This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

CONSTRUCTING COMMITMENT.
BRANDOM'S PRAGMATIST TAKE ON RULE-FOLLOWING

MATTHIAS KIESSELBACH

Abstract:
This paper defends Robert Brandom's attempt to explain normative statuses in terms of practical attitudes. According to a standard criticism, the strategy faces a dilemma: if practical attitudes and their interactions are specified in purely non-normative terms, then they underdetermine normative statuses; but if normative terms are employed within the account, it becomes viciously circular. This essay argues that allowing normative terms into the account need not yield vicious circularity, and that there is hence no dilemma. On the account proposed in this paper, practical attitudes, which are non-normatively specifiable responsive dispositions, can be seen to contribute to their own appropriateness or inappropriateness, if two theses are conceded: firstly, multiple agents' practical attitudes interact with one another in converging feedback loops; secondly, normative claims, including those uttered by us theorists, serve to exhibit their respective authors' practical attitudes and are hence implicated in the feedback loops. The circularity associated with this model does not attach to the explanatory strategy, but to the pragmatic system of interacting attitudes of which we theorists partake. It turns out that normative vocabulary, including the term "rule" itself, serves to enable the calibration of our responsive dispositions against those of others, thereby enabling smooth communication. Once this is understood, rule following ceases to be a problem.

1. Introduction

When Ludwig Wittgenstein asked, in his *Philosophical Investigations,* how an agent's performance can ever count as *obeying* or *violating* a rule,[1] he put new handles on an old problem. The problem is how to understand normative pronouncements, i.e. attributions of appropriateness or inappropriateness, and the new handles are essentially two aspects of a broadly pragmatist methodology. Firstly, Wittgenstein reminds us that the spring of our interest in normativity is an interpretive concern: how do we determine that an agent attempts, but fails, to do something, rather than doing something different from what we initially presumed? Secondly, Wittgenstein suggests that the question what this distinction comes to is a question about the pragmatic point of *making* the distinction in our dealings with other agents, first and foremost in language. Beyond this, however, Wittgenstein does not tell us much – at least not much by way of theory. The rule-following remarks in *PI* are an exercise in philosophical minimalism *par excellence:* their point is to remind us of how we ordinarily use normative vocabulary – no less than that, but no more either.

---

[1] See Wittgenstein (2001 [PI]: roughly §§138–242); see also Wittgenstein (1983 [RFM]: part IV). In this essay, Wittgenstein's works are marked as follows: PI: *Philosophical Investigations;* RFM: *Remarks on the Foundations of Mathematics;* OC: *On Certainty*; RC: *Remarks on Colour.*

This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

This essay will examine and defend a recent attempt at demystifying our capacity to institute and follow norms which takes off from *PI*'s rule-following remarks, but which aims to go beyond Wittgenstein's philosophical minimalism. For the last few years, Robert Brandom has argued that an agent's normative status – i.e. the property of being bound to a norm, so that some of her performances count as correct, and others count as incorrect – can be explained in terms of the practical attitudes that the agent and her peers take towards one another. This idea is standardly criticised as facing a dilemma: either the practical attitudes and their interactions are specified in purely non-normative terms, or normative terms are allowed into the account. In the first case, the approach falls short of yielding determinate normative statuses because there are infinitely many norms consistent with the given data; while in the second case, the account plunges into vicious circularity.

In this paper, I aim to show that allowing normative terms into the account of the emergence of normative statuses from practical attitudes need not result in vicious circularity, and that Brandom's own account can indeed be made to work. If we accept, firstly, that different agents' practical attitudes (of the relevant sort) interact with one another in converging feedback loops, and secondly, that statements about normative statuses – e.g. about the appropriateness or inappropriateness of particular practical attitudes – serve to exhibit their authors' own practical attitudes and are hence implicated in the feedback loops, then the circularity turns from vicious to virtuous. It no longer appears to attach to our attempt at explanation, but can be recognised as a structural feature of the system of interacting practical attitudes of which we theorists partake. It turns out that normative vocabulary, including the term "rule" and its cognates, serves to enable the *calibration* of our responsive attitudes against those of other agents, thereby making possible the smooth running of our complex schemes of social co-operation, up to and including language.

This essay proceeds in four steps. The following section reconstructs Brandom's view of the problem of norms. The third section introduces his attempt at solving the problem, as well as the worry that the attempted solution is viciously circular. The fourth – and main – section contains the defence of Brandom's account of norms against the charge of vicious circularity. The final section shows that the defence amounts to an interpretation of normative talk as a calibration device, and sketches out the theoretical fruitfulness of this idea.

This is a penultimate draft – please cite from the version of the article which appeared in *Philosophical Investigations* 35.2, 2012, pp. 101-126

2. The problem of locating normativity

For Robert Brandom, much of the current interest in norms was sparked by Immanuel Kant.[2] According to Brandom, Kant was the first philosopher to build a theory of intentional content on the concept of norms, and as a consequence, he was also the first philosopher to become seriously vulnerable to the problems surrounding it. Let us follow Brandom on a brief detour to Kant in order to introduce the problem of norms.

Kant's major achievement, in Brandom's story, was to get us to think about evolved intentional beings not as manifesting a particular kind of *substance* – "mental" substance, as scholastic thinking had it –, but simply as subject to judgements of correctness and incorrectness. The core of this approach lies in a novel understanding of the content of judgements and intentions. In both cases, Kant rejects the traditional analysis, which centrally involves a mental representation of a state of affairs and associates it with a particular non-cognitive force. Instead, he interprets judging and intending as taking up *commitments* to future performances, *thereby* laying the intentional agent open to confirmation or rejection. Talk of mental content, on Kant's new scheme, is no more and no less than a way of referring to these commitments.

Taking up a commitment, for Kant, amounts to binding oneself to a law, namely to a law specifying *what* future conduct one is to show in the light of one's (overt or inner) judgements or intentions. With respect to one's judgements, it specifies which further judgements – usually (but not always) taking the form of assertions – are to be made (accepted) and which rejected; with respect to one's intentions, it specifies which actions are to be carried out and which omitted. To characterise this kind of law as self-given is to say that we can only attribute to an agent a failure to follow the law if the agent can, at least in principle, be rationally convinced of the failure herself. If the allegedly mistaken agent persistently rejects the allegation of failure even after careful exploration of potential sources of distortion in her or our data, and if we also have no independent causal, i.e. non-rational, explanation of her failure, then it is generally more plausible that we have misinterpreted the agent. This would mean that the agent was following a different law – gave herself a different maxim, in Kant's words – and we interpreters are charged with finding out what it was, i.e. *what* judgement she made, or *what* her intention was. Here, we can see why Kant associates the laws at the root of intentional content with *autonomy*: they only bind when self-imposed.

---

[2] Brandom's account of Kant's role in the history of philosophy can be found in Brandom (1994: 7ff.; 2001: 602ff.; 2002: 212ff. and 2009: 32ff., among others). Note that Brandom is not overly interested in exegetical issues and often omits references. The main textual evidence of the normative character of intentionality in Kant are passages in which judgement is analysed as involving synthesis, where synthesis is a norm-governed activity, see Kant (1998 [CPR]: B 130ff.), but also see note 4. In spite of the controversial nature of Brandom's reading of Kant, I will sometimes just write "Kant" instead of "Brandom's Kant".

Unfortunately, for all its innovativeness, Kant's idea brings with it a philosophical burden which has been troubling philosophers as diverse as Hegel and Wittgenstein.[3] The trouble starts with Kant's modelling of the relevant concept of law on *positive law*. For Kant, subjecting oneself to a law and thus to judgements of correctness and incorrectness crucially involves the representation of an *explicit rule*. The problem with this idea, apart from the obvious difficulty of naming the medium in which the rule is to be explicitly represented, is that it runs into an infinite regress: since the interpretation of an explicit rule is itself subject to correctness and incorrectness, it would *(ex hypothesi)* rest on another, higher-level, rule which, in turn, rests on a yet higher-level rule, and so on.[4] For Brandom, the upshot of the infinite regress is that there must be a way of understanding commitments to norms, and hence correctness and incorrectness, as being implicit in practice. In Brandom's jargon, this amounts to a rejection of the doctrine of "regulism".

However, if we agree that commitments must be understood as implicit in practice, then there is a strong temptation to think that the correctness of a particular performance according to a norm just consists in its fitting into a *regular* stretch of behaviour. At this point, a second, related, problem – Brandom dubs it the "gerrymandering problem" – begins to threaten us: a finite stretch of behaviour can be interpreted as instantiating an infinite number of norms, no matter how simple or complex or intuitively visible its regularity is.[5] Regularity in finite sets of data – and finite sets of data are all we can ask for without begging the question – is thus no better at picking out the acts which are in accord with a norm than explicit rules are. What this means is that the doctrine of "regularism", to use Brandom's jargon once again, must be false, too.

The challenge which Brandom sees philosophy facing, then, is to show how we can understand commitment, i.e. the liability to judgements of correctness and incorrectness,

---

[3] See Brandom (1994: 18ff.; 2002: 217ff. and 2009: 66ff.)

[4] In a footnote (Brandom, 1994: 657 note 31), Brandom actually admits that Kant is not only aware of this problem, but even registers an explicit warning (in Kant, 1998 [CPR]: A132 / B171). What remains of Brandom's charge, it seems, is that Kant "makes very little" of his own warning.

[5] See Brandom (1994: 26ff.) and, again, Wittgenstein (2001 [PI]: §201).

This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

in a way which locates this status in ongoing practice, and which at the same time does not deflate it into mere regularity. If we accept the theory of content Brandom attributes to Kant, then the question's import is obvious. But even if we don't, the question how norms are instituted and followed is surely a formidable problem in its own right.

3. Norms as instituted by attitudes

Brandom's strategy to meet the challenge, in his opus magnum *Making It Explicit (MIE)* and beyond, rests on the idea that normative statuses can be explained in terms of practical attitudes. From the first couple of chapters of *MIE*, it is clear that Brandom means by "practical attitudes" non-normatively specifiable responsive dispositions[6] characterising the conduct of individual agents, or (to avoid all intentional vocabulary at this stage:) the behaviour of individual world-manipulating systems. His basic idea is that in groups of individual systems exhibiting such practical attitudes vis-à-vis each other, these practical attitudes can, under appropriate circumstances, give rise to normative statuses.[7] They can make make it the case, or contribute to making it the case, that some such practical attitudes can appropriately be said to be *appropriate* and others *inappropriate.*

Brandom begins his account with a series of stories of simple practices in which practical attitudes can be seen at work.[8] While it remains unclear how literally Brandom means his stories to be taken, it *is* clear that he means to convey with them a sketch of different developmental stages. At first, the sketched agents just react to each other by inflicting harm or administering pleasure, thereby reinforcing certain lines of conduct in their peers. Then the sanctioning manoeuvres become increasingly mediated and symbolic, and we begin to interpret them as manoeuvres of recognition of other manoeuvres *as* appropriate or inappropriate. Finally, the exchanges take the form of loops of symbolic recognition, and the symbolic sanctioning manoeuvres become interpretable as actually *worthy* of recognition as appropriate or inappropriate.

---

[6] These responsive dispositions – Brandom dubs them "reliable differential responsive dispositions" (or "RDRDs" in short) – can be thought of as given by algorithms. Brandom does not specify any limits to their allowed complexity or implementation requirements. He accepts, that is, that practical attitudes can be highly complex strategies of interaction with the world, including other agents/systems.

[7] "The theory developed in this work can be thought of as … offer[ing] an answer to the question, What features must one's interpretation of a community exhibit in order properly to be said to be an interpretation of them [sic] as engaging in practices sufficient to confer genuinely propositional content on the performances, statuses, attitudes, and expressions caught up in the practices? *If the practices attributed to the community by the theorist have the right structure, then according to that interpretation, the community members' practical attitudes institute normative statuses and confer intentional content on them*; according to the interpretation, the intentional contentfulness of their states and performances is the product of their own activity, not that of the theorist interpreting that activity." (Brandom, 1994: 61, my emphasis)

[8] See Brandom (1994: 33-46).

It is quite safe to say that Brandom's stories of the exchanges of increasingly sophisticated inter-agent reactions have convinced few serious readers, at least when taken by themselves.[9] Even if we grant Brandom the notion of recognition *as* appropriate (or inappropriate), it is hard not to suspect that Brandom still owes an account of how recognition *as* appropriate leads to *actual* appropriateness, i.e. to *appropriate* recognition as appropriate. Or to say it with Wittgenstein: it is hard not to suspect that Brandom still owes an account of the distinction between a situation in which an agent just *takes* her sanctioning manoeuvre to be appropriate and one in which it *is* appropriate.[10] And if the impulse is to answer this question by recourse to another agent's sanctioning manoeuvres, it is hard not to see this as question-begging: wouldn't we have to ask the same question again with respect to *that* agent's sanctioning manoeuvres?

Brandom is aware of these questions, but remains untroubled by them. His reason, it turns out, is that he does not mean his stories of primitive sanctioning behaviour to stand on their own feet, but to be accompanied by a second idea. This second idea, surfacing in a number of places throughout *MIE*, is that normative statuses are essentially attributed from *within* a normative perspective.

> The work done by talk of deontic [normative] statuses cannot be done by talk of deontic [practical] attitudes actually adopted … nor by regularities exhibited by such adoption… Talk of deontic [normative] statuses can in general be traded in only for talk of proprieties governing adoption and alteration of deontic [practical] attitudes – proprieties implicit in social score keeping practices.[11]

And again:

---

[9] A particularly dismissive reaction is Hattiangadi (2003).

[10] This difference figures prominently in Wittgenstein (2001 [PI]: §202) and is recognised by Brandom as central to any acceptable account of normativity.

[11] Brandom (1994: 626, see also xiii [Preface] and 58ff)

This is a penultimate draft – please cite from the version of the article which appeared in *Philosophical Investigations* 35.2, 2012, pp. 101-126

> We are always already inside the game of giving and asking for reasons. We inhabit a normative space, and it is from within those implicitly normative practices that we … assess proprieties of the application of concepts.[12]

In its circumventing the problem of "regularism", this move clears Brandom of the gerrymandering objection. This was a problem of underdetermination: no matter how complex a non-normative story about the reactive behaviour of primitive agents or systems is, the set of behavioural data provided in it will always be consistent with infinitely many norms, i.e. assignments of correctness and incorrectness. However, it seems that the cost of deflecting the gerrymandering objection is nothing less than vicious circularity.[13] If Brandom admits that the theorist – he – already interprets the practical attitudes visible in the examined agents or systems as appropriately or inappropriately undertaken, does he not simply import into the story what he wants, and claims, to get out of it? In the end, it seems, we are quite justified in shrugging his account off like we shrug off the statement "[T]hat is an authentic Vermeer just in case it is correctly attributed to Vermeer".[14]

Before we look, in the next section, at how Brandom's account can be defended, I ought to emphasise that although Brandom is aware of the worries, he does not address them until the very end of *MIE*. The inferentialist account of meaning filling *MIE's* central 500 pages is nothing less than *full* of (seemingly?) unexplained normative statuses.[15]

4. The theorist as bearer of attitudes

So far, it seems that we are dealing with a dilemma, and that Brandom is stuck on one of its horns. If the account of the emergence of normative statuses from practical attitudes is kept free of normative terms, then it will only underwrite regular patterns

---

[12] Brandom (1994: 648)

[13] See Rosen (1997). Note that not all commentators accept that Brandom has successfully fended off the regularism/gerrymandering charge. See Hattiangadi (2003).

[14] Rosen (1997: 167)

[15] It is also important to realise that if the charge of vicious circularity is legitimate with respect to *MIE*, it is certainly also applicable to Brandom's newer texts. In them, the basic explanatory strategy remains unchanged, the only relevant development being the acknowledgement of historicity, i.e. of the fact that a practical attitude can be a response to *past* attitudes (of other agents and, perhaps, of the responder's own former self) – see Brandom (2002 and 2009: esp. ch. 3). If *MIE*'s account of the emergence of normative statuses is viciously circular, then the later account of the emergence of normative statuses in explicitly transhistorical practices is surely also viciously circular.

This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

and fall short of yielding determinate normative statuses. If, on the other hand, normative attributions are allowed within the account, then it will be viciously circular. If this is true, then it seems that our best bet is to abandon the attempt to get normative statuses out of practical attitudes altogether. Brandom does not do this, and in this section I want to defend his stance. I want to show that allowing normative attributions into an account of normative statuses need not rob the practical attitudes of their explanatory potential, and that Brandom's own account can indeed be made to work.

Let us begin with what Brandom himself has to say. In various places throughout *MIE*, Brandom suggests that his account of normative statuses is not complete until *MIE*'s eighth chapter. This chapter discusses how different speakers' tokens of a singular term or pronoun (and indirectly also their relevant mental states) can come to be interpreted as having *the same representational content* (as representing the same object), *although* different speakers have different inferential dispositions surrounding the non-inferential dispositions within which their singular terms figure (although they face the world with different beliefs[16]). This is a problem, because Brandom's inferentialist – and thus anti-representationalist – account of language rejects the idea that the identity of individual objects, and hence the denotation of singular terms, is given to us prior to, or independently from, our inferentially articulated linguistic practices.

While for Brandom, the divergence of inferential dispositions is indeed an unavoidable aspect of our linguistic situation, he does not think that it rules out successful communication involving singular terms. Brandom's key move in this context is the insistence that sameness of representational content can be achieved through stipulation. The device by which this is done in natural language is the anaphora. With anaphorical constructions, we *take up* and *draw on* other speakers' references, *although* we cannot replicate all the inferential commitments of the speakers from whom we thus borrow. The success of our taking up others' references is guaranteed grammatically: this is simply what anaphors do.

This operation turns out to have philosophically important consequences. With the use of a referential placeholder whose content is anaphorically fixed and then grammatically guaranteed to remain fixed whoever uses it (within a given

---

[16] In these parentheses, I hint at non-pragmatist, non-inferentialist ways of understanding the problem.

This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

conversation), we get a way of distinguishing between what a particular speaker thinks about a particular object – basically, the inferences the speaker in question is disposed to draw from statements involving the referential placeholder –, and what is actually *the case* about that object – basically, the inferences *we* as commentators of the relevant claims are disposed to draw.[17] Whether we are right in what we say is another matter; what is important at this point is that we can distinguish, in language, between what someone thinks about an object and what is the case about the object.

As hinted above, the possibility of making this distinction is a crucial criterion of genuine normative statuses. However, Brandom's point is less than fully satisfactory in that it suggests that it is only meaningful to make the distinction between acknowledged and actual commitments with respect to discursive norms[18] – in fact, only in representational (singular term involving) contexts. Moreover, it is (once again) unclear just how this point is supposed to defeat the worry of vicious circularity, hence leaving commentators unsure whether it is "what [they]'d been gunning for."[19] Don't *we* – the subjects who interpret other speakers' statements about particular objects – have to be shown to be in the *right* when we say what these other speakers are committed to? Amidst all his talk of singular terms and anaphorical chains, Brandom leaves us wondering whether *MIE* contains the resources to dissolve the basic worry of vicious circularity. I now want to show that it does, but that these resources remain partly implicit in *MIE*, and that they operate at a level quite different from that of the representational content of singular terms.

We need to draw on *two* distinct ideas. For now, I will only outline them and sketch their place in Brandom's thought; their discussion will then take up the remainder of this essay. The first idea, which is the simpler and better understood of the two ideas, centrally informs Brandom's sketches of the exchanges of increasingly complex inter-agent sanctioning and recognition manoeuvres. The key thought, already touched on above, is that as brute sanctions are supplanted by symbolic responses with only indirect links with actual (non-symbolic) sanctions, we can expect an emergence of *loops* of symbolic recognition. A symbolic manoeuvre of recognition of another manoeuvre as correct or incorrect is liable to becoming the object of another symbolic recognition manoeuvre, which in turn is liable to becoming the object of yet another such symbolic recognition manoeuvre, and so on. And what is true of individual symbolic recognition manoeuvres is also true of the *dispositions* of carrying out such manoeuvres: in the end, we can speak of symbolic loops in which practical attitudes themselves refer to one another. Typically, these loops will display – indeed, given cognitive and epistemic constraints, require – a convergence on relatively stable patterns of symbols, and so I will speak of *converging feedback loops.*

---

[17] For a (much) fuller account, see Brandom (1994: chs. 7 and 8) and (for its extension to transhistorical practices, as sketched in note 15) Brandom (2002 and 2009: esp. ch. 3).

[18] This is Gideon Rosen's complaint; see Rosen (1997: 169).

[19] Rosen (1997: 168)

Before elaborating on this idea, let me introduce the second one. Although this idea is of the foremost systematic importance, it is usually overlooked or misunderstood, owing (at least in part) to the implicit and often rather metaphorical treatment in Brandom's writings. The relevant thesis, which can be put together from two different remarks of Brandom's, states that we must interpret normative statements, including the theorist's statements about the appropriateness of particular attitudes, as *themselves exhibiting their respective authors' practical attitudes* and hence as *implicated in the converging feedback loops* just introduced. Part of this thesis is covered by Brandom's remark (in the context of conceptual norms) that the representational content of terms is essentially *perspectival*,[20] meaning that it is from the perspective of participants *in* the language game, not from outside, that sameness or difference of representational content is attributed. The other part to the thesis is implicit in Brandom's idea that this attribution is always bound up with confirmation and criticism, which means that it is bound up with a display of the attributors' relevant practical attitudes. Both parts of the second thesis make their first appearance in Brandom's engagement with Daniel Dennett's celebrated idea that for a system to be an "intentional system" is for us to take the "intentional stance" towards it.[21] To be sure, Brandom differs from Dennett in his insistence that the theorist must not look at individual candidate systems (towards which *she* takes a particular stance) but at whole groups of interacting systems.[22] But he retains the fundamental insight that the theorist essentially contributes to the matter under investigation by taking up a particular stance towards it. From here, it is but a small step – requiring only that we equate stances with practical attitudes – to accepting that the theorist is a party to the system she is investigating.

---

[20] See, for example, Brandom (1994: 586ff.).

[21] "A particular thing is an intentional system only in relation to the strategies of someone who is trying to predict or explain its behaviour." (Dennett, 1979: 221) See Brandom (1994: 55ff.).

[22] Part of the reason for this difference is that Brandom is convinced that an attribution of correctness or incorrectness can, itself, be correct or incorrect in a more substantial way than Dennett allows. See Brandom (1994: 58ff.).

This is a penultimate draft – please cite from the version of the article which appeared in *Philosophical Investigations* 35.2, 2012, pp. 101-126

If we put together the idea that practical attitudes interact with one another in converging feedback loops and the idea that to make normative claims is to exhibit practical attitudes and thus to immerse oneself in the feedback loops, we can also say that our practical attitudes, when taken together, are involved in making themselves correct.[23] In the remainder of this essay, I want to elaborate what I mean by this and why the circularity inherent in the idea is not vicious but virtuous. I will start by running through two examples of the institution of normative statuses through practical attitudes, then draw out some interesting consequences, and finally respond to a few counter-arguments.

4.1. A first example: money and value

The first example centres on a relatively simple practice: the use of paper money.[24] It is designed, firstly, to introduce the proposal's basic two ideas, and secondly, to show that a pragmatist elucidation of norms does not rely on representational contexts.

It is quite obvious that in order to carry value, a Dollar note (say) need not be edible (or consumable in any other way). Also, it hardly relies on people with guns. All that a Dollar note needs in order to be valuable is to be widely valued, where "valuing a note" signifies the disposition to exchange it for goods or (other) notes which are designated as of equal value (or more for buyers and less for sellers). The qualification "widely" is meant to block the idea that my privately valuing a piece of paper is sufficient to confer value on it all by itself. What is needed, in addition to me, are other agents with the same disposition (or very similar ones). After all, I can only buy something for a Dollar if there are sellers, and I can only sell something for a Dollar if there are buyers.[25] But once I can do these things, I can speak of *value*.

It is crucial to a correct understanding of the concepts under investigation that the conditions for the meaningful talk of value are at the same time the conditions for the

---

[23] In his newer writings, Brandom credits Hegel with the view that multiple practical attitudes can be "jointly sufficient" for normative statuses to arise, see Brandom (2009: ch. 3, §1), see also Brandom (2002). It is clear that Brandom endorses this view. What is unfortunate is that in the newer "Hegelian" writings, Brandom does not make it clear that statements *about* normative statuses themselves exhibit attitudes (this, of course, is the second of the two theses which, I claim, stand behind *MIE*'s treatment of norms.)

[24] Of course, there is a tradition of taking money as a philosophical showcase, originating with David Lewis or, perhaps, with David Hume. However, it has traditionally focused on its conventionality – see Lewis (1969). I want to highlight different aspects.

[25] The same is true, of course, for these buyers and sellers.

This is a penultimate draft – please cite from the version of the article which appeared in *Philosophical Investigations* 35.2, 2012, pp. 101-126

meaningful talk of *misjudging* value. Being wrong about the value of a particular note or object – something that happens quite regularly[26] – is also made possible by the multitude of subjects involved in the business of exchanging, and hence valuing, notes and goods.

By now, it is probably clear that the idea I am after is that a paper note's *value* exemplifies a normative status, while *valuing the note* exemplifies a practical attitude, and that in an unproblematic way, the latter explains the former.[27] A little more precisely, it looks as though multiple agents' *attitudes* towards paper money are jointly sufficient for the latter to attain the *status* of being valuable. Its having the *status* of being valuable, in turn, is sufficient for it to be appropriate for each agent to take the *attitude* of valuing it. What we see, here, is a circle, but not a vicious circle, for it is not our explanation which manifests it, but the system within which valuing figures and gives rise to value.

Note how the proposal elucidates the status of being valuable, and thus of making certain attitudes appropriate and others inappropriate, without appealing to sanctions. Once the circular pragmatic system is up and running, sanctions are quite unnecessary.[28]

Of course, the circular pragmatic system rests on various layers of non-normative matter, and does so in a quite non-circular way. There are people who produce and consume goods and who have access to markets. They act on algorithms regulating their conduct on and off the markets. The algorithms, in turn, are implemented by immensely complex webs of interconnected neurons. And so on, and so forth. In an obvious way, the emergence of value relies on the stability of these layers of structured matter.

---

[26] If this seems strange to you, start by thinking about *blatant* mistakes of evaluation. Surely, it is wrong for me to say that a new Rolls Royce is worth (or "really" worth) only 1'000 $, even if *I* would (stupidly) not pay more to have it. (Stupidly, because I could go on to sell it for a multiple of 1'000 $.) Then continue to less blatant mistakes. Surely they exist. If I look for an expert to estimate the value of some real estate I want to buy or sell, I will aim for someone who is usually right about real estate value.

[27] It might be complained that it is unnatural to attribute normative statuses to inanimate objects; so far, we have only used the term to apply to people. I will stick to the use because it strikes me as a very modest extension, and because there is a principled way to turn the example into one where it is people, not pieces of paper, which bear normative statuses. (Think of a community of unusually honest people who, instead of carrying actual money, keep mental score of how much they "have", and make purely linguistic transactions.)

[28] In fact, in the money case, sanctions are hardly to be expected, because an agent's evaluative mistakes are usually *good* for other agents. This fact seems to me to be an important corrective to many philosophers' intuition (Brandom himself is no exception) that rule-following behaviour must start with sanctions. It seems to be better to think that the evolution of rule-following starts with the danger of loss.

This is a penultimate draft – please cite from the version of the article which appeared in *Philosophical Investigations* 35.2, 2012, pp. 101-126

However, to say this is not to claim that value can be reduced to any of it. In fact, we have reason to be wary regarding the latter claim – not just because the value of a single note depends on the *whole* (potentially vast) circular system, but also because talk of reduction is oblivious to the pragmatic role of attributions of value *within* the circular system.[29] Here, we come to the second thesis at work in Brandom's account of norms.

One way to make the point is this: an agent's claim that a Dollar note is (now) worth, say, one loaf of bread, occupies the same logical space as other agents' valuings of it and is thus best interpreted as an explication of her own valuing. Of course, the agent might be a theorist, and her claim might be a line in a philosophical paper of hers. But even so, if she is talking about an actual note in an actual marketplace she is observing, then if other agents consider the note worth *two* loaves, those agents would rush to make deals with the theorist – bringing it about, if the market is small enough, that the note *becomes* worth something like one loaf (or if it's big, that she becomes poor and eventually drops out). She would have no reason to ward them off with the remark that her claim was "meant in a purely descriptive way": if her statement about the note's value is truthful, and if she is rational, then she would be happy to make the deals.

In preliminary conclusion, if we recognise that multiple agents with particular dispositions to exchange paper notes for goods and other notes *interact* with one another to form systems of paper money circulation with relatively stable convergence points, and that it is only within such systems that an individual exchanging disposition can be proper or improper, and if we allow further that a statement *about* a disposition's propriety or impropriety – in other words, a value statement – serves to make its author's own exchanging dispositions explicit, thereby affecting the system as a whole and contributing towards its own propriety, then there is a sense in which the *value of a paper note* can indeed be explained in terms of *agents' dispositions of dealing with the note* (and with other notes of its kind). We start[30] with a world containing only exchanging dispositions, and end up with a world containing value.

---

[29] Wittgenstein was fond of making the point that rules rely on regularity without reducing to it. See, for example, Wittgenstein (2001 [PI]: §§207f). I will say more about the problems facing the attempt to derive normative facts from non-normative facts below. At this point, I am only concerned to introduce the main ideas.

[30] I should stress that my use of temporal terms is purely for didactic reasons. I am *not* talking about the genesis of paper money circulation, I am *only* talking about the structure of such a system once it is up and running. The same goes for the next example.

This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

4.2. A second example: language and meaning

Let us now jump to the context which has inspired philosophical thinking about norms more than any other: language and meaning. We will see that with respect to the exchange of linguistic signs, it is fruitful to speak of practical attitudes and normative statuses in the very same way as the one sketched above. My private disposition to use an expression in a particular way – inferentially, substitutionally or non-inferentially, to go with Brandom's theory of language – by itself hardly confers *meaning* on the expression. But if *my* disposition interacts with the relevant dispositions of *other* speakers in such a way that we can speak of an ongoing conversation, then it is quite appropriate for me to say that the expression carries a meaning which determines the correctness or incorrectness of my uses of it. Moreover, my statement *about* the meaning of the expression exhibits a linguistic disposition of mine, thereby defeasibly affecting – in fact, defeasibly settling – the meaning of the expression and hence contributing towards its own appropriateness.[31] We thus have another circular but converging pragmatic system of attitudes (this time: linguistic dispositions) and statuses (this time: meanings), and once again, the theorist's place is right within it.[32] This comes out most clearly when we look at communication problems.

Our communication by linguistic signs is a rather smoothly running practice. However, there is always a possibility of communication problems: every now and then, someone makes an utterance which we find surprising and slightly off the mark, perhaps even flat out wrong. For example, we are used to people talking about the "Sahara Dessert" or saying that "dolphins are majestic fish". We are even used to people complaining to their doctors about having "arthritis in the thighs". In order to distinguish cases in which an interpreted speaker just speaks weirdly – i.e. makes malapropisms – from cases in which the speaker makes factual errors – i.e. shows inconsistencies in her framework of inferential, substitutional and non-inferential linguistic dispositions – and to determine the precise locations and scopes of the malapropisms and errors, respectively, we have developed a range of examination

---

[31] The idea that meaning statements have special perlocutionary consequences which set them off from ordinary descriptive claims is also featured in Lance and O'Leary-Hawthorne (1997), who underline the inference-licensing and the censuring role of meaning statements (see 58ff., 64ff.). Unfortunately, they do not say how the *correctness* of meaning statements is determined.

[32] We can thus say: *meaning* (the verb) is to *meaning* (the noun) as *valuing* is to *value*.

This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

techniques.[33] One of these techniques[34] is the use of *meaning talk*. Consider a case where you have ordered water, but the waiter brings tea, and on your complaint that you actually ordered water, the waiter says (with a puzzled face): “Well, that’s what you’ve got.” In this kind of situation, it is quite natural for you to ask something like: “By ‘water’, do you perhaps mean any liquid which has at least 99% of H_2O in it?” And the waiter’s answer may well be: “Yes, that’s what ‘water’ means”. If this is indeed his reply, the case looks like a complex malapropism – in other words, a problem which can be remedied by re-interpretation alone.

What is interesting in this context is the direct analogy to what we have seen in the money example. By replying something like “But in a restaurant setting, ‘water’ means *clear* H_2O” you make it the case (albeit defeasibly) that the exhibited use of the relevant terms becomes their appropriate use – namely by inducing your interlocutor to use them in the way exhibited by your meaning statement, at least when dealing with you. Not only do you induce the waiter to bring you *water* when you order “water”, you also enable him to translate his own remarks into your ways of speaking. What he formerly (or normally) meant by “water”, he will, when talking to you, express (perhaps) as “water *or* tea” – given, of course, that he is interested in successful communication, and given that his weird category ever becomes salient in a conversation with you. In short: now, “water” *really* means *clear* H_2O – just as you said!

In order to deflect two anticipated complaints, let me enter two clarifications. Firstly, in my scenario, I made common parlance win out against weird parlance. However, I could just as well have had you reply something like: “Oh, if *that* is what you mean by ‘water’, then I would like to order cold bottled H_2O *without any added substances,* please.” Within the little language game between you and the waiter, something being “water” would then be compatible with it being “tea”, again showing that the practical

---

[33] There is actually a third possibility between malapropism and mistake: an utterance which initially seems strange can turn out to reflect a formerly unknown fact. Although this point strikes me as the main reason behind our need for the linguistic manoeuvres sketched in this section, I want to avoid overburdening this essay with the topic of linguistic evolution via increase in empirical knowledge. For some roots of this idea, see Wittgenstein (1972 [OC]: §§210f.); for a more extensive treatment, see Brandom (2008: lecture 6, §§3f.); see also Brandom (2002 and 2009: ch. 3).

[34] Another technique is the formulation of what Wittgenstein has famously called “grammatical propositions”. We make statements which are so obviously true that any denial on the part of another speaker points not towards differences in opinion, but merely towards differences in meaning. An example of a grammatical statement in the scenario sketched below would be the sentence: “(Let’s see, we do agree that) there is a difference between tea and water (right?)”. See Wittgenstein (2001 [PI]: §242); see also Wittgenstein (1978 [RC]: I:32), which confirms that grammatical statements are an everyday manoeuvre, not just a therapeutic philosopher’s trick.

This is a penultimate draft – please cite from the version of the article which appeared in *Philosophical Investigations* 35.2, 2012, pp. 101-126

attitude exhibited by the meaning statement – this time the waiter's – contributed to its own appropriateness. Secondly, while in small and transitory language games, people are substantially more free to adjust even to idiosyncratic understandings of their interlocutors, in larger games with more than two speakers, there are usually good reasons to refrain from giving in too much. Here, it is much more likely that a speaker's individual meaning statement fails in the sense that it is not validated by other speakers. Of course, my proposal is not meant to block this obvious point.

It is worth spelling out the second clarification in more detail, as this provides a welcome opportunity to discuss the question of when to use our power to make a disposition correct – by adapting to it, by adopting it – and when not to. As I said, in the rare cases of language games involving just two speakers, it is usually not overly costly for the speakers to adapt to their respective interlocutor's linguistic dispositions. Of course, such an adaptation takes some cognitive energy, and there may be a loss in linguistic elegance, but supposing that both speakers' original dispositions have the same expressive power, there are no further costs involved. The situation is different when there are more than just two speakers. Normally, each speaker must not only aim for harmony with one interlocutor's linguistic dispositions, but with multiple interlocutors' dispositions. In such a situation, a speaker's decision to adapt to a weird interlocutor's linguistic dispositions actually constitutes a decision to allow the language game to split up into separate language games – after all, the speaker is most likely unable to do anything about the fact that other speakers continue talking in the ordinary way. Needless to say, this comes with extra costs all speakers involved, and so it may be best simply to reject the weird speaker's linguistic disposition as wrong, and leave it at that.[35]

What we see here, I contend, is analogous to the case of money and value. If we allow that the appropriateness of a linguistic disposition is an aspect of its interaction with other speakers' dispositions within systems of linguistic exchange with relatively stable convergence points, and if we allow further that the point of labelling a disposition "appropriate" is to exhibit the disposition oneself, to the end that other

---

[35] Since the linguistic world described in this essay is idealised in that linguistic differences do not stem from different stages in empirical learning (see note 33), the adaptation to another speaker's disposition comes out as a rare exception. As Davidson (1986) rightly stresses, however, some degree of dispositional difference is ubiquitous and unavoidable, and linguistic competence thus cannot exist without some capacity to adapt to foreign dispositions – even in large language games.

This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

speakers adjust their linguistic conduct accordingly, then there is a sense in which the *appropriateness of a speaker's linguistic conduct* – and hence *the phenomenon of meaning* – can be explained in terms of the *linguistic dispositions* of the speaker and her peers. We start with a world containing only linguistic dispositions, and end up with a world containing meaning.

4.3. Three points about meaning

Before looking at a few worries regarding the circulatory account of norms just sketched, let us briefly consider three of its most salient features. I will discuss them in the context of language and meaning, although they could, with certain modifications, also be discussed in the context of money and value. The features are, firstly, the possibility of extending the account to cover third-person meaning statements; secondly, its counsel of caution regarding the thesis that semantic facts or statements can be derived from (or reduced to) non-semantic facts or statements; and thirdly, its allowance of claims to the effect that a whole linguistic community violates a meaning.

Let us start with the first feature right away. Above, I have only spoken of first-person meaning statements ("I / we mean A by B") and non-personal meaning statements ("B means A")[36] and treated them as synonymous (at least in the kinds of scenarios imagined). However, this should not be taken to imply that third-person meaning statements are somehow illegitimate. In fact, allowing these statements constitutes an interesting extension of the account.

Consider statements like the following: "By 'collateral damage', NATO generals mean dead civilians", "By 'arthritis', Tyler means arthritis or gout" and also "'Schnee' means snow in German." The relationship between these statements and non-personal meaning statements is quite close – however, their import is rather obviously limited to contexts involving the speakers referred to (NATO generals, Tyler, and the Germans, respectively). Just as before, the function of these statements is to prevent or redress communication problems, although the communication problems are now located in the mentioned third party contexts. Just as before, these statements explicate attitudes, and

---

[36] It should be obvious that what I am referring to are not strictly-speaking grammatical categories. Of course, "non-personal" meaning claims are, grammatically speaking, third person statements. What makes them "non-personal" is that they do not refer to *speakers*.

This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

are themselves involved in the business of validating these attitudes. Here, however, it is important to note two points.

Firstly, there are two different attitudes involved: that of the speaker making the third-person meaning statement, and that of the speaker or speakers referred to *in* the statement. The former is explicitly limited to contexts which involve the speakers referred to in the statement, while the latter is not explicitly limited. Secondly, while both attitudes rely on their being validated, the validation of the third party's attitude can be *one-directional*. The point is that it is logically possible to add to a third-person meaning statement the counsel not to *adopt* the way of speaking exhibited in it: "By 'arthritis', Tyler means arthritis or gout, but we really should not adopt this way of speaking". In cases of this sort, talk of validation may only refer to an audience's interpretation of the speakers' utterances, not its adoption of these speakers' ways. It should be obvious that this way of extending the circulatory account of meaning to third person statements brings us right into the domain of everyday intentional talk with its characteristically intensional[37] nature ("Tyler thinks that...", "The books says that..." etc.).

Here is the second feature. From what has been said, it should be clear that anybody who talks about meaning, including a theorist of meaning, must be located within the system giving rise to it. It is not just that we would hesitate to interpret a certain sequence of sounds as *speech* unless the person (or organism) emitting it took part, or could take part, in some appropriate sense, in the linguistic and non-linguistic practices making up our lives, crucially including the trading of goods and the bringing of cups of tea, to continue with the examples given above. This philosophical point is routinely stressed by Wittgensteinians. But what is of comparable importance, and only made explicit in the account just given, is the fact that by saying that A means B, a speaker indicates that *she* uses and interprets the terms accordingly – and thereby contributes to making it the case that *everyone* who uses the terms in this (substitutional, inferential or non-inferential) way is *justified* in doing so. Unless we are prepared to grant a sound-emitting organism this kind of authority over our linguistic conduct, we ought not speak of a *speaker* and hence not of her following, or knowing, or making explicit, a *meaning*.

---

[37] For more on the inten*s*ionality of inten*t*ionality, see Brandom (2001: 588ff.). Analogous statements in the field of value are claims such as "Peter values it highly" or also "Peter would never sell it for anything less than $100".

It is here that the circulatory account of meaning speaks against the reduction of semantic statements to non-semantic statements.

Note that even an ideally situated Martian observer with extreme information-handling capacities and with data of the locations of all particles on and around our planet (over a long time) need not be characterised as knowing the *meanings* of our expressions, or as being in the position to infer them. After all, it is most likely that the omniscient Martian's inhuman nature bars it from entering into the practices within which meaning talk has its practical home and within which we are prepared to accept others' authority over our linguistic conduct. Importantly, however, this is not to deny the Martian's capacity to *predict* utterances *of meaning statements* on the part of actual human speakers. It is also not to deny that it can quote these utterances, as it is not to deny that it can quote dictionary entries. Sure: in some restricted contexts, for example in game shows, "knowing the meaning" means nothing more than being able to blurt out such quotes. In this restricted sense, the Martian *can* know or infer the meanings of our terms. But if the pragmatist account of meaning talk given above is correct, and if the Martian does not somehow limit her cognitive powers and become more like us humans, then there is a sense of "knowing the meaning" on which only humans – namely, participants in those of our practices within which ordinary meaning talk figures – can claim such knowledge.

Finally, a third consequence of the circulatory account of meaning is the logical admissibility of statements to the effect that whole communities go wrong (by their own standards). If a whole imaginary community, through whatever quirk of fate, from one day to the next starts to treat "water" as covering *tea* in restaurant settings, there is room for us to say that all of its members go wrong by their own standards. We, after all, can imagine ourselves as community members, and as such we have every logical right to pronounce on the meanings of their terms. However, two points are immediately clear in this context – lessening the excitement somewhat, but strengthening the analogy with money. Firstly, by making our pronouncement, we show that not the whole community went wrong. We, after all, are still right. Secondly, another theorist could always come along and argue that instead of going wrong, the members of the community shifted from one norm (and thus from one concept) to another norm (and thus to another

This is a penultimate draft – please cite from the version of the article which appeared in *Philosophical Investigations* 35.2, 2012, pp. 101-126

concept). Unless the new way of speaking involves an inconsistency about which we can agree with the other theorist,[38] any debate about this is quite gratuitous.

4.4. Three worries and their rejection

I propose to end this section with a discussion of three worries regarding the defence of Brandom's strategy to explain normative statuses in terms of practical attitudes.[39] The first worry is that the proposal illegitimately blurs the line between knowledge and meaning, i.e. between theoretical statements or beliefs (about the world) and merely grammatical[40] commitments (which just settle the meanings of the expressions appearing in them). The second worry is that the proposal is inconsistent with what many philosophers have come to regard as one of the greatest achievements of recent philosophy of language, namely semantic externalism. The third – and perhaps main – worry is that the proposal fails to address the infinitary nature of meaning.

Let us start with the first worry. The counter-argument at its centre is based on the premise that the picture drawn here blurs the line between substantive and grammatical commitments. With each inferential, substitutional or non-inferential commitment, it is claimed, we can say whether the commitment counts as a piece of actual knowledge or as just settling the meanings of sentences or words. However, one can cast doubt on the dividing line between knowledge and meaning. At least for a Brandomian inferentialist, the line is not only blurred; the idea of such a dividing line actually rests on a mistaken view of language. According to Brandomian inferentialists, any sentence which counts as informative, i.e. contentful, in some scenarios can be employed, in other scenarios, as a grammatical statement. The status of a particular sentence depends on what *other* sentences are taken as its background, not on how it is constructed or on how it was originally integrated into the speaker's web of beliefs. Since I cannot, of course, argue for the holistic view of language in the confines of this essay, I must contend myself with registering that (at least) adherents of Brandomian inferentialism will not be

---

[38] I have in mind an inconsistency within the framework of inferential and non-inferential (or evaluative, in the case of money) dispositions of the speakers (traders) in question. Of course, any alleged inconsistency can *logically* be explained away using a more complex disposition; hence the insistence on prior consensus about what is and what is not an actual failure of consistency.

[39] As in the previous subsection, I will set the discussion in the context of language and meaning, although it could – with certain modifications – also be set in the context of money and value.

[40] Once again, I borrow from Wittgenstein's technical apparatus. See note 34.

This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

convinced by the counter-argument, and that a lot more needs to be said to settle the issue.

The second worry is that the account of meaning sketched here is inconsistent with Saul Kripke's and Hilary Putnam's widely celebrated theses about the externality of semantic content.[41] I want to claim that the circulatory account of meaning sketched in this paper is not only compatible with semantic externalism. In fact, it might open up a perspective on the matter from which even some of those who have, up to now, rejected the externalist position, can revise their view of it. The core of the externalist thesis is that the meanings of speakers' expressions are not fully controlled by the speakers using them, so that the latter can find out – *a posteriori* – what their terms mean, for example what their terms denote. A simple (and this-worldly) example is the term "gold" as used by Archimedes a long time ago. Externalists point out that when Archimedes' used the term, he referred to the very same range of things as *we* do when we use the term "gold", although Archimedes' did not know nearly as much as we do about how to test whether a particular shining object is or is not gold. And even most of us today, externalists continue, know very little about how to do that.

Up to this point, of course, there is no hint of an incompatibility between the account of meaning sketched here, and semantic externalism. The appearance of incompatibility enters the stage with the next step. When we ask what it is that determines the denotation of Archimedes' and our term "gold", we notice that what seem to matter are chemical facts – facts to which Archimedes had no access, and to which even most of us only have a highly indirect access (if any access at all): "gold" denotes objects made up of atoms with mass number 79 (perhaps with some permissible degree of impurity) – whether we know this or not, and whether we have the means to identify such mass numbers or not. In a way, it is the *world*, in the guise of our familiar objects' microstructure (or what else is their essential nature) which determines what our term "gold" denotes and hence what it means.

This morale of the story is hard to swallow for adherents of many varieties of meaning-as-use-theories; in particular, it seems to be incompatible with the circulatory account of meaning sketched above. The latter, after all, holds that meaning is the product of interactions of speakers, not of the "world". However, there is a way of reading the externalist story which is compatible with the circulatory account of meaning. The externalists' story need not be interpreted as pointing to the semantic power of the metal gold and the other substances making up the "world", but can be seen as pointing to a quite mundane aspect of our natural language.[42] The aspect to which externalists draw attention is that speakers of natural language can be interpreted as implicitly deferring, in matters of meaning, to intellectual authorities.

---

[41] The reading to be sketched in this paragraph is inspired by Ebbs (1997), who minimises the "causal determination of reference" story and maximises the "linguistic division of labour" story in the externalists' writings. Ebbs's story is consequently based more on Putnam (1975) and (particularly) on Burge (1979) than on Kripke (1972).

When a speaker of our natural language uses the word "gold", she implicitly grants that just *what* is thereby picked out is settled, in part, by current or future experts *in metallurgy*. Of course, in order for a metallurgist to know *what* she is supposed to pick out as constituting or belonging to the extension of some term of ours, there must be the implicit understanding of an original encounter with a particular natural kind and a chain of expressions for which co-reference is guaranteed. But while the former (the natural kind) is just a matter of the taxonomies used by the intellectual authorities to which we defer,[43] the latter (the chain of co-referring expressions) is just a grammatical aspect of natural language. There is, in other words, a way to read the intuitions elicited by externalist thought-experiments in which all determinants of meaning are firmly located *within* the reach of speakers – only not within the reach of *every* speaker. Indeed, sometimes a determinant of meaning lies only in the reach of the interpreter. If we put the externalists' insight in this way, then the circulatory account of meaning is no longer at odds with it. On the contrary: the externality of semantic content is captured remarkably well by an account featuring linguistic dispositions which validate themselves in concert in the way described in this section.[44]

Finally, there is the third worry that the proposal misses a key part of the problem in that it fails to address the *infinitary nature*[45] of norms. Perhaps this is the main worry. Wittgenstein's original presentation of the "rule-following" problem[46] involves a teacher who tries to get his pupil to continue a series which he begins as follows: "2, 4, 6, 8…".

---

[42] The word "our" here presumably includes Archimedes.

[43] Of course, these taxonomies are shaped and re-shaped in the course of complex interactions with the physical world, but this is an interaction that is innocent from the point of view of critics of semantic externalism.

[44] I ought to stress that the story sketched out here is not limited to natural kinds terms, but to all sorts of kind terms, and also to names.

[45] See Boghossian (1989: 509).

[46] See Wittgenstein (2001 [PI]: §185 among others).

This is a penultimate draft – please cite from the version of the article which appeared in *Philosophical Investigations* 35.2, 2012, pp. 101-126

The pupil continues in the way we would intuitively consider correct (“10, 12, 14…”), but on hitting the number 1000, starts to diverge, writing “1000, 1004, 1008…”, thus provoking the philosophical question how to understand and establish the incorrectness of his steps by the series as the teacher *meant* it. It might seem that our proposal does little to solve *this* puzzle. On the basis of this essay’s proposal, we would just say that *the teacher’s series* does not include the fragment “1000, 1004, 1008”, but the fragment “1000, 1002, 1004”. By saying this, we would display a practical attitude whose content, among other things, is that on encountering the fragment “1000, 1004, 1008”, we judge that it does *not* belong in *the teacher’s series*. Thereby, we contribute to it actually being *inappropriate* to consider “1000, 1004, 1008” part of “the teacher’s series” (in the way explained above – note that we are dealing here with an internal relation between two linguistic items). So far, so good. However, it is crucial to note the “among other things” formulation. When thinking about the question which other judgements we are disposed to make, we soon realise that our practical attitude is as infinitary as the mathematical series. Wasn’t this the core of the problem?

The answer is no. The problem was not to explain how we can have infinitary dispositions, the problem was to explain how we can say that particular performances are correct or incorrect. Infinitary dispositions, which we also find in machines such as clocks or computers – at least, if we allow that a disposition can be called “infinitary” although the machine manifesting it has a limited life-span and is less than perfectly error-proof, quite like us – are as unproblematic as our ability to count and keep on counting. The question to which Wittgenstein draws our attention, namely how people can make *mistakes* in counting – or more exactly: how we can describe people as purporting (or trying) to count but making mistakes – is adequately answered by this essay’s proposal. Correct counting, for example, does not involve the step from 28763542 to 28763544. How do I know? Well, partly because I am, like any computer, equipped to add 1 over and over again, remember the result, and check whether a particular step has been made. (I am also quite good at recognising patterns.[47]) But however I know it, and however I do it: if this essay’s proposal is correct, then this

---

[47] Note that human pattern-recognition capacities are actually far too complex to be technically reproducible at this point. While we can make computers count, our current technological limits are already manifest in areas like the recognition of facial expressions. The immense complexity of human capacities of pattern-recognition, the consequent limits of our attempts at making their inner workings explicit, and the resultant de facto impossibility of checking whether two speakers’ pattern-recognition dispositions are reliably congruous, have led some philosophers to express uneasiness about the ensuing precariousness of mutual understanding. See, e.g., Cavell (1969: 52).

This is a penultimate draft – please cite from the version of the article which appeared in *Philosophical Investigations* 35.2, 2012, pp. 101-126

attitude of mine is part of the system that gives rise to the actual incorrectness of the step from 28763542 to 28763544 given the ordinary understanding of the term “counting”.[48]

5. Normative talk as a calibration device

It thus appears that Brandom’s strategy to explain normative statuses in terms of practical attitudes can be made good, if we accept that practical attitudes essentially interact with one another in converging feedback loops, and if we accept that claims about normative statuses themselves exhibit practical attitudes and are hence implicated in the feedback loops.

With this defence of Brandom’s project against the charge of vicious circularity, however, we have changed the status of his account. At first, we innocently asked how we can be committed to a norm, noting the problem that the usual answers either plunge into an infinite regress or turn commitment into mere regularity. Initially, we tried to construct an answer in the form of a straightforward explanation, resting on facts about practical attitudes as the explanans. In noting, however, that claims about commitments to norms are best interpreted as themselves exhibiting the very practical attitudes that our explanatory attempt mentioned, we came to see that the original question was a bad question to ask. The philosophical worries behind it turned out to be best addressed by trying to answer the question what it is that we – individually and collectively – *do* when we make claims about commitments to norms. It is at this point that we transformed our account from one purporting to explain a phenomenon into one elucidating a region of discourse.

But our elucidation seems to be doing slightly more than Wittgenstein’s minimalism which is also, as we noted in the beginning, aimed at reminding us of how we ordinarily use our words. While Wittgenstein memorably warned against all pretensions of philosophical *theory*, i.e. of claims about language whose truth is not already obvious to speakers solely in virtue of their linguistic competence,[49] this is just what we seem to have done. Our story did not restrict itself to surveying our uses of words, but

---

[48] Surely, if the question is how *counting* came to be more salient than, say, *quounting* (which is like counting, except that it does allow the step from 28763542 to 28763544), then we should tell a story about customary practices and their natural selection. See Dennett (2006: esp. 8ff.).

[49] See Wittgenstein (2001 [PI]: § 128)

This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

comprised an account of the virtuously circular social mechanism to which our individual uses of language add up. Indeed, from our account of this mechanism, we can derive a plausible *functional* characterisation of normative talk: since the appropriateness of a particular disposition is the effect of its interaction with other dispositions, and since the affirmation of a disposition as appropriate (or its rejection as inappropriate) *itself* exhibits such a disposition and thereby influences the system as a whole, we can characterise normative talk as essentially *calibrational.* We use normative vocabulary to get the dispositions of others into line with ours, so that we can then employ them to meaningfully confront the world of things we trade (using money) or talk about (using language).[50] This conclusion, it seems, deserves to be called *theoretically interesting:* it is a conclusion that is less than fully obvious to speakers simply in virtue of their linguistic competence, and yet, in the light of the sketched account, plausible.

I think that the worries that Wittgensteinians might have about this style of philosophising can be assuaged by distinguishing between two commitments which usually stand behind minimalist programs in philosophy. We find them both in John McDowell, a self-avowed “quietist” himself, who has contributed much to the literature on rule-following. The first is his insistence that philosophers ought, when attempting to dissolve philosophical worries, to restrict themselves to “appealing to social interactions, described in a way which does not presuppose the material to be reconstructed”,[51] not to aspects *beyond* social practice. They ought not, that is, to erect standards of correctness of linguistic (and other practical) manoeuvres which are unknown to, and out of reach for, ordinary speakers (agents), or – in McDowell’s words – look at social interactions “from sideways-on.”[52] The second commitment is the principle that when faced with questions about aspects of our language games which are integral to our conception of ourselves, the appropriate response “is something like a shrug of the shoulders.”[53] For example, when faced with philosophical puzzlement

---

[50] Note that the account given in this paper is consistent with, indeed complementary to, the thesis of Brandom’s latest book, *Between Saying and Doing.* The book’s main point regarding normative vocabulary is that the capacity to employ normative vocabulary can be algorithmically elaborated from the capacity to engage in autonomous discourse, while the employment of normative vocabulary serves to make commitments at work in autonomous discourse explicit which would, without it, remain implicit – see Brandom (2008: lecture 2). This idea can plausibly be read as an answer to the question *how* normative talk succeeds in effecting a mutual calibration of different agents’ dispositions.

[51] McDowell (1994: 92)

[52] McDowell (1994: 34ff., 83, among other places)

[53] McDowell (1994: 178)

This is a penultimate draft –
please cite from the version of
the article which appeared in
Philosophical Investigations
35.2, 2012, pp. 101-126

concerning *meaning*, the task of philosophy, according to McDowell, is "to dislodge the assumptions that make it look difficult to find a place for meaning in the world. Then we can take in our stride meaning's role in shaping our lives. We do not need a constructive legitimizing of its place in our conception of ourselves."[54]

It may come as a relief to Wittgensteinian or McDowellian minimalists that at least one of minimalism's two commitments – namely, the first – is fully shared by the circulatory account of norms given in this paper. Admittedly, our account falls somewhat short of living up to the second commitment. But then, it is much more dubious than the first. One way to "legitimise" the place of meaning "in our conception of ourselves" is, surely, to say just how the phenomenon of meaning is systematically connected with, or produced by, what we collectively do. And in the sense that the relevant systematicity is both somewhat complex and not fully transparent to all participants of meaning talk, this "legitimisation" surely constitutes what McDowell calls "constructive philosophy". But as long as the warning against looking at our social practices "from sideways-on" is heeded, it is unclear why this should be so bad. Besides, our account can actually be given an interpretation on which it is not so far from instantiating the second commitment, too. Firstly, the circulatory account of norms can be read as focussed on "dislodging assumptions" about normative claims, in particular the assumption that making a normative claim is *not* to exhibit a practical attitude (and thus to *influence* the normative status the claim is about). Secondly, the account does not tell us something of which we were completely unaware. Rather, it rearranges things we have already known into a potent response to our philosophical puzzlement about norms. After all, the calibratory nature of normative vocabulary is reflected by the – well known and quite transparent – fact that the less in line we are with one another, the more we need normative talk, and the more in line we are, the less we need it.

With the potential Wittgensteinian or McDowellian reservations out of the way, let me close by sketching the fruitfulness of the circulatory account with respect to Brandom's Kantian idea of building intentionality on normativity. We can now formulate the Brandomian thesis about intentionality as follows: to say that a being displays intentionality is to say that it is a potential or actual participant in one or more of the practices which lend themselves to the social calibration of individual attitudes of the sort discussed in this paper.[55] If we want to make room for different kinds of intentionality, then we can divide the field along the different qualifying practices, and we can – of course – reserve a special place for those who master the actual calibrational practice of normative talk: speakers – *us*.

---

[54] McDowell (1994: 176, cf. 95 and 174)

References

Boghossian, Paul (1989): "The rule-following considerations", in *Mind* 98, 507–549
Brandom, Robert B. (1994): *Making It Explicit. Reasoning, Representing, and Discursive Commitment,* Cambridge Mass.: Harvard University Press
Brandom, Robert B. (2001): "Modality, normativity, and intentionality", in: *Philosophy and Phenomenological Research* 63, 587–609
Brandom, Robert B. (2002): "Some pragmatist themes in Hegel's idealism", in his: *Tales of the Mighty Dead,* pp. 210–234, Cambridge Mass.: Harvard University Press
Brandom, Robert B. (2008): *Between Saying & Doing. Towards an Analytic Pragmatism,* Oxford: Oxford University Press
Brandom, Robert B. (2009): *Reason in Philosophy. Animating Ideas,* Cambridge Mass.: Harvard University Press
Burge, Tyler (1979): "Individualism and the mental", in: *Midwest Studies in Philosophy* 4, 73–121
Cavell, Stanley (1969): *Must We Mean What We Say?,* Cambridge: Cambridge University Press
Davidson, Donald (1986): "A nice derangement of epitaphs", in: E. Lepore (Ed.), *Truth and Interpretation. Perspectives on the Philosophy of Donald Davidson,* Oxford: Blackwell
Dennett, Daniel C. (1979): "Intentional systems", in his: *Brainstorms. Philosophical Essays on Mind and Psychology,* Sussex: Harvester Press.
Dennett, Daniel C. (2006): "The evolution of 'why?'", accessed online at http://ase.tufts.edu/cogstud/papers/Brandom.pdf on 31 March 2009.
Ebbs, Gary (1997): *Rule-Following and Realism,* Cambridge Mass.: Harvard University Press
Hattiangadi, Anandi (2003): "Making it implicit. Brandom on Rule-Following", in: *Philosophy and Phenomenological Research* 66(2), 419–431
Kant, Immanuel (1998): *Critique of Pure Reason,* Cambridge: Cambridge University Press
Kripke, Saul A. (1972): *Naming and Necessity,* Oxford: Blackwell
Lance, Michael N. and John O'Leary-Hawthorne (1997): *The Grammar of Meaning. Normativity and Semantic Discourse,* Cambridge: Cambridge University Press
Lewis, David K. (1969): *Convention. A Philosophical Study,* Cambridge Mass.: Harvard University Press
McDowell, John H. (1994): *Mind and World,* Cambridge Mass.: Harvard University Press

---

[55] It should be noted that ascribing a *particular* intentional state to some being, on a Brandomian account, amounts to *entering* into the calibrational practice (with other observers and, depending on the kind of intentionality ascribed, the being to which the particular intentional state is being ascribed).

Putnam, Hilary (1975): “The meaning of ‘meaning’”, in his: *Mind, Language and Reality*, pp. 215–271, Cambridge: Cambridge University Press

Rosen, Gideon (1997): “Who makes the rules around here?”, in: *Philosophy and Phenomenological Research* 57(1), 163–171

Wittgenstein, Ludwig (1972): *On Certainty. Über Gewissheit,* New York, Hagerstown, San Francisco, London: Harper & Row

Wittgenstein, Ludwig (1978): *Remarks on Colour. Bemerkungen über die Farben,* Oxford: Blackwell

Wittgenstein, Ludwig (1983): *Remarks on the Foundations of Mathematics,* Cambridge Mass.: MIT Press

Wittgenstein, Ludwig (2001): *Philosophical Investigations. Philosophische Untersuchungen,* Oxford: Blackwell