Normativity and Judgement

Author(s): David Papineau and Julia Tanney

Source: *Proceedings of the Aristotelian Society, Supplementary Volumes*, 1999, Vol. 73 (1999), pp. 17-43+45-61

Published by: Oxford University Press on behalf of The Aristotelian Society

Stable URL: https://www.jstor.org/stable/4107061

NORMATIVITY AND JUDGEMENT

David Papineau and Julia Tanney

I—David Papineau

ABSTRACT This paper disputes the common assumption that the normativity of conceptual judgement poses a problem for naturalism. My overall strategy is to argue that norms of judgement derive from moral or personal values, particularly when such values are attached to the end of truth. While there are philosophical problems associated with both moral and personal values, they are not special to the realm of judgement, nor peculiar to naturalist philosophies. This approach to the normativity of judgement is made possible by naturalist views of truth, that is, views which do not presuppose normativity in explaining truth.

I

Introduction. It is widely assumed that the normativity of conceptual judgement poses problems for naturalism. Thus John McDowell urges that 'The structure of the space of reasons stubbornly resists being appropriated within a naturalism that conceives nature as the realm of law' (1994, p 73). Similar sentiments have been expressed by many other writers, for example Robert Brandom (1994, p xiii) and Paul Boghossian (1989, p 548).

In this paper I want to dispute this widespread view. I am not going to deny that there are norms of judgement. My thesis will only be that such norms raise no special problems for naturalism. The normativity of judgement can as easily be accommodated by naturalists as by anybody else.

Let me begin by fixing an agreed subject matter. By norms of judgement I mean true statements about the way judgements *ought* to be made.[1] I take it to be uncontentious that there are many truths like this.

However, while I accept that there are many such truths, I do not accept that there is a species of normativity peculiar to judgement. My general strategy in this paper will be to argue that all norms of judgement are derived from moral or personal

1. And by judgements I mean beliefs and their linguistic expression (though until Section VII I shall use 'judgement' to refer specifically to beliefs). In this paper I shall stand pat on the issue of exactly which mental attitudes qualify as beliefs, and in particular on whether moral attitudes do.

values, and do not involve any sui generis species of conceptual normativity.

A central role here will be played by the end of truth. I shall contend that the most significant[2] norms of judgement can be viewed as prescriptions to the effect that, in order to achieve the truth, you ought to judge in such-and-such ways. In my view, there is nothing constitutively normative about the end of truth itself. So I take the force of these prescriptions to derive from independent moral or personal reasons for attaching value to truth.

Thus truth may in some (or possibly all) contexts have a *moral* value. Alternatively, people may sometimes simply *want* the truth, for whatever reason. In either case, these evaluations will imply truths of the form 'you ought to judge in such-and-such ways'. In the former case, these will be derived moral oughts, about how to achieve the morally valuable end of truth; in the latter case, they will be derived hypothetical imperatives, about how to achieve the personally desired end of truth.

Of course, both these 'oughts' raise further questions for the naturalist. Most obviously, in the moral case the naturalist needs to explain how moral value itself is possible. There is also a less obvious difficulty about the second case, namely, to explain the status of the 'you *ought* to X' that seems to follow from 'you want Y' and 'action X is necessary for Y'.[3]

2. At the end of Section V I shall say something about norms which relate to truth-independent aspects of judgement.

3. When I first started working on this paper, I assumed that these hypothetical 'oughts' at least could be taken for granted by naturalism. More specifically, I assumed that there is a sense in which (i) 'A ought to X' is an analytic consequence of (ii) 'A wants Y' and (iii) 'X is necessary to Y'. Given that there is nothing naturalistically problematic about facts of kinds (ii) and (iii), I therefore took it that there can be nothing naturalistically problematic about claims of kind (i): maybe such claims don't feature a *moral* 'ought' (maybe they involve a kind of pun on 'ought'), but even so they are analytic consequences of (ii) and (iii). More recently, however, I have been persuaded by a paper by John Broome (forthcoming) that this doesn't work. Broome shows that there is no special non-moral 'ought' that can be analytically derived from (ii) and (iii), because no 'ought' of any kind can be so derived. The only prescription in play is 'You ought, if you want Y (and X is a means), to perform X'. The non-conditional prescription 'You ought to X' cannot be detached. Broome draws an analogy. 'Necessarily, if John is married, he has a wife.' We ordinarily then say, knowing that John is married, that 'John must have a wife'. But this obviously doesn't mean that 'John *necessarily* has a wife'. It is simply an elliptical statement of the necessary conditional. No non-conditional modal claim can be detached. Similarly, Broome shows, 'You ought to X' is often elliptical for 'You ought, if you want Y (and X is a means), to perform X', and not itself detachable. This thus leaves us with the task of explaining the unreduced 'ought' which governs this prescriptive conditional.

However, neither of these difficulties is peculiar to the context of conceptual thought. The issue of moral value arises generally, not just in connection with norms of judgement, as does the more subtle problem of understanding hypothetical 'oughts', as in 'you ought to tune your car well, if you want it to go fast'. So, if there are problems facing the naturalist here, they are not problems occasioned by the special nature of conceptual judgement.

Moreover, both these problems arise for philosophers of all kinds, and not only for naturalists. The nature of moral values presents a puzzle to all philosophers, as does the problem of understanding the 'oughts' in hypothetical imperatives. So naturalists will have no more problems than any other philosophers, if they can reduce the normativity of judgement to derived moral or hypothetical prescriptions.

For both these reasons, I will be content if I can persuade you that norms of judgement are derived prescriptions orientated to moral or personal values. I accept that there are philosophical problems attached to both these kinds of derived prescriptions. But these problems arise across the board, and for all philosophers.

The plan of this paper is as follows. In the next section I shall explain how my approach to the normativity of judgement presupposes a naturalist view of content. In Sections III-V I shall argue that my approach avoids three problems facing alternative views. Section VI elaborates some of these points within the context of specific theories of content. Sections VII then asks whether the linguistic embodiment of judgements raises further problems for the naturalist.

Questions relating to linguistic norms will be postponed to this final section. Until then I shall assume that judgements are mental items, and moreover that their normativity is independent of any norms of language. This is a matter of expository strategy, not of any metaphysical commitment to a realm of pre-linguistic thought. As the final section will make clear, I am happy to allow that some judgements are in part constituted by norms of language. However, it will be helpful to proceed in stages, first showing how naturalism might account for non-linguistic norms of judgement, and then adding in the complexities that arise with language.

II

Norms and the Theory of Content. The approach I have just outlined assumes there is nothing essentially evaluative about truth itself. Truth is in the first instance a descriptive property, like car-speed or celibacy. Like car-speed or celibacy, it may be personally desired or morally valuable, in given contexts, but that is additional to its nature.

Here I am taking a implicit stand on the analysis of truth-conditional content. In particular, I am taking a stand against those approaches to content which place normativity *inside* the analysis of content, in the sense that they presuppose sui generis norms governing judgement in explaining truth-conditional content, and hence truth.[4] Let us call theories of this kind 'non-naturalist' theories of content. Neo-verificationist and Dummettian approaches to content are of this kind, since they take content to depend on the conditions in which you are *entitled to assert* a claim. So also are Davidsonian theories of content, which take content to depend inter alia on facts about when it is *reasonable to form* a belief.

From the point of view of such non-naturalist theories of content, the approach to normativity outlined in the last section will seem deficient. For that approach takes truth-conditional content and truth as given, prior to any issues of normativity, and then analyses norms of judgement as arising only when moral or personal value is attached to truth. Non-naturalist theories will regard this approach as necessarily ignoring the prior norms which play a role in constituting truth. Instead of explaining truth in terms of normativity, non-naturalists will complain, I am trying to explain normativity in terms of truth.

Still, this charge does not worry me, given the availability of a number of accounts of truth-conditional content which do not

4. Let us agree, against the deflationists (cf. footnote 6 below), that a substantial account of truth-conditional content is presupposed by the notion of truth. We can also agree that, once we have such a substantial account of truth-conditional content, then truth itself falls out without further ado. Thus those who are prepared to reify truth conditions can simply say that a judgement is true just in case its truth condition obtains. And those who would rather not reify truth conditions, like myself, can note instead that a substantial account of truth-conditional content will imply, together with further facts, claims of the form 'this judgement is true iff p', for any given judgement, and will therewith specify what is required for any given judgement to be true. (Cf. Papineau, 1993, pp. 83–6.)

assume normativity in explaining truth-conditional content. I shall call theories of this kind 'naturalist' theories of content, and have in mind here such theories as indicator semantics (Stampe, 1977; Dretske, 1981), success semantics (Whyte, 1990, 1991), teleosemantics (Millikan, 1984, 1993; Papineau 1984, 1993),[5] and Fodor's 'asymmetric dependence' theory of content (Fodor, 1987, 1990). All of these theories offer to explain truth-conditional content without any commitment to prior norms governing judgement.[6]

Given that views about normativity interact with general questions about content in this way, an obvious dialectical problem faces this paper. It seems that I won't be able to say anything useful about normativity without first establishing the right overall theory of content. Yet clearly there is no question of adjudicating fully on the right theory of content in this relatively short paper. Because of this, my primary thesis will perforce be

5. Some readers may be wondering why I have classified teleosemantics as a theory that analyses content independently of norms. For many teleosemanticists take the view that their theory does indeed trade in norms, in particular 'biological norms', and that for this reason teleosemantics offers an immediate solution to the apparent conflict between naturalism and normativity (cf. Millikan, 1990; McGinn, 1989). However, I find this implausible. Although I myself uphold the teleosemantic theory of content, and so regard it as the right source for the sort of naturalist content I need for my own approach to normativity, I don't think that the teleosemantics itself says anything about norms, nor therefore anything about how to reconcile normativity with naturalism. Whatever norms are, I take it that they must involve some kind of *prescription*, some kind of implication about what *ought* to be done. This simply isn't true of the biological facts on which the teleosemantic theory rests. There is nothing wrong, I suppose, in talk of 'biological norms', if by this you mean simply that we have been designed by natural selection to operate in a certain way. But these biological 'norms' aren't norms in any prescriptive sense. It is a vulgar, and indeed dangerous, error to infer, from the premise that X has been biologically designed to Y, that in some sense X *ought* to Y. My knuckles have arguably been biologically designed to hit people with, but it doesn't in any sense follow that I ought so to use them. Again, a number of human male traits have undoubtedly been designed to foster sexually predatoriness, but it doesn't follow that men ought to be sexually predatory. Similarly with teleosemantics and judgement. As a teleosemanticist I hold that our beliefs have been biologically designed to track their truth conditions. But I don't think that this does anything to show they *ought* to do this.

6. What about 'deflationary theories', by which I mean theories which combine the thought that truth falls out of content (cf. footnote 4 above), with the suggestion that the device of disquotation (plus the idea of meaning as use) removes any need for a substantial account of truth-conditional content? (Cf. Horwich, 1990.) I doubt that such theories will work (Papineau, 1993, 84–5), but as far as I can see there is no reason why deflationists should not agree with me on all the issues in this paper. After all, they don't assume anything in explaining content, so a fortiori don't assume prior norms.

conditional: *if* you have a naturalist of content, *then* you can explain norms of judgement as derived prescriptions orientated to the end of truth. Still, having said this, I do think that the line developed in this paper gives us some purchase at least on the larger issue of which approach to the theory of content is correct.

Most obviously, my line directly blocks the argument that naturalist theories of content are inadequate because they cannot account for the normativity of judgement. From the perspective of this paper, this argument begs the question. I am happy to agree that *non*-naturalist theories of content imply the existence of prior judgemental norms that resist appropriation 'within a naturalism that conceives of nature as the realm of law', in that non-naturalist theories postulate a peculiar species of content-constituting norms which appear mysterious from a naturalist perspective. But naturalist theories of content don't accept the existence of such special content-constituting norms, since naturalist theories explain content without any reference to such prior norms. They place the norms of judgement *outside* the theory of content, and hold that content and truth are constituted independently of facts of the form 'you ought to judge in such-and-such ways'. Given this, they can then account for such norms of judgement straightforwardly, holding that all the 'oughts' involved are derived oughts, arising when moral or personal value is attached to the independently constituted aim of truth.[7]

So the line defended in this paper directly blocks an argument *against* naturalist theories of content, namely, that they cannot account for the normativity of judgement. However, I also feel that it provides some positive support *for* such theories. This is because the naturalist approach dissolves a number of awkward problems about normativity facing non-naturalist theories of content. This isn't of course a conclusive argument. There may be further overriding considerations in favour of non-naturalist theories of content, in which case we will simply need to face up

7. To avoid terminological confusion, let me make it clear that this is meant as an explanation of how naturalists *ought* to deal with normativity, not of what they actually say about the issue. I don't want to count a writer like Millikan as a non-naturalist, just because she adopts what I take to be the mistaken view that judgemental norms are present *inside* her theory of content (cf. footnote 5 above). In my terminology, then, a theory of content is naturalist if it doesn't in fact commit us to prescriptive norms, whatever its proponents may say about this.

to the problems of normativity they generate. But it surely counts against non-naturalist theories of content that they generate difficult problems that do not arise for their naturalistic rivals.

In the next three sections I shall identify three such problems. The first relates to the *universality* of the requirement to seek the truth. The second relates to the *status* of norms of judgement. The third relates to animals, children and other *unrefined beings*.

In this section I have explained how my naturalist approach to normativity stands opposed to non-naturalist theories of content, theories which explain content and truth in terms of norms of judgement, and so cannot in turn explain those norms in terms of truth. However, my approach to normativity will also be opposed by some philosophers who do not embrace such non-naturalist theories of content. I have in mind here those epistemologists who defend 'internalism', in the sense of upholding normative requirements on knowledge or justification that cannot be explained in terms of the pursuit of the external aim of truth. Such philosophers need not suppose that these 'internal' norms play any part in constituting content and truth (indeed they could embrace one of the naturalist theories of content listed above). But they would still deny my claim that the normativity of judgement can be fully explicated in terms of the pursuit of truth.

My main target in this paper is the resistance to naturalism arising from non-natural theories of content. In the space available I cannot also seriously address the motivations for epistemological internalism, and I shall not refer explicitly to such internalist views in what follows. Even so, some of the problems I shall pose seem to me to raise difficulties for epistemological internalists as well as for non-naturalists about content. In particular, the next two sections seem to me to raise such difficulties, even though the third problem, of unrefined beings, does not.

III

Is There a Universal Requirement to Seek the Truth? I have contrasted naturalist theories, which place normativity *outside* the theory of content, with non-naturalist theories, which place it *inside*. In this section I want to point to a prima facie reason for thinking that the naturalists are putting normativity in the right place.

Consider people who aim deliberately to mislead themselves. Suppose an elderly man realises that he is likely to be upset if he learns about the real probability of his developing cancer, and so arranges to avoid any evidence that might undermine his sanguine belief that this probability is low. Or suppose an adolescent youth learns that people with an inflated view of their own worth are generally happier and more successful, and so deliberately seeks out evidence which will make him think overly well of himself. Of course, there are familiar psychological difficulties about deliberately arranging to have false beliefs, but examples like this suggest they are not insuperable.[8]

Are these people acting wrongly? Of course, they aren't doing what they need to, *if* they want their beliefs to be true. But by hypothesis they don't want their beliefs to be true. So is there any other sense in which they are proceeding improperly?

It is not obvious to me that there is. I would say that it can sometimes be quite proper, as in the above examples, not to be moved by the aim of truth. However, if this is right, then there is a problem for those non-naturalist theories of content that explain content in terms of normativity, and so place the norms governing judgement inside the theory of content. For such theories make it constitutive of your possessing a belief with a certain truth condition that you be subject to norms which are apt to guide you to the truth.[9] In particular, if you believe that you have a low chance of cancer, or that you are are blessed with above-average attributes, then the norms which constitute these beliefs will apply to you, and so you won't be thinking as you ought to, if you have formed those beliefs with due regard to normatively authorised evidence.

Let me be clear about the difficulty I am raising here. The complaint is not that non-naturalist theories of content suppose

8. Note that these examples involve the manipulation of belief by the deliberate avoidance or pursuit of evidence, not any deliberate refusal to tailor your beliefs to the evidence you already have. I agree that cases of this latter kind are of doubtful psychological possibility. It is intriguing to consider exactly why this should be so.

9. Here I assume that the non-naturalists' putative content-constituting norms will include truth-conducive norms. Some possible non-naturalist theories of content might lack this feature, as I shall observe in Section VI. But it is common to all actual non-naturalist theories, and in particular to theories in the Dummettian and Davidsonian traditions.

an *indefeasible* requirement to seek the truth at all costs. Prescriptions can compete, and there is no reason why the prescriptions required by non-naturalist theories of content should not be overridden by other considerations, such as the desirability of not upsetting yourself unnecessarily. My objection is to the weaker claim, to which non-naturalist theories of content do seem to be committed, that there is always *some* reason to seek the truth, even if it can be overridden. For even this weak claim seems to me implausible. Why should we suppose that the elderly man, or the adolescent youth, are in any sense obliged to comport their beliefs to the evidence? It doesn't seem to me that they are violating any prescriptions at all by adopting their entirely sensible strategies.

Note how there is no difficulty here if you adopt the naturalist view that all 'oughts' relating to judgements are derived oughts, arising because personal or moral value is attached to truth. For there is no obvious reason to suppose that derived 'oughts' of either these kinds apply to cases like the elderly man or the youth. They don't want the truth, and it is not obvious that they are transgressing any moral boundaries in avoiding it. The naturalist view thus makes space for the possibility that people who aim to avoid the truth may nevertheless be acting entirely properly.

Perhaps some readers will wonder whether it is as easy as I am suggesting to lack personal or moral reasons for seeking the truth. Isn't there always at least some personal or moral reason for wanting the truth? I am doubtful about this, but note that, even if I am wrong here, this won't really help non-naturalism. For non-naturalists are committed to the view that the elderly man and adolescent youth are proceding improperly, *whether or not* there are moral or personal reasons for them to seek the truth. After all, non-naturalists believe in sui generis norms of judgement, which prescribe the pursuit of truth independently of moral or personal considerations. It is this underlying commitment to sui generis norms that I am objecting to, not the idea that there may always be reasons of some kind for pursuing the truth.

Still, even though it is not directly relevant to my line of argument, it will be instructive to digress briefly, and consider whether it is indeed possible to lack any moral and personal

reasons for seeking the truth. On the question of personal reasons, it might seem that there is at least one sense in which everybody has a personal reason to pursue the truth (even if they themselves suppose, like the elderly man or adolescent youth, that they will be better off without it). For doesn't it remain the case that everybody will be better able to satisfy their desires, whatever they may be, if their beliefs are true? Well, I agree that *if* you act appropriately on true beliefs, then your actions are guaranteed to satisfy your desires, and indeed I take this pragmatic connection to be a crucial component in the analysis of truth-conditional content (Papineau, 1993, ch. 3.6-3.7). And this pragmatic connection does mean that there is always a species of derived personal value to truth in beliefs that are relevant to action, for such truth will always help you to find a way of satisfying whatever desires you have.

However, this prescription only applies to those beliefs that *are* relevant to your actions, and so provides no reason at all for your wanting truth in beliefs which are not so relevant.[10] For example, if the elderly man cannot do anything about cancer, then his belief about his chance of developing it is unlikely to influence any of his actions, and it won't therefore matter to his success in satisfying his desires whether this belief is true or not. More generally, while we can agree that everybody has an interest in seeking the truth on action-relevant matters, such pragmatic considerations do nothing to show that people cannot quite sensibly avoid the truth on other issues.

So our personal desires need not always give us reason to pursue the truth. What about morality? Might it not be that the truth is always to be morally valued? Well, while it is uncontentious that the truth is morally valuable in some contexts (it is clearly morally valuable that we should have true beliefs about the dangers of smoking, for example), it is by no means obvious that it is *always* morally valuable. Does it matter *morally* if musicologists are mistaken about the drummer on Elvis's first

10. Note also that the pragmatic connection with successful action is by no means the only possible reason for personally wanting truth. You could want the truth for any number of reasons, or indeed for none. (In Section VII below I shall explain how the institution of language in particular provides one specific reason for wanting true beliefs.)

Sun session? Or again, does it matter morally if the elderly man remains hopeful that he won't get cancer, or that the youth thinks he is more estimable than he is?

Of course this issue deserves further discussion. It may well be that reflection on the overall role of truth in human life, its importance for social institutions, and so on, might show us that we always have a moral duty to respect the truth. Still, as I said earlier, even if this is so, it won't really help my non-naturalist opponents. For they are committed to truth having a value independently of any such considerations. They put the value of truth inside the theory of content, not outside with moral values. And this seems wrong. If further reflection were to show us that truth is *not* always of moral value, then in what sense would the elderly man or adolescent youth necessarily be acting wrongly in ignoring it?

IV

The Status of Judgemental Norms. I am aware that the points made in the last section will carry little dialectical force against non-naturalists. For clear-headed non-naturalists can simply repeat that their putative 'oughts' of judgement are special, quite different in kind from moral or personal 'oughts'. So, whether or not the elderly man and adolescent youth are not doing anything wrong by ordinary moral or personal standards, they will still be acting wrongly in a further sense, by violating the sui generis norms of judgement.

Even so, it has been worth rehearsing these issues, for they serve to highlight the peculiar status of the non-naturalists' putative norms of judgement. If these norms are quite distinct from moral or personal 'oughts', then where do they come from? What kind of fact is it that we categorically 'ought' to reason in certain ways? And whence does the guiding force of these 'oughts' derive—why *should* we reason in these ways? I know that I should do what is morally required. And there is an obvious sense in which I should do what will get me what I want. But I find myself in difficulty understanding why I should be moved by the non-naturalists' putative sui generis norms of judgement.

Thinkers in the non-naturalist tradition have in effect addressed this issue, but it cannot be said that they have arrived at any canonical answers. I have in mind here the discussion

surrounding Wittgenstein's 'rule-following considerations'. This is not the place to unravel all the different strands of that discussion, many of which are to do with issues of language. But I take it that one central issue which remains, after we put issues of language to one side (cf. Boghossian, 1989, p 514), is the question of why we ought to judge so-and-so on specific occasions. Why ought we to judge that *these* things are *green*? Why ought we to judge that *1,002* is the result of adding 2 to 1,000?

When we focus on this issue, non-naturalist theories seem to have little room to manoeuvre. It doesn't seem right to explain these 'oughts' in terms of individual dispositions. The fact that I am typically disposed to judge that certain things are green, say, doesn't make these judgements right, for I may be systematically prone to error. Nor does it help to widen the issue to some larger community, for whole communities too can be prone to error. In the end there seems to be little alternative for the non-naturalist except to hold that the norms of judgement are primitive and not to be further explained (McDowell, 1984). This seems unsatisfactory. Even if we can't reduce judgemental norms to other kinds of facts, it is surely desirable that we should have some kind of understanding of the peculiar force that judgemental norms are supposed to exert on us.

None of these difficulties about the status of judgemental norms arises on the naturalist line I am advocating. On my view, what makes it the case that you ought to judge in certain ways on specific occasions is that this will be a means to your judging *truly*. There is nothing circular about this analysis, provided truth itself can be analysed without appealing to norms of judgement, and the adoption of truth as an aim in turn explained by reference to moral or personal value attached to truth.

From my perspective, then, difficulties about rule-following are peculiar to non-natural theories of content. If you want to explain truth-conditional content in terms of norms governing judgement, as non-natural theories of content do, then you need to explain the source of these norms before you explain content, and it is not obvious there is any good way of doing this. But this problem simply does not arise for naturalist theories of content. They can explain content without mentioning norms. And then, given this, they can happily place norms of judgement outside the theory of content, as prescriptions which follow once moral or personal value is placed on truth.

Some readers may be wondering whether the naturalists' norms can suitably *guide* thinkers. We need norms of judgement not just to set a standard, but also to show thinkers how to conform to this standard. But bare prescriptions such as 'you should judge truly', or even such more specific imperatives as 'you should judge *green* of green objects', 'you should judge *n plus 2* when asked to add 2 to n', would seem not to fit this bill. They may articulate the aim of judging truly, either generally or on more specific matters, but they are not going to help thinkers who need guidance on how to achieve this aim.

This query calls for much detailed analysis, some of which will be sketched in Section VI below. But at this stage let me merely observe that the adoption of a given end will standardly generate derivative norms about appropriate means to that end. Just as the adoption of car speed as an end generates the prescription 'you ought to tune your car well', so will adoption of truth as an end generate prescriptions about means of arriving at truth. The reliabilist tradition in epistemology becomes relevant at this point. This tradition is concerned to identify belief-forming methods that generally produce true beliefs as output. These are just the kind of methods that you *ought* to use, if value is attached to truth. I would argue that inductive logic, deductive logic, standard observational procedures, and many other modes of reasoning are reliable methods in this sense. So, by derivatively prescribing these methods as means, the naturalist approach to normativity does not just assume truth as an aim, but also generates advice about how to achieve it.

At this point it is worth observing that reliability for truth is not the only desideratum that can be placed on belief-forming methods. At least two others offer themselves. It will normally also be desirable that our belief-forming methods are *important*, in the sense of delivering informative beliefs on matters of concern, and *economical*, in the sense of not using large amounts of time or other resources. Both these desiderata will generally pull against the requirement that belief-forming methods should deliver only true conclusions. Moreover, any moral or personal considerations that call for reliability are likely also to call for importance and economy, and so that some joint optimization of these mutually conflicting desiderata will usually be required.

For naturalists, then, reliability for truth is simply one potentially valuable feature of belief-forming processes, alongside

importance, economy, and possibly others. This marks a further contrast between the naturalists and their opponents, a further sense in which those norms of judgement which relate to the aim of truth are not sui generis for naturalists. For naturalists, truth-conducive norms of judgement constitute just one dimension among a number of possible ways of evaluating belief-forming methods. These norms have no special status which distinguishes them from evaluations orientated to further desiderata like importance and economy. This might seem odd to philosophers brought up to focus on truth-related issues, but it seems to me just the right attitude. There seems no good reason to privilege the desideratum of reliability for truth over others when we evaluate methods of belief-formation. True, an economical method that generates lots of important falsehoods is not generally worth much. But neither is a reliable method that consumes costly resources in generating trivialities.

Having said this, naturalists will do well to concede a significant terminological point to their opponents. When we consider whether a belief is '*justified*', or constitutes '*knowledge*', we evaluate the methods behind it purely from the perspective of reliability for truth, and in abstraction from such issues as importance and economy. Imagine a man who spends a month counting the individual blades of grass in his garden. We will no doubt feel this is a complete waste of time, and that the conclusion is of no possible interest to anyone, yet we will not say on this account that he does not *know* how many blades of grass there are, nor that his belief in their number is not *justified*.

It is striking how these central epistemological notions are exclusively truth-focused in this way. I am inclined to think that this tells us something about the central importance of truth-seeking for human society (Papineau, forthcoming). Still, we can recognize this importance, and the epistemological terminology it carries in train, without supposing that there is some special type of normativity attaching to the aim of truth.

V

Unrefined Thinkers. Let me turn now to another respect in which naturalist theories of content fare better than non-naturalist

ones. On the face of it, many beings who do not observe truth-conducive norms of judgement nevertheless seem to have contentful beliefs. The most obvious examples are higher mammals and very young children, who often act in ways that seem to call for the attribution of beliefs, but seem in no obvious sense to be observing norms of judgement. This poses a *prima facie* problem for those non-naturalist theories which make sensitivity to such norms a procondition of the possession of contentful beliefs.

In addition, some grown-up human beings pose a similar problem for non-naturalist theories of content. I am thinking here of the large body of psychological literature which suggests that normal humans characteristically violate principles of logic, elementary probability theory, and so on (cf. Cherniak, 1986, Stein, 1996). In this case it is not just that the believers lack the sophisticated abilities required to observe norms. They have the intellectual wherewithal, yet simply fail to apply it to the pursuit of truth. If this is right, then again we have a *prima facie* problem for those theories which make sensitivity to truth-conducive norms constitutive of the possession of contentful beliefs.

Non-natural theories of content seem to have little room to manoeuvre. Some philosophers, it is true, have argued that the psychological research just alluded to should not be taken at face value, and that, when the matter is looked into fully, the supposed violations of rational norms can be explained away (Cohen, 1981). I myself doubt that anybody without philosophical preconceptions would find this line plausible, but, even if it is right, the non-naturalist still has the infants and animals to deal with.

Here the non-naturalist has two options. Either infants and animals are guided by norms, or they lack beliefs. Neither option seems attractive. It is hard to see how animals and pre-linguistic children could in any sense be said to be observing norms. So non-naturalists have tended to go for the other option, suggesting that animals and infants do not really have beliefs. ('Dumb animals are natural beings and no more.' McDowell, 1994, p 70.) There seems little to recommend this move. It is one thing to recognise that there are significant differences between the beliefs of non-linguistic creatures and those of normal adult humans. It is another to inflate this into a fundamental division, with genuine thought on one side of the line, and nothing but mechanical response on the other.

Again, note how none of this is a problem for the naturalist. Naturalist theories of content regard judgement as prior to norms, and so have plenty of conceptual space for beings who have beliefs yet do not observe norms. Of course, specific naturalist theories of content will have specific views about the nature of judgement, and it will depend on the details exactly whom they recognise as believers. But by definition no naturalist theories will make sensitivity to norms itself a requirement for believing. (Moreover, for what it is worth, the specific naturalist theory I favour, teleosemantics, has no difficulty allowing that animals, infants, and systematically irrational adults all have beliefs.)

From the naturalist point of view, sensitivity to norms of judgement is an addendum to the possession of beliefs itself. Such sensitivity can arise in beings with beliefs, but does not have to. More specifically, given what I have said so far, it will arise in beings who adopt the end of true beliefs, for whatever reason, and consider what steps they need to take to achieve this end.

As it happens, I think adult human beings do this a great deal, and indeed that this plays a crucial role in human culture (Papineau, forthcoming). But note that the deliberate pursuit of true beliefs involves a sophisticated meta-representational concern *about* beliefs, and is therefore not something that we should expect to be present in all beings who have beliefs. Since young infants, and probably all animals, lack the notion of true belief, they will be incapable of sensitivity to such norms. And even adult human beings, who do have the notion of true belief,[11] don't always take pains to ensure that they are getting their beliefs from reliable sources, and often simply take them as they come.

These points now allow a better understanding of the psychological literature on widespread human irrationality. In particular, they suggest a answer to one obvious query about this literature, namely, that it seems to be belied by the success of human enterprises which call for high levels of doxastic accuracy

11. Footnote 4 implies that I am here taking ordinary adults to have a 'substantial account of truth-conditional content'. However, this needn't mean a sophisticated philosophical theory of content, like a teleosemantic theory, say, or a Davidsonian analysis, but only some way of telling which people have which beliefs with which contents.

('If we're so dumb, how come we sent a man to the moon?'). The answer is that humans normally reason badly, but can set themselves to do better. We can suppose that humans normally get their beliefs from 'quick and dirty' processes, which no doubt worked well enough in our evolutionary history, but all too often lead us astray in the modern world, and in particular in the experimental circumstances studied by psychologists. On the other hand, this is only part of the story. For humans can also deliberately set themselves to use reliable methods instead of the 'quick and dirty' tricks. Even if humans don't always follow norms that offer reliable routes to the truth, they are perfectly capable of this, and characteristically set themselves to do so when it matters.

According to the naturalist position, then, infant and animal believers lack the intellectual wherewithal to follow truth-conducive norms of judgement, and even adult human believers, who are intellectually capable of observing such norms, don't always do so. This position may be philosophically unfamiliar, but it has the virtue of plausibility. The usual condition, even for adult humans, is surely to think and act without worrying about what you are doing. Sometimes, I agree, we adults do make an effort to observe truth-conducive norms of judgement, and it is important that we be able to do this. But it overintellectualises ordinary mental life to suppose that this kind of effort accompanies all judgement.

VI

Some Specific Theories of Content. My remarks so far have contrasted 'naturalist' and 'non-naturalist' theories of content. In this section I want to attend to some different species within these broader genera, and note briefly how the details affect some of the points made so far. I think these refinements are of some intrinsic interest, but they will not affect my overall line of argument, so those who have been comfortable with the level of abstraction adopted so far may wish to skip to the next section.

The distinction between natural and non-natural theories of content interacts with that between 'output-based' and 'input-based' theories. By 'output-based' I mean those theories, like success-semantics and teleosemantics, which analyse the content of

beliefs in terms of the *actions they prompt*; and by 'input-based' I mean those theories, like verificationism and Dummettian theories, which analyse the content of beliefs in terms of the *conditions that produce them*. In more detail, output-based theories aim to equate the truth conditions of beliefs with those conditions that guarantee that resulting actions will satisfy desires (with success sematics then differing from teleosemantics in its account of desire satisfaction[12]), while input-based theories aim to equate truth (or assertibility) conditions with the circumstances that typically give rise to beliefs.

When I first started thinking about these matters, I saw the conflict as primarily between 'output-based' naturalistic theories and 'input-based' non-naturalistic ones. This was because the naturalist theories I favour are in fact output-based, while the most familiar non-naturalist theories are in fact input-based. However, further reflection, showed me that I was running two issues together here. The input versus output issue is orthogonal to the natural versus non-natural issue. After all, there are also naturalistic input-based theories, like indicator semantics and Fodor's asymmetric dependence theory. And, even if there is no clear-cut example of a non-natural output-based theory, there are elements of this position in Davidsonian approaches to content.

These distinctions complicate what I have said so far. Take first the division of *naturalist* theories of content into output-based (teleosemantics, success semantics) and input-based (indicator semantics, asymmetric dependence). I have suggested that naturalist theories of content can regard truth-conducive rules guiding belief-formation as specifications of means appropriate to the valued end of truth. But this comes out somewhat differently for output-based and input-based naturalist theories. For output-based theories, all such means will be *contingently* connected to the end of truth: since what makes a belief true are circumstances picked out by the belief's relation to action, it will be an independent further fact that any given way of arriving at

12. See Whyte, 1991, Papineau, 1993, ch 3.8. In this paper I am ignoring philosophical questions about satisfaction conditions for desires. There are non-naturalist as well as naturalist answers to these questions, as there are for truth conditions and beliefs.

the belief should track the truth. With input-based theories it works differently. Since the circumstances that make a belief true will now be equated with some subset of the possible causes of that belief, any recommendation that this belief should issue only from those causes, while still construable as a derived imperative orientated to the valued end of truth, will now be recommending a course that is *constitutively* connected with the end of truth.

Given this, we might expect a related contrast on the output side. That is, we might expect output-based naturalist theories to view some aspect of output as constitutively related to truth, which input-based natural theories regard as only contingently connected. There is indeed such a contrast. To see this, note first how input- and output-based naturalist theories are a kind of mirror image of each other. Just as an input-based theory needs to distinguish, among all the circumstances which can actually cause beliefs, those canonical conditions which constitute their truth conditions (this is why Fodor's asymmetric dependence, or something similar, is needed), so also does an output-based theory need to distinguish, among all the behavioural choices which can actually be occasioned by beliefs, those canonical behavioural consequences that play a role in constituting truth conditions: for it is the specific circumstances in which *those* canonical choices lead to successful actions that constitute the truth conditions. (The point is that we need to winnow out choices occasioned by rushes of blood, acting in haste, and so on, or we will get the truth conditions wrong. It is a challenge for a naturalistic output-based theory to make this distinction naturalistically, a challenge which I think is best met by teleological considerations. Cf. Papineau, 1993, 3.7.7.)

Now consider the recommendation that people with true beliefs *ought* to choose in those canonical ways, *if* they are to satisfy their desires, first from the perspective of an output-based naturalist theory, and then from the point of view of a input-based one. From the output-based perspective, this hypothetical imperative falls constitutively out of the nature of truth: if your belief is true, and you so choose, then, by the nature of truth, you will satisfy your desires. For, on the output-based theory, true beliefs are in effect identified as those which lead to desire satisfaction specifically when they combine with such canonical choices. On an input-based theory, by contrast, this hypothetical

imperative (that you *ought* to choose in the canonical ways, *if* you are to satisfy desires) will depend on contingent facts. For there is nothing in an input-based analysis of content (truth conditions = canonical *sources*) to ensure that if you have true beliefs, and choose in the canonical ways, your desires will be satisfied. This may be true, even for an input-based theorist, but if so it will be because of contingent facts about the world, and not because of the nature of truth.

Let me now quickly consider these matters from the point of view of non-naturalist theories of content. On input-based non-naturalist theories, like verificationism or Dummettian theories, recommendations about appropriate sources for beliefs will be taken to be primitive norms which fix truth (or assertibility) conditions for those beliefs. These theories are thus akin to input-based naturalist theories like Fodor's, in fixing truth conditions by reference to appropriate sources, except that they hold that there is an essential normativity involved in the reference to 'appropriate'. Moreover, they seem to come out exactly like Fodor-style theories on the output side. When it comes to recommendations that people with true beliefs *ought* to choose in such-and-such canonical ways, if they are to satisfy their desires, an input-based normative account offers no constitutive rationale. Just as with Fodor, they can still say that these choices are sensible because they will ensure satisfaction of desires, if your beliefs are true. But there is no constitutive tie here, and this claim will be true, if it is, because of contingent facts about the world.

What about output-based normative theories? Nobody seems actually to hold such a theory, but it is easy enough to imagine one. It would be like success-semantics or teleosemantics in fixing truth conditions as those circumstances in which appropriate choices will succeed, except in holding, again, that 'appropriate' here is primitively normative. So on the input side this theory wouldn't disagree with its output-based naturalist counterpart: it would view belief-forming recommendations as derived imperatives specifying that certain sources are an effective means, as a matter of contingent fact, when value is attached to the end of truth.

Even if there are no exclusively output-based non-natural theories of content, the Davidsonian approach to content can be viewed, as I said, as embodying elements of this approach. For

we can view Davidson as in effect upholding a non-naturalist theory of content which is *both* output-based and input-based, in that he identifies beliefs as states which ought primitively to arise from such-and-such epistemological inputs *and* which ought primitively to lead to such-and-such choices.

Won't a Janus-faced theory like Davidson's at some point face an embarrassing doubling of primitive and derived norms? In its intial identification of beliefs, the theory invokes both primitive input and output norms. But there seems no barrier to a Davidsonian also recognizing derived norms: derived input norms arising whenever moral or personal value is attached to the end of truth, and derived output norms about which choices you ought to make, if you are to satisfy your desires. This seems like an overabundance of norms. Surely we do not want to allow *both* primitive *and* derived norms prescribing effectively the same strategies for forming beliefs and making choices.

I think this is indeed a genuine problem, but it is not peculiar to Davidson's theory. Rather it is a problem that arises for all non-naturalist theories, and indeed is simply a special case of the problem I raised in Section III. We can put the point like this. Naturalists (whether input- or output-based) want to construe all norms relating to beliefs as derived imperatives: input norms are orientated to moral or personal values placed on truth, while output norms are orientated to the satisfaction of desires. Non-naturalists, by contrast, hold that at least some norms (input, or output, or both) are prior, sui generis and non-derived, and moreover that these prior norms underpin the constitution of content for beliefs. But this means that non-naturalists of any kind are going to be threatened with some doubling of norms. For there will always be room for corresponding derived norms, orientated to the value of truth, or to the satisfaction of desires, alongside whichever norms non-naturalists take to be primitive and content-constituting. (Of course non-naturalists, by definition, will suppose that these primitive norms come first, so to speak, in that they are needed to constitute truth-conditional content; but the trouble is that, once content has been so constituted, non-naturalists won't be able to stop extra derived norms arising, because certain ways of reasoning are additionally valued as a means to the end of truth, or because certain ways of choosing are additionally valued as a means to the satisfaction of desires.)

In Section III I pressed the implausibility of there being primitive norms for truth-seeking, as required by input-based non-naturalist theories, in addition to derived norms which arise when value is attached to truth.

We can now see that the point generalizes. An output-based non-naturalist theory would have primitive output norms (you ought to choose in certain canonical ways) *and* the hypothetical recommendation that you ought to so choose, if you are to satisfy your desires. And a Davidsonian Janus-faced theory will have primitive norms for both outputs and inputs, *and* in addition both derived input and output norms.

I take the general point to argue against non-naturalist theories of any kind, just as the special case in section 3 argued against non-naturalist theories with primitive input norms. Primitive norms are an embarrassment, given the extent to which derived prescriptions are available anyway.

VII

Norms of Language. In this final section I want briefly to consider the lingusitic dimension of judgement. Obviously, in the space available, much will have to be taken for granted. Let me cut through a number of issues by adopting the following simple two-part model for an assertoric language L. (1) *Knowledge of truth conditions*. For all sentences s in L, speakers agree on a claim of the form: s is true if and only if p.[13] (2) A *rule of truth*. Utterances are governed by the prescription, 'You should utter s only if (you know that) p'.[14]

Given anything like this model, linguistic judgements will uncontroversially be constitued in part by a normative rule of truth. I have argued that content for non-linguistic judgements is not constituted by norms. By contrast, assertoric utterances would clearly not have their contents without a rule of truth. If

13. Perhaps they need also to know that such claims are derived canonically from facts about the semantic roles of s's parts, if this is to constitute knowledge of truth conditions. Alternatively, they may derive such claims from claims of the form that *s means that p*, or *s expresses the belief that p*, or *s would be true iff p*. I shall skip these refinements.

14. See Williamson (1996) for arguments in favour of inserting the knowledge requirement here.

it were not the rule that s is only to be asserted when p, then utterances of s would not have the significance they do. (Maybe s would still in some sense be *true* if and only if p, given its agreed truth condition, but it is only in virtue of the rule of truth that we can regard those who utter s as *committed* to p, as offering the judgement that p—as opposed to conjecturing that p, say, or denying that p, or whatever.)

So I accept that linguistic judgments are in this sense normatively constituted, if anything like the simple model above is right. However, I do not accept that there is anything in this to trouble naturalism. To start with, note that the issue of normativity arises rather differently in the present context from the way it did in earlier sections of this paper. In earlier sections I was concerned to explain how naturalists can *uphold* claims of the form 'you ought to judge in such-and-such ways'. That is, I wanted to show how naturalism can accommodate the *truth* of such normative claims. But there is no need for me, as a naturalist theorist, to *uphold* the norm which I take to constitute linguistic assertions. What matters is that the speakers of the relevant language accept the rule of truth for assertions, not that this rule is true. It doesn't matter whether they *really* ought only to say what is true. Their assertoric practice will still contain a rule of truth, which will constitute certain utterances as assertions, as long as they *think* they ought only to say what is true.

In sociological parlance, what matters is that the rule of truth is a *social fact* in the linguistic community in question, not that it is in fact a good rule. As such, it is akin to social facts proscribing theft, extra-marital sex, broken promises, driving on the right, or splitting infinitives. Such social facts are constituted by the evaluative attitudes of the members of the relevant community, quite independently of whether those attitudes are normatively justified. Even if extra-marital sex is in fact quite proper in many circumstances, it would still be true that in many societies there is a social norm prohibiting it, in the sense that the members of those societies disapprove of it and will exert pressure to prevent it.

Now, there are many interesting philosophical questions about the constitution of social facts. For example, there seems reason to distinguish social facts which depend on the community's moral attitudes (to sex, theft, promises), from those which rest

on feelings about etiquette or style (splitting infinitives), and again from those which derive from a common interest in coordination (driving on the left). And it is by no means clear in which of these three categories the rule of truth ought to be placed. But, wherever we put it, there seems nothing to trouble the naturalist. Social facts of all kinds are constituted by the psychological attitudes of the members of the relevant society. In the earlier sections of this paper I have sought to rebut the most common argument for thinking that naturalism cannot accommodate psychological attitudes. If I am right, and naturalism can accommodate psychological attitudes, then it should be able to accommodate social facts in general, and the rule of truth in particular.

I have argued that the constitution of languages presupposes only that the relevant rule of truth is a social fact, and not that is true. But it is worth observing that, if a rule of truth is indeed a social fact, then members of linguistic communities will generally have some kind of pro-attitude towards saying only what is true, and this in itself will imply hypothetical imperatives about what they ought to do to to get the truths they want. Moreover, if it is morally valuable that languages exist, as it arguably is, then this in turn will imply that people ought morally to avoid false utterances.

It is also worth observing that any such prescriptive truths about *utterances* will also generate prescriptions about the avoidance of false *belief*. If there is moral or personal value in speaking truly, then there will be derived moral or personal value in believing truly, since aiming to believe truly is a presupposition of aiming to speak truly. In preceding sections of this paper I have considered what follows if true beliefs have moral or personal value. We now see how the institution of language can add to the reasons for which true beliefs might have such value.

Let me finish by briefly addressing the issue of the relation between thought and language. My overall expository strategy may have seemed to commit me to the priority of thought over language. After all, I have explained representation, in the first instance, by applying naturalist theories of content to beliefs, without attending to any involvement beliefs may have in language. And, more specifically, when I have turned to language, I may seem to have credited speakers' with language-prior abilities to *think* that p, in explaining their understanding of sentences in terms of beliefs of the form *sentence s is true iff p*.

However, I don't think there is in fact anything worrisome here. Despite any initial appearances to the contrary, I am happy to concede that thought depends on language in ways I have yet to address. But I do not accept that this dependence invalidates my overall analysis. More specifically, it does not preclude the application of naturalist theories of content to beliefs, nor does it preclude the explanation of assertoric language in terms of speakers' beliefs about truth conditions.

It is worth distinguishing two ways in which belief depends on language: psychodevelopmentally and constitutively. To start with, linguistic training obviously exerts a developmental influence on psychological abilities. Many thoughts are clearly too complex for humans whose perceptual sensivities and inferential associations have not been shaped by a relevant linguistic upbringing. In addition, there is also a more fundamental, constitutive dependence of thought on language. I am thinking here of those now familiar cases where my possession of a concept does not depend on individual factors like my perceptual sensitivities and inferential associations, but rather on the structure of my linguistic community. Names of people (*Aristotle*) or kinds (*arthritis*) provide obvious examples. Here it is not that my linguistic upbringing has caused me to develop new internal states. Rather, my possessing such concepts is *constituted* by my being embedded in a linguistic community. If you change the community, while leaving me the same, my concept changes.

I deny that there is anything in either of these kinds of dependence that is inconsistent with the overall story I have told in this paper. The first point to address is my use of naturalist theories to analyze belief representation. Now, such theories need clearly have no difficulty with the fact that linguistic training affects psychological sensitivities and abilities. For they can simply embrace these linguistically developed abilities as part of what fixes the representational contents of beliefs. Moreover, they can make just the same move with respect to the social consititution of content. Again, they can simply embrace this, and agree that certain of your states represent such-and-such, not just because of intrinsic facts about you, but also because of your relationships to other people, and in particular because of your embedding within a lingusitic community. Of course, exactly how these points might work will depend on the specifics of different

naturalist theories of content. But it is not hard to see in outline how my development of special psychological abilities, or my membership of a linguistic community, might affect the informational significance of my states, or their teleological purposes, or their success conditions. All in all, there seems no obvious reason why naturalist theories of belief content should not simply accept that beliefs depend on language in the two ways I have indicated.

The other point where I might have seemed to be committed to the priority of thought over language was where I assumed that assertoric language depends on speakers' beliefs that sentences s are true iff p. The worry here is that I apparently need to assume that speakers can think p *before* they learn sentences which mean that p.

However, I do not in fact need any such assumption of temporal priority. I need only claim that once the linguistic institution is up and running, as it were, and speakers are capable of thinking that p, then the different parts of the institution, thinking and speaking, are coordinated by beliefs of the form: s is true iff p.

True, this will leave us with an interesting kind of circularity in some cases. Speakers have concepts like *arthritis* because they belong to a linguistic community. This community exists in part because speakers follow rules for the linguistic expression of these concepts (they aim to assert 'arthritis' sentences only when certain arthritis facts obtain). But speakers will only be able to follow these rules if they already have the concepts (they won't be able to grasp the rules for 'arthritis' sentences unless they are capable of *arthritis* thoughts). So the concepts are constituted by the rules which are constituted by the concepts.[15]

However, while I accept there is a circle here, I do not think it vicious. It is just a special case of certain phenomena (here, mastery of concepts) being constitued by rules which in turn presuppose those phenomena. Many familiar social phenomena display this feature. There wouldn't be soccer matches unless there were rules about who wins soccer matches. There wouldn't be

15. It is specifically the constitutional, not psychodevelopmental, dependence of thought on language that generates this interesting circularity. It is scarcely noteworthy that a new generation's concepts should causally depend on the linguistic practices of older people with those concepts.

bank balances were there not regulations about who has access to bank balances. Certain concepts wouldn't exist were there not norms governing the linguistic expression of those concepts. This kind of circularity is certainly interesting, but it does not present any more of a special philosophical problem than soccer matches or bank balances do.[16]

REFERENCES

Boghossian, P., 1989, 'The Rule-Following Considerations', *Mind* 98.
Brandom, R., 1994, *Making it Explicit* (Cambridge, Mass.: Harvard University Press).
Broome, J., Forthcoming, 'Relative Rationality'.
Cherniak, C., 1986, *Minimal Rationality* (Cambridge, Mass.: MIT Press).
Cohen, L. J., 1981, 'Can Human Irrationality be Experimentally Demonstrated?', *Behavioural and Brain Sciences* 4.
Dretske, F., 1981, *Knowledge and the Flow of Information* (Oxford: Basil Blackwell).
Fodor, J., 1987, *Psychosemantics* (Cambridge, Mass.: MIT Press).
Fodor, J., 1990, *A Theory of Content* (Cambridge, Mass.: MIT Press).
Horwich, P., 1990, *Truth* (Oxford: Basil Blackwell).
McDowell, J., 1984, 'Wittgenstein on Following a Rule', *Synthese* 63.
McDowell, J., 1994, *Mind and World* (Cambridge, Mass.: Harvard University Press).
McGinn, C., 1989, *Mental Content* Oxford: Basil Blackwell.
Millikan, R., 1984, *Language, Thought, and other Biological Categories* (Cambridge, Mass.: MIT Press).
Millikan, R., 1993, *White Queen Psychology and other Essays for Alice* (Cambridge, Mass.: MIT Press).
Papineau, D., 1984, 'Representation and Explanation', *Philosophy of Science* 51.
Papineau, D., 1993, *Philosophical Naturalism* (Oxford: Basil Blackwell).
Papineau, D., Forthcoming, 'The Evolution of Knowledge' in *The Evolution of Mind*, ed. P. Carruthers (Cambridge: Cambridge University Press).
Stampe, D., 1977, 'Towards a Causal Theory of Linguistic Representation' in eds. P. French, T., Ueling and H. Wettstein *Midwest Studies in Philosophy* 2.
Stein, E., 1996, *Without Good Reason* (Oxford: Clarendon Press).
Whyte, J., 1990, 'Success Semantics', *Analysis* 50.
Whyte, J., 1991, 'The Normal Rewards of Success', *Analysis* 51.
Williamson, T., 1996, 'Knowing and Asserting', *Philosophical Review* 105.

16. Earlier versions (some very early) of this paper were read at the Scottish Postgraduate Philosophy Society, King's College London, University of Nottingham, University of St Andrew's, and the Cambridge Moral Sciences Club. Thanks are due to all who contributed on these and other occasions, and in particular to Robert Black, John Broome, Edward Craig, Peter Goldie, Christopher Hill, Paul Horwich, Neil Manson, Alan Millar, Gabriel Segal, Scott Sturgeon, Julia Tanney, and Timothy Williamson.

NORMATIVITY AND JUDGEMENT

David Papineau and Julia Tanney

II—Julia Tanney

NORMATIVITY AND THOUGHT

ABSTRACT This paper attempts to describe why it is not possible to account for normative phenomena in non-normative terms. It argues that Papineau's attempt to locate norms of judgement 'outside' content, grounded in an individual's desires or reasons, mislocates the normativity that is thought to resist appropriation within a 'world that conceives nature as the realm of law'. It agrees, however, that a theory of content that locates norms 'inside' content will not be forthcoming—at least if this is to require fashioning the norms that in some sense govern judgment or thought into individually necessary conditions for contentful states.

I

In 'Empiricism and the Philosophy of Mind',[1] Wilfrid Sellars claims that the attempt to analyze epistemic facts in terms of non-epistemic facts is 'a radical mistake—a mistake of a piece with the so-called "naturalistic fallacy" in ethics'.[2] I suspect that the attempt to reduce the phenomena or otherwise account for epistemic episodes or states in terms of non-epistemic ones would also be a mistake, and that, generally speaking, the attempt to account for conceptual, intentional, or semantic phenomena in terms of non-conceptual, non-intentional, or non-semantic phenomena would be instances of a similar mistake. For Sellars, what distinguishes cognitive episodes, or conceptual awareness, for example, from the having of impressions or sensations (which, properly understood, are states of the perceiver) is the former's role in the 'framework of thoughts'.[3] In characterizing

1. Sellars, 'Empiricism and the Philosophy of Mind' originally published in *Minnesota Studies in the Philosophy of Science*, vol. 1, ed. Herbert Feigl and Michael Scriven (Minneapolis, University of Minnesota Press, 1956) and reprinted with an Introduction by Richard Rorty and a Study Guide by Robert Brandom as *Empiricism and the Philosophy of Mind* (Cambridge, MA, University of Harvard Press, 1997).

2. ibid.; 19.

3. ibid.; 110.

a state in cognitive or epistemic terms, he says, '...we are not giving an empirical description of that episode or state, we are placing it in the logical space of reasons, of justifying and being able to justify what one says'.[4]

The thought is echoed by a number of subsequent writers: notably by Davidson in his arguments that rationality has no echo in physical theory,[5] by McDowell who claims that 'the structure of the space of reasons stubbornly resists being appropriated within a naturalism that conceives nature as the realm of law',[6] and by Brandom in his programme for a normative pragmatics.[7] The nature of the radical mistake that is supposedly involved in accounting for epistemic, conceptual, intentional, semantic, or for short, 'normative' phenomena in non-normative terms is, however, notoriously obscure.

II

David Papineau denies that the normativity of conceptual thought or judgement poses a problem for naturalism. He wishes to defend the idea that norms of judgement are 'derived from moral or personal values and do not involve any sui generis species of conceptual normativity'.[8] The force of prescriptions such as 'in order to achieve truth, you ought to judge in such-and-such ways' derives, he maintains, from 'independent moral or personal reasons for attaching value to truth'. Papineau's strategy in defending this idea involves assuming a naturalized view of (truth-conditional) content (one which does not itself presuppose norms and one in which the notion of truth itself is construed as a purely 'descriptive' property). The strategy also involves pointing to obstacles for accounts that locate normativity 'inside' content.

The idea that there could be a 'natural' theory of content is, however, precisely what should be challenged by philosophers

4. ibid.; 76.

5. *Essays on Actions and Events* (Oxford, Clarendon Press, 1980); especially essays 11 and 12.

6. McDowell, *Mind and World* (Cambridge, MA, Harvard University Press, 1994); 73.

7. Brandom, *Making It Explicit* (Cambridge, MA, Harvard University Press, 1994).

8. Papineau, pp. 17–18.

who are impressed with the normativity of thought and its ostensible recalcitrance vis-à-vis the 'realm of law'.[9] They should not be willing to concede that the 'real norms' are ones that can be located 'outside' a theory of content. But they should also share Papineau's pessimism about the availability of non-circular theories of norm-governed content. No theory of content (reductive or non-reductive) can be given if this is supposed to involve fashioning the norms that (in some sense) govern thought or interpretation into non-circular necessary and sufficient conditions for the application of mental concepts.

The idea that the 'real' norms are ones that can located 'outside' content should be resisted because it misidentifies part of the normative space that opens up when an item is identified in cognitive or epistemic terms. This can be seen by considering the aspect of Papineau's proposal that invites us to construe the 'real' norms as hypothetical imperatives whose prescriptive force is conditional on a believer's desires.[10] There are general questions whether reason explanations (and in particular, those relevant to an action's moral evaluation) need a conative or motivational element in addition to a cognitive one in order to be truly explanatory. But this familiar dispute about hypothetical vs. categorical prescriptions cannot even get off the ground when the content of the norms with which the agent is in supposed cognitive touch are the alleged norms that govern truth or rationality. Whatever there is to the idea that conceptual space is a norm-governed space or, more specifically, to the thought about the normativity of beliefs crystallized in the phrase 'beliefs aim at the

9. Papineau has anticipated this response and has attempted to structure his argument so that the move is question-begging. He claims that his defense of the conditional claim ('*If* you have a naturalist theory of content, *then* you can explain norms of judgement as derived prescriptions orientated to the end of truth') (p. 22)) 'blocks the argument that naturalist theories of content are inadequate because they cannot account for the normativity of judgement.' (p. 22) He agrees that theories that locate norms 'inside' content call for a species of judgemental norms that resist appropriation within the realm of law. But he rejects the existence of such special content-constituting norms, since he hopes to be able to explain content without any reference to such prior norms in a theory that locates them 'outside' content. Techically, then, I agree with Papineau's conditional thesis, but for disappointing reasons: because the antecedent is false. The burden of my discussion in this section is to show why it is not open to Papineau to regard the content-constituting norms as artefacts of the wrong theory of content.

10. The following arguments also tell against the thought that norms are categorical imperatives (that, for instance, derive from the ostensible moral value of truth).

truth',[11] this cannot be construed as a command issued to persons to conform or mould their judgements to meet the requirements of truth, or as giving them a reason (categorical or hypothetical) to form only true beliefs.

Belief-acquisition is not an action. Although one can form strategies (for example, by avoiding relevant evidence) that will maximize or minimize one's chances of acquiring some particular belief, forming beliefs is not something one chooses to do. Furthermore, the acquisition of true or false beliefs, even if this were construed as a quasi-action, could not possibly be explained by attributing to an individual a desire or an aversion to conform to a norm. In order for such an explanation to get off the ground, the individual would have to have beliefs about the norm—how else could it be the object of her desires?—and presumably questions regarding the truth of her beliefs about the norm would simply resurface. It ought to be clear in any case that the explanatory mileage gained by attributing to an individual the desire to seek the truth would be negligible, at best (even if this does not involve putting her in cognitive contact with a norm). The desire for truth can only explain one's decision to stop visiting the astrologer if one comes to believe, for example, that the interpretation of horoscopes is not a reliable process for gleaning the truth. But how is the truth-aspiration of the belief about the unreliability of astrology supposed to be understood? It is presupposed.[12]

Contentful states cannot play a role in reason-explanation unless it is taken for granted that beliefs 'aim at the truth'. The contingent desires of the agent have nothing to do with this supposition: it is a conceptual truth about beliefs. There are, of course, special circumstances in which it makes sense to talk about a person who wants truth, or who wants to avoid it, just as there are special circumstances in which it makes sense to talk about a person who might need to be convinced that some of her belief-forming processes are unreliable. The fact that these are

11. Bernard Williams, 'Deciding to Believe' in *Problems of the Self* (Cambridge, Cambridge University Press, 1973); 136–151.

12. Although I have framed this as a problem inherent in the idea that someone might be said to obey or to want to accord with semantic norms, the problem easily generalizes to include the attempt to construe rationality as a matter of obeying norms of rationality. See my 'De-Individualizing Norms of Rationality', *Philosophical Studies*, vol. 79 (1995); 237–258.

special circumstances can be seen by reflecting on the fact that 'the game of giving and asking for reasons', the practice of justification, and, incidentally, the whole notion of *explanation* presupposes truth as both an ideal and as the norm. The framework of thought, or the logical space of reasons, comes 'prepackaged' with truth and its preservation as standards. Although instances can be be isolated in which a person fails to conform to *modus ponens*, or seeks to avoid the truth, or acts against her better judgement, this is accomplished against a general background in which truth, inference and rationality remain intact.[13]

At least part of the normativity that constitutes the 'logical space of reason' attaches to this *background* space; it attaches to what is presupposed when items are identified in conceptual space, or in cognitive or intentional terms. The attempt to locate the 'norms of thought' or the 'norms of judgement' as grounded in a person's desires or reasons therefore mislocates part of the normativity that is said to resist appropriation within a world that conceives nature as the realm of law.

III

I have argued that Papineau misconstrues the normativity that attaches to what Sellars calls the 'framework of thought' or the 'logical space of reasons' and I have suggested that the right way to conceive it is in terms of the relations that are taken for granted in reason-explanation. I have also hinted at the futility of construing the norms that are extracted by studying these relations as playing a reason-giving role in an explanation of an individual's rationality. In other words, attributing to an individual beliefs about, or the desire to conform to, a norm of rationality cannot explain—on pain of regress—her ability to act rationally or her ability to form true beliefs: these abilities are

13. This should not be taken to imply that the background space (or any one of its features) is itself something that cannot be put in jeopardy; but this could only be done once the framework for dismantling it no longer presupposes it. See Sellars's claim (p. 79, op. cit.) that '...empirical knowledge, like its sophisticated extension, science, is rational, not because it has a *foundation* but because it is a self-correcting enterprise which can put *any* claim in jeopardy, though not *all* at once.' Correlatively, the fact that this framework is presupposed should not be taken to imply that it is necessary: human patterns of thought might have been very different. See Barry Stroud, 'Wittgenstein and Logical Necessity', *Philosophical Review*, vol. 74 (1965); 504–518.

presupposed in the very idea that she has beliefs and desires. This last thought is related to the pessimism I endorsed about the availability of non-circular accounts of items identified within the space of reasons, since it seems that the phenomena to be explained are presupposed by the kind of explanation on offer. I shall come back to this thought later. First, more needs to be said about why in particular it is a mistake to attempt to reduce items within the 'framework of thought' to the realm of law.

One way of approaching the question would be to see what is missing when such a reduction is attempted. Calibrating something against a norm, ideal or standard allows it to be evaluated in terms of its relative success or failure in achieving this ideal or standard. The appeal to semantic norms (e.g., that *vert(e)* means green) or to principles of rationality (e.g., that a person tends to act in accordance with her better judgement) allows a diagnosis of, for example, what has gone wrong in a particular case when a language-user misunderstands a word or when a rational agent acts irrationally. Items identified within a normative space admit of this special kind of error—'internal' error—which allows both that the one who commits the error is still participating in the practice that the norms govern and that her mistaken moves are still moves within that practice. Someone who makes mistakes in inference is still a thinker; someone who mistakes the meaning of a word might still say something meaningful; someone who acts irrationally is still acting in a way that is rationalizable; and someone who believes something falsely still has a belief. The phenomena to be explained—thinking, acting, saying, believing—allow 'internal' mistakes of this sort; mistakes, like an *error* in baseball, that do not disqualify the player and is still considered to be a move within the game.

Various attempts to assimilate normative phenomena to law-governed phenomena have run into difficulty accommodating this kind of error. (In contemporary philosophy of mind, the difficulty shows up in the computationalist's attempt to construe the algorithmic/heuristic rules that govern transitions between representations as themselves instantiated causally (in the neuronal structure of the brain) and it shows up in the representationalist's attempt to give causally necessary and sufficient conditions of what it is to be a content-bearing state.[14] It also shows

14. I argue for both these points in 'How to Resist Mental Representations', (A Critical Notice of Tim Crane's *The Mechanical Mind*), *International Journal of Philosophical Studies*, vol. 6, no. 2 (1998); 264–278.

up in certain functionalists' construal of mental terms as theoretical posits in a causal/nomological 'theory of mind'.[15]) But the difficulty accommodating 'internal error' is not merely a problem for naturalists: it is a problem for anyone who misunderstands the nature of the platitudes or principles that are presupposed with the application of concepts that have their home in the 'logical space of reasons'. It is a difficulty, that is, for any philosophical account of the mental that misunderstands the kernal of truth contained in the idea, for example, that it is 'constitutive of [one's] possessing a belief with a certain truth condition that [one] be subject to norms which are apt to guide [one] to the truth.'[16] If the 'instantiation' of such an norm (as it were) is conceived *either as a causally or a logically necessary condition* of having a belief, being an action, or being a perception, then false beliefs, irrational actions, and misperceptions will be impossible. Take any norm that allegedly governs, say, intentional action: if the instantiation of that norm is a necessary condition for intentional action, then irrational actions cannot be described as the failure of the instantiation of that norm. This idea gets to the heart of the philosophical paradox of irrationally,[17] and the way out is to recognize that the norms that govern a practice do not provide rigid conditions of what is to count as a move within that practice. The appeal to norms supports the distinction between successful moves and unsuccessful ones: true and false beliefs, misperceptions and veridical ones; rational actions and irrational ones.

The idea that 'psychological principles' can be extracted from an investigation of our reason-giving practices is nonetheless attractive. (As is the idea that there are underlying principles which might be extracted by looking at our ethical practices.) Suppose we met with some success in extracting such principles or norms by watching how people correct themselves and others

15. See Tanney, 'Why Reasons May Not Be Causes', *Mind and Language* vol. 10, nos. 1/2 (1995); 103–126.

16. Papineau, p. 24.

17. In 'De-Individualizing Norms of Rationality', op. cit., I suggest that this might be interpreted as the following argument in which seemingly acceptable premises lead to an unacceptable conclusion: 1) The instantiation of rational norms is necessary for intentional actions. 2) But irrational actions are nevertheless intentional actions. 3) Irrationality is the violation of a rational principle. 4) Therefore, irrational actions are both possible and impossible. A paradox is generated if the first premise reads that for every norm, its instantiation is necessary for intentional action.

and how they justify these corrections. The list of generalizations that results might form a kind of normative guide for participating in the practice: a list of dos and don'ts. And knowledge of the norms might give us some understanding of the nature of the practice. But the sense in which this normative guide forms a *theory* would probably turn out to be similar to the sense in which adumbrating the dos and don'ts of etiquette or classical piano-playing forms a theory. One would not think that the norms are exhaustive, that they suffice to govern allowable moves, that they are exceptionless, or that they are even stable. But this is a very different conception from that which is involved in discovering natural laws.

IV

A controversial question (even among those sympathetic with this general line of thinking) is whether one can view the nature of our reason-giving practices independently of the choices, decisions, and reactions of the participants; the answer would be 'no' according to those who are sympathetic with the idea that the choices, decisions, and reactions may be partially constitutive of the protean phenomena under investigation. This is certainly the case for etiquette and piano-playing, and plausibly the case for aesthetics and ethics.[18] If this intuition were true for our reason-giving practices as well, it would pose another threat to a 'naturalist' programme which construed its investigation of the mental along the lines of a scientific model, since the whole point of a scientific investigation is to prescind from subjective features as much as possible.

I suggested earlier that at least part of the normativity that constitutes the 'logical space of reason' attaches to what is presupposed when we identify something as an item in conceptual space or in cognitive or intentional terms. The normativity attaches to this background space, I argued, quite independently of considerations having to with particular reasons of particular individuals. But I think that another part of the normativity that

18. I note here that I am passing quickly over an important and interesting distinction between individual choices, decisions, and reactions, and group choices, decisions, and reactions.

attaches to the logical space of reasons will concern particular individuals and their 'points of view'.

In an earlier work, Papineau suggests that the only serious candidate for an argument that there are certain principles of thought necessarily common to all thinkers is the argument that 'the interpretation of human thought is a distinctive matter of *understanding* (or *verstehen*...).'[19] He glosses this in terms of the familiar argument that if what is involved in making sense of a person amounts to appreciating how certain beliefs, judgements or actions might be rational, given the other things the subject believes, desires, and so forth, then principles of rationality must be in play in the very identification of an item as one with content. But Papineau suggests that 'the whole idea of *verstehen* as a distinctive mode of understanding loses its appeal once we reject the idea of incorrigible private access to mental 'givens'.[20]

Now it would indeed be ironic if it turned out that any part of the special understanding that is available through the identification of items within what Sellars calls the 'space of reasons' depends upon accepting the very myth that Sellars deserves so much of the credit for repudiating. In any case, it is difficult to see why the special mode of understanding that is generated by attributing reasons *should* depend upon accepting some version of the mych of the given.

Sellars's attack on the myth of the given is an attack on the empiricist's appeal to various items in order to justify, logically ground, or provide a foundation for claims to empirical knowledge. These items include 'immediate experiences', 'self-authenticating episodes', 'knowings in presence', objects of 'direct acquaintance' or other 'private' episodes, and indeed, 'immediate observables' within the public domain or even facts, if the capacity to know these facts is supposed to figure in an explanation of the the acquisition of ordinary empirical concepts. Sellars famously allows for the existence of inner episodes, immediate experiences, or appearings, as long as it is understood that these cannot be the raw data for the application of concepts; according to him, we can only make sense of *looks-talk* as parasitic on and thus presupposing (in the order of analysis) *sees-talk*

19. *Reality and Representation* (Oxford, Basil Blackwell, 1987); 119.
20. ibid.; 121.

or awareness of ordinary objects. To the extent that there are non-verbal thoughts, these cannot be what (in the order of analysis) accounts for language: the intentionality of thoughts, is, according to Sellars, inherited from and hence parasitic upon the semantical categories pertaining to overt verbal performances.

The sense-data theorist is an example of someone who makes the category mistake of supposing that the having of a sensation, which is not supposed to tax conceptual resources and is available to brutes and language-users alike, is magically supposed to carry with it the inferential structure (or 'articulation') which would enable it to play a justificatory role (i.e., be 'usable in formulating premises and conclusions of inferences assessable as correct or incorrect'.[21]) The mistake, or fallacy, or myth is to suppose 'that the space of reasons, the space of justifications or warrants, extends more widely than the conceptual sphere'[22] or that 'what is Given can serve as justification, without its being given requiring the exercise of conceptual capacities'[23] or that anything 'can count as a reason for a belief except another belief'.[24]

It is true that the kind of understanding that is brought about by identifying an item as something that falls within the space of reason is very different from the kind of understanding that is brought about by identifying an item as something that falls within a causal nexus; presumably, it is this fundamental difference that accounts for the suspicion that there is no recipe (via analysis or reduction) for identifying an item identified within one space as an item within the other. Papineau is right in supposing that some of what might account for the difference will take into account the point of view of the person whose thoughts and actions are under scrutiny. But a point of view in this sense does not introduce particular episodes of consciousness that are supposed to be independent of our conceptual resources and

21. Brandom, in the commentary in Sellars's *Empiricism and the Philosophy of Mind*, op. cit.; 138n.

22. McDowell, *Mind and World*; 7, quoted in Brandom's commentary, op.cit.; 122n.

23. Brandom's commentary, op. cit.; 122.

24. Davidson, 'A Coherence Theory of Truth and Knowledge' , reprinted in *Truth and Interpretation*—Perspectives on the Philosophy of Donald Davidson, E. Lepore, ed. (Oxford, Basil Blackwell, 1986); 310. This is quoted and discussed in McDowell, *Mind and World*; see especially 14ff and 137ff, and in Brandom's commentary, op.cit.; 122.

provide the sort of foundation sought by empiricists for claims to knowledge.

There is an obvious sense in which, when we ascribe a belief to someone as part of a way of rationalizing *her* actions and the rest of *her* mental life, we are bound to be considering things from *her* point of view. But there is a different way of coming to see the importance of the individual's point of view which I would like to consider; one I think Sellars is getting at when he suggests (in the passage cited in Section I) that the logical space of reasons concerns 'justifying and being able to justify what one says'. I have in mind the distinction that we often need to make when we are considering whether to regard someone as a fully-fledged participant in some practice or not. Sometimes, for example, we need to be able to distinguish between an individual who is able to make a move in a game, accidentally, as it were, and an individual who makes a move in a game because it is a move in the game.[25] The distinction is familiar enough when it comes to deciding whom to accept and whom to reject as players in the 'game of giving and asking for reasons'. We tend to reject without difficulty inanimate objects, forms of plant life, simple biological creatures and simple mechanical devices, even when, on occasion, they behave in ways that accord with rational norms. But difficulties arise with more complicated biological creatures likes cats, dogs and young children, and with the possibility of artificial intelligence.

Papineau, of course, is right to suggest that it makes sense to attribute conceptual awareness to animals and to children, and that it makes sense to ascribe beliefs and desires to them, since doing so enables us to understand their behaviour by showing why it would be rational for them to behave in that way. But there are obvious limits to their rationality. Crucially missing in the case of animals and small children, but present in paradigm cases of agency, is the agent's own explanation of or 'take' on her reasons for acting. Now the importance of the agent's self-conception has long been discounted—both by those (e.g., following Freud) who are attracted to the idea that a person's beliefs, desires, etc., might be unconscious and therefore not

25. See Sellars's 'Some Reflections on Language-Games', reprinted in *Science, Perception, and Reality* (London, Routledge & Kegan Paul, 1963), and my 'Playing the Rule-following Game' (ms).

accessible for reporting purposes and by those philosophers of language and mind (e.g., those following Chomsky and Fodor) who invoke a notion of tacitly held beliefs or quasi-intentional informational states which are also not accessible for reporting purposes. Those who discount the paradigm cases in which an agent or a speaker is able to cite the reasons for which she acted, or explain the meaning of her words, tend to do so by pointing out that agents and speakers are often wrong about their reasons and meanings and that often a person's sincere explanation of her actions or words needs to be rejected.

It is certainly true that a person's authority about her reasons or meanings does not provide a guarantee that what she says will rationalize her actions or constitute the meaning of her words. Nonetheless, what she says about these things cannot be ignored, even when she is wrong, since what she says is crucial to understanding her. It is not simply that we would be unable to diagnose a person as self-deceived unless we had some access to her own (mistaken) self-understanding. But insofar as an agent's conception of herself sets its own rational constraints on her future actions (and reinterpretations of herself) we would also lose a whole dimension of what is involved when we construe an individual as a participant in the 'game of giving and asking for reasons.'

It will be worth elaborating these ideas. Consider first cases of self-deception where it should be noted that there are at least two strands of behavioural patterns manifested. One is the pattern that belies the agent's self-conception and uncovers her ignorance about her own mind. The other is the pattern—often of denials, of protestations, of avoidance—that is a straightforwardly rationalizable outcome of this self-conception. Constant's Adolphe, who is self-deceived about the obstacles to his worldy success, is not merely wrong to blame his relationship with Ellénore. His false conception about their life together feeds into a whole pattern of behaviour leading to a tragedy that is itself only rendered comprehensible by this conception.

The importance of a person's self-conception is reflected in other ways besides in cases of systematic self-deception. When I want to hear that my beloved loves me, it is not merely because I like the sound of the words.[26] It is because his own conception

26. This example was suggested by a point made by Richard Moran in his reply to Bar-On and Long at the American Philosophical Association Pacific Division Meeting in Los Angeles, Spring, 1998.

of his love for me will account for a much richer pattern of behaviour than his unavowed love for me would. Imagine the difference between a relationship in which the beloved's love forms a part of his self-understanding which is expressed comfortably not only in his plans and actions but in his thoughts about these plans—and a relationship in which the beloved's love for the other—though it finds expression in certain ways—does not form a part of his own 'narrative'.

Most famously, this distinction plays a fundamental role when we attempt to evaluate the moral worth of an action. Kant probably went too far in suggesting that actions that accord with duty but which are not performed from duty have no moral worth.[27] But there is something right about the idea that the capacity to justify or to provide the right sort of reasons for one's action is an integral part of what it is to see someone as a moral agent. There is some plausibility, as well, in the idea that the capacity to justify or to provide reasons for one's actions or for one judgements is an integral part of what it is to see someone as an agent or as someone who has a claim to knowledge.[28]

It is this capacity to justify that is missing when the individuals whose actions are being evaluated are animals and young children. There is nothing strange about accepting them as

27. Kant, *Foundations of the Metaphysics of Morals*, trans. Lewis White Beck (Indianapolis, Bobbs-Merrill, 1959); First Section.

28. This idea is integral to Sellars's conception of what counts as knowledge. He concedes that '...the correctness of a report [say, 'This is green'] does not have to be construed as the rightness of an *action*' and that 'it can be correct as being an instance of a general mode of behaviour which, in a given linguistic community, it is reasonable to sanction and support'. Nonetheless, there is a second strand to the authority of a sentence token 'This is green': it must not only have authority in the sense that we can infer from the utterance the existence of a green object, which we could do, by the way, for a well-trained pigeon. What distinguishes a fully paid up member in the practice (someone we can credit with knowledge) is that 'this authority must in *some sense* be recognized by the person whose report it is' (*Empiricism and the Philosophy of Mind*; 74). And, as Sellars acknowledges,

> this is a steep hurdle indeed. For if the authority of the report 'This is green' lies in the fact that the existence of green items appropriately related to the perceiver can be inferred from the occurrence of such reports, it follows that only a person who is able to draw this inference, and therefore who has not only the concept *green*, but also the concept of uttering 'This is green'—indeed, the concept of certain conditions of perception, those which would correctly be called 'standard conditions'—could be in a position to token 'This is green' in recognition of its authority. (ibid.; 75).

It follows from this that one could not have observational knowledge of any fact unless one knew many *other* things as well. In Section VI, I suggest that the same sort of considerations apply to rational agency.

fledgling members in our reason-giving (and epistemic practices) if this merely involves identifying their behaviour as conforming to a rational pattern—even, in certain cases, one which comes about as a result of training—but the sense in which they are not full members, and might never become them, must not be forgotten. They are simply not responsible to the norms of rationality even if, in some sense, they can be seen as responsive to them.[29]

VI

Normative phenomena cannot be converted into non-normative phenomena without the loss of the special explanatory apparatus that the appeal to norms makes possible. I have suggested that this will involve a dimension of assessment which allows that a mistaken move can still be a move within the practice and that an individual who makes the mistake can still be considered a participant in the practice. I have also suggested that it will involve (not unconnectedly) the individual's own point of view or self-conception. Science, conceived as the realm of law, makes no room for 'internal error' in the same sense, nor for subjective points of view. The explanatory aims of science are different.

Even if one were to try to give some modified scientific account of this notion of 'self-conception' that seems such an important strand in conceiving an individual as a rational agent, how could this possibly be achieved? The normal way of deciding whether someone is acting in accordance with the rules of a game merely by accident, because of training, or because she is *following* those rules, is to decide whether it makes sense to see her as having a cognitive grasp of the rules. The usual way of deciding this is to determine whether she is capable of appealing to them in an attempt to justify or to explain what she has done. Obviously, however, if the capacity to see oneself as acting for reasons is part of what it is to be a rational agent (or if the capacity to justify one's beliefs is part of what it is to be a full member of our interpretive practices) then no non-circular account or explanation of what it is to act rationally, have reasons, or beliefs can

29. See Michael Root's discussion in 'Davidson and Social Science' in *Truth and Interpretation—Perspectives on the Philosophy of Donald Davidson*, E. Lepore, ed.; 272–304.

be given. To paraphrase Sellars,

> ...while the process of becoming a rational agent may—indeed does—involve a long history of acquiring *piecemeal* habits of response to behavioural patterns, there is an important sense in which one has *not yet committed* an action unless one is already a rational agent.[30]

This line of thinking, I believe, gives some substance to the refusal to offer constitutive accounts of normative phenomena, or to the refusal to 'locate "bedrock" lower than it is'.[31] Papineau, along with many others, is dismissive of the those who hold that 'the norms of judgement are primitive and not to be further explained.'[32] But if I am right, the refusal to offer such an explanation is justified by recognizing the impossibility of giving an account of something that is at the same time presupposed in our very method of accounting for it.

I would like to end by lodging a complaint against the use of the term 'non-naturalist' as a description of those who take epistemic or conceptual episodes to be irreducible to particulars that operate within the realm of law. Note that Sellars's dispute with the empiricists was about the incommensurability of items *within different areas of discourse*: the extent to which he believed that 'science is the measure of all things',[33] ought to be consistent with his mockery of the idea that there is a way to get beyond all discourse.[34] Note too that Sellars himself has been described as

30. This is a paraphrase of:

> ...while the process of acquiring the concept green may—indeed does—involve a long history of acquiring *piecemeal* habits of response to various objects in various circumstances, there is an important sense in which one has no concept pertaining to the observable properties of physical objects in Space and Time unless one has them all—and indeed, as we shall see, a great deal more besides (Sellars, *Empiricism and the Philosophy of Mind*; 45).

31. McDowell, 'Wittgenstein on Following a Rule', reprinted as essay 11 in *Mind, Value and Reality* (Cambridge, MA, Harvard University Press, 1998); 252.

32. Papineau, p. 28.

33. Sellars, *Empiricism and the Philosophy of Mind*; 83.

34. 'Does the reader not recognize Jones as Man himself in the middle of his journey from the grunts and groans of the cave to the subtle and polydimensional discourse of the drawing room, the laboratory, and the study, the language of Henry and William James, of Einstein and of the philosophers who, in their efforts to break out of discourse to an *arché* beyond discourse, have provided the most curious dimension of all.' (*Empiricism and the Philosophy of Mind*; 117).

explaining 'why we can be naturalists without being behaviourists, why we can accept Wittgenstein's doubts about what Sellars calls "self-authenticating non-verbal episodes" without sharing Ryle's doubts about the existence of such mental entities as thoughts and sense-impressions'[35] and as giving us a way to understand 'mind as gradually entering the universe by and through the gradual development of language, as part of a naturalistically explicable evolutionary process, rather than seeing language as the outward manifestation of something inward and mysterious which humans have and animals lack.'[36]

The philosophers who follow Sellars in supposing that normative phenomena are basic are 'naturalists' fighting a battle against scienticism. One familiar manifestation of scientism is the idea that in order to render the phenomena under investigation understandable or explicable, we have to be able to take a point of view that prescinds as much as possible from human sensibilities and we have to reconstrue areas of discourse in which the application of their proprietary concepts (evaluative, conative, and secondary quality concepts, in particular) seem to be constitutively tied to human sensibilities. It is becoming recognized, however, that 'evaluability' or 'normativity' extends wider than is usually supposed,[37] so it is no longer clear what sense there is in programmes designed to reconstrue or 'legitimatize' this element.

McDowell has long urged that not only *realists* of a certain breed, but also certain *irrealists* make a mistake in wishing to 'legitimatize' or reconstrue domains of discourse by relating them to areas less immune to human sensibility.[38] He might be right that the conceptual is unbounded,[39] that the escape from human sensibilities is a misconceived project in any case, and that there cannot be a project of giving an account of normative discourse if that means doing it from a non-normative point of view. Conceding all of this, it seems to me that there are residual differences between areas of discourse worth exploring that might give some

35. Rorty, in the Introduction to Sellars's *Empiricism and the Philosophy of Mind*; 6.

36. ibid.; p. 7.

37. See, for example, Brandom, *Making It Explicit*.

38. See, for example, 'Non-Cognitivism and Rule-Following', reprinted as essay 10 in *Mind, Value and Reality* (Cambridge, MA, Harvard University Press, 1998).

39. McDowell, *Mind and World*.

substance (back) to the idea that in certain domains, individual human judgement plays more or less of a role than it does it others.[40] In the domains that involve 'giving and asking for reasons', as my remarks above indicate, I suspect the individual's contribution plays a very large role.[41]

40. Cf. Crispin Wright, *Realism, Meaning and Truth*, Second edition (Oxford, Blackwell, 1993).

41. I discuss this further in 'A Constructivist Picture of Self-knowledge', *Philosophy*, vol. 71, 405–422. Thanks to John Flower, Richard Moran, David Wiggins and Crispin Wright for helpful comments.