will also pose a problem for this linguistic proposal. In this chapter I have used 'true' broadly, in conformity with the principle that one is justified in asserting that *p* is true whenever one is justified in asserting that *p*.

'world', etc. But what does it mean to say this? Nothing more, I claim, than is captured in the distinction between receptivity and spontaneity. William James held that "Anything is real of which we find ourselves obliged to take account in any way".[22] The less pragmatist and more Kantian attitude advocated in this chapter starts instead from the distinction between constraints spontaneously acknowledged and facts receptively discovered.

Its irrealism shows in what philosophical arguments about the normative it puts out of work. For example, the irrealist will insist that though the convergence thesis applies to fundamental normative judgements, it is philosophically misguided to 'explain' that by appeal to the argument from representation (see above, p. 443). For such 'explanation' postulates a normative reality and a receptivity to that reality whose output accurate representers, given the same input, will share – a picture which is nothing but metaphysical illusion. Likewise, the irrealist will insist that the idea that there might be such a thing as a theoretical reduction of reason relations to naturalistic relations, or that such a reduction is desirable to achieve 'ontological economy', is wholly misguided. There is no such reduction – but that is not because reason relations are an irreducibly non-naturalistic kind of worldly relations, to be recognised in a full list of what the world contains. Naturalism is the thesis that the natural facts are all the facts. Only when combined with metaphysical realism about content does it force one to question the cognitive status of normative claims.

[22] James, *Some Problems of Philosophy*. (Greenwood Press, 1968), p. 101.

INDEX

This volume brings together two papers given at the 1998 RATIO conference on normativity, by Peter Railton and John Skorupski, and further papers from John Broome, Christopher Hookway, Frank Jackson, and Joseph Raz. This prestigious group of philosophers comes together to consider:

- What is the nature of normative force?
- Is normativity itself subject to sceptical attack?
- What are the relations between moral and epistemic normativity?
- How best should we understand the nature of a normative requirement?
- Can there be an expressivist conception of epistemic normativity?
- Can we suppose that there are normative beliefs or judgements without admitting the existence of normative facts?

The papers bring out many interesting interconnections (and some disagreements) in the field of normativity, which is the subject of intense current debate.

Cover design by Raven Design

Printed in Great Britain

visit our website at: http://www.blackwellpublishers.co.uk