

Meaning, Dispositions and Supervenience

Leonardo Marchettoni

ciclvm@tin.it

Jura Gentium: Centre for Philosophy of International Law and Global Politics,
<<http://www.juragentium.unifi.it/>>

Department of Theory and History of Law
University of Florence

Abstract: In this paper I raise some doubts about Brandom's pragmatic strategy of explanation of norms. I argue that Brandom's attempt to explain normative statuses through recourse to normative attitudes does not succeed in preserving a hiatus between norms and regularities of behaviour. Since calibrating one's own behaviour as a consequence of normative assessments can be described, at least in principle, in non-normative vocabulary, the upshot of Brandom's pragmatism about norms is an account of normative phenomena—and especially of semantic phenomena—that does not require reference to normative notions.

My aim in this paper is to discuss some issues about the thesis of the supervenience of norms that Brandom defends in *Making It Explicit* (henceforth *MIE*). The basic idea is that the pragmatist strategy of explanation of the normative aspects of intentional phenomena does not succeed in distinguishing itself from naturalistic approaches: pragmatism about norms explains normativity through a reduction of norms to notions that are in the end—contrary to what Brandom asserts—non-normative.

The plan of the essay is as follows: in the first section I consider Brandom's arguments against the reducibility of attitudes to dispositions and argue that they are less than conclusive. In the second section I will then contend that Brandom's thesis of the supervenience of norms on deontic statuses and hence on normative attitudes is untenable, because norms are indeed reducible to attitudes and, for what I have argued in the first section, to dispositions.

1. Attitudes and Dispositions

Brandom's solution to the problem of rule-following is centred around the idea that we can explain the existence of rules if we focus on our activity of treating performances as correct or incorrect. In this way normative statuses are taken to supervene on normative attitudes, which in turn are deemed to be non-describable in purely naturalistic terms.¹ Is this assumption reasonable? To answer this question, we need to examine his account of normative attitudes.

Normative attitudes are assessments, “assignments to performances of normative significance or status as correct or incorrect according to some norm” (*MIE*, p. 35). But assessments can be understood as dispositions to sanction, that is, to reward appropriate and to punish inappropriate performances, as Brandom recognizes.² Therefore, it seems that normative attitudes are completely explainable in naturalistic terms: they seem reducible to clusters of dispositions. When we speak of one's attitudes we are actually speaking of her dispositions to react to certain performances sanctioning them, positively or negatively. But

¹ See *MIE*, ch. 1.

² Cf. *MIE*, pp. 34ff.

if this hypothesis were correct, we could conclude that there is a naturalistic description of norms that is couched entirely in non-normative terms.

This line of reasoning, however, is quickly dismissed.³ In fact, Brandom offers two arguments against the reducibility of attitudes to dispositions.

(A) The first argument is based on the observation that the normative character of the metalanguage in which norm-instituting social practices are specified is irreducible to naturalistic accounts:

... it is important to realize that it is one thing to understand practical assessment as sanctioning, and quite another to understand sanctioning in non-normative terms such as reinforcement. ... Defining normative attitudes in terms of dispositions to apply sanctions does not by itself reduce the normative to the non-normative—it just trades off one sort of norm for another. At the most basic level, to reward someone is to offer some good ..., and to punish them is conversely to inflict something bad. Benefit and harm, desirable and undesirable, are concepts that also have normative senses. Indeed, these senses would seem to be primary, so that some sort of reductive hypothesis would be needed to naturalize them. (*MIE*, p. 42)

According to Brandom the reduction of attitudes to dispositions to apply sanctions cannot be a proper reduction of normative notions to non-normative ones, because the description of attitudes as dispositions to sanction is not entirely couched in naturalistic vocabulary. In fact, the concept of ‘sanction’ is a normative notion, something that refers back to a normative theory of what benefit and harm consists in.

(B) The second argument against the naturalizability of normative attitudes starts by noticing that positive and negative sanctions need not consist in rewards and punishments. Indeed, they “may consist in acclaim and censure that itself has only a normative significance” (*MIE*, p. 43). A correct performance can be rewarded by release from an obligation; in the same way, an incorrect execution might be punished by withholding a license. In such cases there is no direct shift from normative evaluations to bestowals of benefits or impositions of harms; we simply face “a change in normative status rather than natural state” (*MIE*, *ibidem*). Following this remark it is possible to distinguish between *external* and *internal* sanctions. External sanctions are those sanctions that are expressible in non-normative terms, like offering something good or inflicting a physical punishment. Instead, internal sanctions are those sanctions that involve only a change of normative status. Now, for Brandom

[i]t is possible to interpret a community as instituting normative statuses by their attitudes of assessment, even though each such status that is discerned is responded to by sanctions that involve only other normative statuses. ... Such an interpretation would not support any reduction of normative status to non-normatively specifiable dispositions, whether to perform or to assess, whether individual or communal. (*MIE*, p. 42)

³ Even though Brandom criticizes explicitly the identification between assessments and dispositions to sanction, several commentators have attributed to him this view: “So put it appears as though Brandom is offering a dispositionalist account of the determination of correctness—since the starting point includes nothing more than behavioural dispositions. Moreover, ... nothing is added that would distinguish the account from dispositionalism” (Hattiangadi 2003, p. 425). However, this seems to me a hasty conclusion, because to establish the thesis one would have first to address the arguments Brandom adduces against the naturalizability of normative attitudes. See also Grönert 2005, p. 163-4.

Brandom takes the possibility of a community that dispenses with external sanctions to show the non reducibility of normative statuses to non-normative notions. Therefore, Brandom's defence of the irreducibility of attitudes to dispositions seems to entail the following modal claim:

- (1) Necessarily, normative attitudes, *at least when internal sanctions are at work*, are not epistemically reducible to clusters of behavioural dispositions.⁴

From this follows that, if there are reasons to reject (1), Brandom's defence is unsuccessful.

The simplest way to refute (1) consists in providing some counterexample. Since (1) is a necessary statement it suffices to exhibit a putative case in which (1) fails. To secure this outcome we must understand how is it possible to reduce a normative attitude concerning an internal sanction to a disposition articulated in naturalistic terms. The major obstacle seems to be that of expressing a change of normative status without reference to normative notions. In fact, for what concerns the argument (A), it is possible to notice that, since individuals who share the same conception of benefit and harm may disagree in their attitudes towards a given performance, and conversely, subjects having different conceptions of good and evil may express the same attitudes, the choice of the sanctions is not relevant to define the status of a performance as correct or incorrect.⁵ If this were not precluded by the variability of sanctions, we could express normative attitudes—as conceptually separate from the specification of the behavioural outputs in which they result—as dispositions, in purely non-normative terms. For example: we can consistently imagine a community in which people are disposed to adopt a single type of positive sanction and a single type of negative sanction. In such a case a dispositionalist account of normative attitudes entirely couched in naturalistic terms would be available.

But the difficulty posed by the argument (B) is not so easy to cope with. To overcome the problem linked to the distinction between external and internal sanctions we could first note that once an internal sanction is applied it is reasonable to expect that the community members begin to adjust their behaviour in accordance with the change of normative status following the sanction: for example, by declaring what additional duties the transgressor is now required to accomplish. Basically, a behavioural disposition is a way of associating a certain behavioural output with a given circumstance without appealing to inner states. Therefore, it is still possible to conceive of normative attitudes pertaining to internal sanctions in dispositional terms, not as dispositions to sanction, but as sets of dispositions to calibrate one's behaviour as a consequence of changes of normative status.

Consider again the previous example. Instead of a community in which all individuals are disposed to sanction in the same way, we can imagine a community in which every

⁴ Where a is epistemically reducible to $b =_{df}$ everything that can be said about a and its properties, can be said by speaking only of b and its properties. This formulation aims to capture Brandom's idea of reduction as the activity of eliminating, "in favour of non-normative or naturalistic vocabulary, the normative vocabulary employed in specifying the practices that are the use of language" (*MIE*, p. xiii). There is no attempt to reduce *ontologically* attitudes to dispositions: this issue is beyond the scope of this essay.

⁵ It is useful to compare this case with that of communal assessment theories about rule-following. Brandom maintains that communal assessment theorists make illicit reference to a normative concept when they define correctness in terms of another normative concept, the concept of 'expert' (*MIE*, p. 39). This objection seems reasonable, because the reference to the normative notion of 'expert' is necessary in order to determine the extension of the set of correct performances—since this set is by hypothesis composed of just those performances that are so evaluated by experts. So in this case—but not in the case of the reduction of attitudes to dispositions to sanction—the reference to a normative concept is *necessary* in order to determine the content of attitudes towards performances.

member must perform some task from a given set A . When someone fails to carry out her duty, her normative status changes and she is committed to performing some task from another set, say B . If she fails again, she is assigned a task from a third set, C . Eventually, if inaccuracy persists, the community member is charged with an external sanction S . If we should describe the normative attitudes of these folks in evaluating other community members' performances, we could avail ourselves of the dispositionalist jargon as follows: everyone has a set of *first-order* and *second-order* dispositions. First-order dispositions are instructions that specify how to accomplish certain tasks and how to behave towards other community members; second-order dispositions are instructions that specify how to react to others' practitioners performances. If we assume that each individual has some default dispositions and define a second-order disposition as a disposition to alter one's own pre-existing dispositions, we can interpret normative attitudes as second-order dispositions, that is dispositions to keep track of others' performances and modify one's own default dispositions in accord with new evidence concerning how a certain performance has been carried out. In this way the original behaviourist motivation behind the dispositionalist interpretation of attitudes can be preserved without reintroducing normative notions. In fact, once it is defined a convenient set of first-order dispositions and of rules that explain in which way an agent alters her first-order dispositions as a consequence of getting new information, it becomes possible to display a purely naturalistic account of the activity of assessing performances in which changes of normative status are expressed in terms of second-order dispositions. Moreover, in this account the difference between internal and external sanctions is made harmless, because it responds to a difference in the sets of second-order dispositions.

It is not difficult to adapt this dispositional model to the case of linguistic practice. In order to accomplish this task it is necessary to take into account inferential as well representational dimensions of linguistic activity.⁶ This can be done, following the trace of Brandom's scorekeeping semantic, as follows:

- (2) For a given atomic sentence φ , the meaning of φ , M^φ , consists in an ordered quintuple, $M^\varphi: \langle C^\varphi, P^\varphi, E^\varphi, I^\varphi, A^\varphi \rangle$, where C^φ is the set of commitment-preserving inferences in which φ serves as conclusion, P^φ is the set of commitment-preserving inferences in which (along with other auxiliary hypothesis) φ plays an essential role as premise, E^φ is the set of entitlement-preserving inferences in which (along with other auxiliary hypothesis) φ plays an essential role, I^φ is the set of the sentences that are incompatible with φ and A^φ is the set of the states of affairs of which φ constitutes an appropriate observation report.⁷

Each speaker in her use of words follows what she believes to be the conditions of correct use. So, she must have some ideas about what assertions commit to a given sentence φ , to what further claims she is committed as a consequence of her assertion of φ , to what assertions she is entitled as a consequence of entitlement to φ , what sentences are incompatible with the assertion of φ , what states of affairs can be described by means of φ .

However, two or more speakers can differ in their judgments about what inferences are licensed by a given utterance or about what states of affairs can be correctly described by means of a given sentence. This complication is acknowledged by Brandom himself when he recognizes the perspectival character of conceptual content, that is, the fact that what is assumed to constitute the content of a sentence varies from speaker to speaker, according to

⁶ Actually, we should consider also 'practical commitments', for example commitments to act following from the utterance of a given sentence (see *MIE*, ch. 4). However, I will ignore this problem in what follows.

⁷ See *MIE*, pp. 188-189.

the linguistic customs each speaker has acquired.⁸ Noticing that content is perspectival imposes a slight reformulation of our previous definition of meaning in terms of sets of sentences and states of affairs. Since each speaker can attribute a different meaning to a given assertion, it appears reasonable to index meaning in relation to speakers. As result, we have that the meaning of a given sentence φ is expressible as an indexed family of sets, $M_i^\varphi: \langle C_i^\varphi, P_i^\varphi, E_i^\varphi, I_i^\varphi, A_i^\varphi \rangle$, in which each M_j^φ represents the meaning of φ in the perspective of a single speaker j .

Once we have defined what performances are comprised in the activity of classifying correct and incorrect uses of sentences, it is possible to state more accurately the content of a pragmatist approach to semantics. In the general case the presence of rules was explained through the recourse to attitudes instituting normative statuses of performances. If we want to adjust this style of reasoning to semantic norms we have to identify the corresponding normative statuses and normative attitudes. The preceding discussion has suggested that normative statuses pertaining to semantic norms are connected with the activity of individuating, for a given sentence φ , and for a given speaker j , what inferences form the sets $C_i^\varphi, P_i^\varphi, E_i^\varphi$, what sentences form the set I_i^φ , and what states of affairs are collected in the set A_i^φ . The upshot is that we can easily transpose the abstract analysis of rules to the case of semantic norms if we take into account the five components of linguistic meaning, $\langle C_i^\varphi, P_i^\varphi, E_i^\varphi, I_i^\varphi, A_i^\varphi \rangle$, and the corresponding deontic attitudes of treating someone as committed or entitled to or interdicted from further assertions as a consequence of the utterance of a given sentence.

Consider, for example, a speaker j and a set of n sentences S . Since meaning is perspectival, for each sentence $\varphi \in S$ there is a quintuple M_j^φ that supplies with the conditions of correct use for that sentence in that speaker's perspective. M_j^φ indicates what inferences follow from φ and what states of affairs can be reliably reported by means of it in j 's perspective and so explains how j is prepared to use φ in her linguistic exchanges with other speakers. We can call $M_j^S = \{M_j^\varphi \mid \varphi \in S\}$ the set consisting of all the quintuples M_j^φ for all the sentences included in S . Now, we can extend this rough model in a simple way. We can imagine that j , when enters into contact with another speaker k , is prepared to behave linguistically according to the conditions stated by M_j^S and also to evaluate k 's utterances along the same lines. This last condition implies that j will keep track of k 's entitlements and commitments and of hers and will form corresponding deontic attitudes that specify to what further performances k is committed or entitled, thus regulating her further linguistic behaviour on these bases. Moreover, it requires that j will note the divergences between k 's further utterances and her expectations and revise his future expectations⁹ and, eventually, if there is any reason to recognize to k some special linguistic authority, modify her own attitudes to linguistic behaviour and evaluation.

What does Brandom's treatment of these issues in *Between Saying & Doing* (henceforth *BSD*) add to our discussion? In his later book Brandom is primarily concerned with the relationship between languages and practices in a more abstract setting than that deployed in *MIE*. Nothing is said about attitudes and dispositions, even if Brandom repeatedly quotes Huw Price's thesis that normative vocabulary is irreducible to naturalistic one but that one can specify in naturalistic terms what a given subject must do in order to deploy some non-naturalistic vocabularies.¹⁰ As for what regards the problem I am considering the most

⁸ See also *MIE*, p. 185.

⁹ This means that j will keep two separate registers, concerning respectively the objective meaning of k 's assertions—objective in j 's perspective, obviously—and the meaning that k herself attributes to her assertions.

¹⁰ Cf. Price 2004. I think that thi thesis is in tension with the pragmatic strategy adopted in *MIE*, but I cannot pursue this issue in this essay.

important reflections are contained in the third lecture. Here Brandom criticizes the tenets of AI functionalism. More particularly, Brandom critiques the idea that

there is a set of practices-or-abilities meeting two conditions:

1. It can be algorithmically elaborated into (the ability to engage in) an autonomous discursive practice (ADP).
2. Every element in the set of primitive practices-or-abilities can intelligibly be understood to be engaged in, possessed, exercised, or exhibited by something that does *not* engage in any ADP. (*BSD*, p. 75, italic in text)

If Brandom's worries are well founded, it can be noted, the project of semantic naturalism is not doomed to failure, since one could equally maintain that the primitive abilities that precede ADPs bear some other relation to ADPs (for example, one could hold that ADPs *emerges* from non-semantic abilities). However, it is clear that, if Brandom is right in his contention, the naturalistic project should be rethought. Surely, in fact, the identification between attitudes and dispositions cannot be preserved if ADPs are not algorithmically decomposable into primitive abilities.

Against this idea, Brandom recognizes to have no "knock-down arguments". However from the text it is possible to elicit the following qualms:

(A') The practice of adjusting one's other beliefs in response to a change of belief is intrinsically holistic; this raises the problem of revising and updating one's commitments and entitlements in the right way, that is in a way that be sensitive to one's other collateral commitments and entitlements. Since "*any* change in *any* property of one changes *some* of the relational properties of *all* the rest" (*BSD*, pp. 80-81, italics in text), "it is *not* plausible ... that *this* ability can be algorithmically decomposed into abilities exhibitable by non-linguistic creatures" (*BSD*, p. 81, italics in text), because each attempt to deal with this difficulty should face the problem of finding a rule to determine what factors are to be ignored. This trouble is assuaged in linguistic creatures, for the latter have semantic, cognitive, or practical access to the complex relational properties they would have to distinguish to assess the goodness of many material inferences.¹¹

(B') Brandom contrasts algorithmic decomposition into primitive abilities with training by an expert. A *course of training* can be thought of "as having as its basic unit a stimulus (perhaps provided by the trainer), a response on the part of the trainee, a response by the trainer to that response, and a response to that response by the trainee that involves altering his dispositions to respond to future stimuli" (*BSD*, p. 87, italic in text). Moreover, the abilities interested by this process "vary wildly from case to case, and depend heavily on parochial biological, sociological, historical, psychological and biographical contingencies" (*BSD*, p. 85). Finally, the question of what algorithmic elaboration is sufficient for a particular creature, in a particular context cannot be settled empirically.

To these worries it is possible to reply as follows. For what concerns the issues grouped under the heading (B'). The way in which Brandom treats the phenomenon of training suggests that there is nothing magical or mysterious in the manner trainees are instructed by their trainers. After all, Brandom himself proposes an abstract model of what a course of training should consist in:¹² a series of responses from the part of the trainee to which the trainer reacts with appropriate corrections. It seems to me that in this succession of events

¹¹ Cf. *BSD*, p. 83.

¹² Brandom's description of training, it can be noted in passing, reminds what Donald Davidson called, in his later writings, *triangulation*. See, for example, Davidson 2001.

there is nothing that cannot be algorithmically decomposed: it suffices to set a sequence of stimuli arranged in a proper way, so that each stimulus be related to the preceding responses of the trainee. (Obviously the practical implementation of this model can pose almost insurmountable difficulties; but this problem is connected with the troubles that come out from the argument A'.) Moreover, it is true that the abilities involved vary dramatically from case to case and that the success of a particular course of training cannot be predicted in advance. But these empirical limitations do not affect the algorithmic decomposability *in principle* of the process of training. They are rather to be viewed as contingent features of the training process: since a single course of training can be implemented by activating different sets of abilities, the variability of the abilities involved in each particular case is connected with the multiple realizability of the overall process, in accord with the functionalist thesis.

If this is correct, the point that emerges from the objection (B') is strictly related to the worries posed by the argument (A'). It is the practical intractability of the holistic character of our activity of revising our beliefs that motivate Brandom's discomfort with the possibility of an algorithmic decomposition of the ability to engage in an ADP. But, whereas it is possible to agree with him that a proper treatment of semantic holism poses a formidable obstacle to our efforts of creating a computational system capable of engaging in linguistic practice, no evidence is offered for the stronger claim that such a treatment is *in principle* impossible. From the fact that current researchers are not able to *simulate* ADP, does not follow that this task cannot be accomplished.¹³

Moreover, the idea that only those creatures which have something like a semantic access to the complex relational properties they would have to distinguish to assess the goodness of many material inferences are able to engage in ADPs, is in tension with the semantic project pursued by Brandom in *MIE*. In fact, from that work can be elicited a conception of meaning which exploits, as its basic ingredients, the performances of the speakers that are caught in the game of adjusting their beliefs in response to the assertions of other speakers. For these reasons it seems to me that the scepticism that Brandom now shows towards algorithmic decomposability of the performance of revising one's beliefs contrasts with the role that in *MIE* is assigned to deontic scorekeeping: if an access to semantic properties is necessary in order to exhibit those abilities that are required to engage in an ADP the semantic project pursued in *MIE* becomes irredeemably circular, because one cannot see what utility possesses a semantic explanation that employ as its basic ingredients notions that are just semantic in advance.

My conclusion is that Brandom's critiques to AI functionalism are less than decisive. But if there is no conclusive argument against the algorithmic decomposability of the ability to engage in an ADP, and the analogue arguments offered in *MIE* against the identification of normative attitudes with dispositions equally fail, there is a naturalistic reading of Brandom's pragmaticist semantic that starts with the interpretation of the basic attitudes as dispositions and ends with a reduction of norms to naturalistic items, as I am going to show.

2. Supervenience of Norms

I have argued that Brandom's arguments against the naturalizability of normative attitudes are far from being conclusive. But if we cannot dismiss the idea that normative attitudes are, at least in principle, reducible to behavioural dispositions—that can be accounted for in

¹³ The issue is the object of a large body of work, from connectionist approaches to fuzzy logic, that far exceeds the scope of this paper. At present, I would like to mention only the recent and promising attempt by Andy Clark of challenging Fodor's frame problem. See Clark 2002.

non-normative vocabulary—we must also admit, via the thesis of supervenience of normative statuses on normative attitudes, that normative statuses can be taken to supervene on dispositions to regulate one’s own behaviour as a consequence of normative assessments. And this conclusion poses some problems, because it seems to entail a new variety of naturalism about norms. But to verify whether this suspicion is a sensible one we should first try to understand what the supervenience thesis exactly implies.

To see how this happens, it is useful to reformulate the thesis of the supervenience of normative statuses in a slightly more technical fashion. Saying that normative statuses supervene on normative attitudes means, in Brandom’s words, that “settling all the facts concerning normative *attitudes* settles all the facts concerning normative *statuses*” (*MIE*, p. 47, italics in original).¹⁴ This description suggests a *global supervenience* of statuses on attitudes.¹⁵ However, since normative statuses are instituted by attitudes, it seems reasonable to interpret the kind of dependence Brandom has in mind as asserting also that if two individuals entertain the same normative attitudes they institute the same normative statuses.¹⁶ This kind of dependence can be expressed as a *weak supervenience* of statuses on attitudes.¹⁷ Indeed, global supervenience formally does not entail weak supervenience. However, it has been shown that it does if we consider only intrinsic properties, and the property of entertaining a given attitude towards a performance so and so is certainly intrinsic. In fact, global supervenience entails *strong supervenience* either—where strong supervenience is the thesis that if two individuals, whether in the same or different possible worlds, entertain the same normative attitudes they institute the same normative statuses—if we limit our attention to intrinsic properties, so in what follows I will assume that Brandom is committed to the thesis of the strong supervenience of normative statuses on normative attitudes.¹⁸ Moreover, since attitudes are—at least in principle—reducible to behavioural dispositions, this thesis entails that normative statuses strongly supervene on dispositions.

This way of stating the matter, however, does not seem completely correct. In fact, in the last chapter of *MIE*, Brandom advances a different explanation of the relationship between normative statuses and attitudes. Brandom stresses the fact that normative statuses are not instituted by *actual* attitudes but only by *correct* attitudes. The institution of statuses should be understood

in terms of the implicit practical *proprieties* governing such scorekeeping—not how the score is actually kept but how, according to the implicitly normative scorekeeping practices it *ought* to be kept, how scorekeepers are *obliged* or *committed* to adopt and

¹⁴ It should be noticed, however, that Brandom’s thesis that the conceptual proprieties implicit in discursive practices incorporate empirical dimensions (see *MIE*, pp. 119-120 and 331-332) could revoke into doubt this global supervenience thesis, since earthlings and twin-earthlings count as instituting different conceptual contents even though entertaining the same attitudes.

¹⁵ See Rosen 1997, pp. 164-165.

¹⁶ While it is commonly agreed that Brandom’s theory of supervenience of rules can be expressed as a weak supervenience thesis, it is a more disputed issue whether it entails a strong supervenience thesis either—that is a relation of covariance that holds *necessarily*, for all possible worlds—too. See Loeffler 2005, p. 58.

¹⁷ More precisely, saying that *A*-properties weakly supervene on *B*-properties implies that necessarily (that is, in every possible world), if two objects possess the same *B*-properties they share also the same *A*-properties. For further details, see Kim 2003, p. 559.

¹⁸ *A*-properties strongly supervene on *B*-properties iff if two objects, whether in the same or different possible worlds, possess the same *B*-properties they share also the same *A*-properties. See again Kim 2003, p. 560.

alter their deontic attitudes, rather than how they actually do. (*MIE*, p. 628, italics in original)

What distinguishes the attitudes that are capable of instituting statuses is their correctness. The institution of statuses is a consequence of keeping the deontic score in the right way.¹⁹ So, not every attitude institutes a corresponding status but only those attitudes that are adopted according to a proper scorekeeping activity. In this sense normative statuses are a product of an idealization of the actual scorekeeping practices.²⁰

Therefore, we have to take into account not actual normative attitudes but only attitudes that are correct in the sense explained above, that is consistent normative attitudes and deontic attitudes that are correctly taken as a consequence of the encounter with other community members' performances. To translate this definition in the dispositionalist jargon, we should be able to express the idea that normative statuses strongly supervene on sets of dispositions correctly displayed in patterns of social interaction. Recalling the previous discussion about the dispositionalist interpretation of attitudes, it can be suggested that we may attempt to formulate this point in terms of second-order dispositions. For example, if we interpret the family of sets M_j^S as determining j 's default first-order dispositions to linguistic behaviour—that is, the way in which, at the beginning of the conversation, j is prepared to behave and to evaluate another speaker's assertions—we can conceive of deontic attitudes as second-order dispositions that specify in which way a given speaker j , who keeps track of others' commitments and entitlements, will modify her previous dispositions in accordance with new evidence concerning how a certain linguistic performance has been carried out—for instance, adjusting the deontic status of her interlocutor (intended as a cluster of first-order dispositions).²¹ Therefore, the supervenience thesis can be expressed as the thesis according to which normative statuses strongly supervene on correct sets of second-order dispositions—where 'correct' can be obviously defined in non-normative terms through reference to the way in which second-order dispositions keep track of previous performances—and indirectly on sets of first-order dispositions. (In fact, a second-order disposition must refer to the criteria set up by first-order dispositions to individuate the performances that require to be sanctioned and the first-order dispositions that are to be modified.)

¹⁹ In the same vein, Ronald Loeffler has recently maintained that the right way to intend the supervenience thesis is as asserting that “not *de facto* normative attitudes, but only attitudes that *should* be adopted or that are *properly* adopted, determine semantic norms” (Loeffler 2005, p. 62, italics in original).

²⁰ Now, however, one could ask what are the norms according to which the judgments concerning the correctness of attitudes are made. But Brandom's answer to this last question could hardly be considered satisfying. The only partial response, in fact, comes at the very end of the book and seems to consist in the rather disappointing admission that identifying the parameters of correctness is entirely up to the interpreter who attempts to reconstruct the discursive scorekeeping practices. According to Brandom the norms that determine when it is correct for an agent to attribute a certain doxastic commitment to someone else are not available in advance as a set of explicit principles “but are implicit in the particular practices by which we understand one another in ordinary conversation” (*MIE*, p. 646). Moreover, since the external interpretation of a linguistic community is not qualitatively different from ordinary scorekeeping activity, “[t]here is never a final answer as to what is correct; everything, including our assessments of such correctness is itself a subject for conversation and further assessment, challenge, defense, and correction” (*MIE*, p. 647). See also Laurier 2005, pp. 156-158.

²¹ It is reasonable to imagine that there should be a set of dispositions that specify in which cases the speaker has to recognize her interlocutor some kind of linguistic authority, but this is a point I will not pursue further.

Now, we have to understand whether this way of linking normative statuses to behavioural dispositions matches Brandom's anti-reductionist premises. I have recalled that Brandom is committed to the supervenience of normative statuses, and hence of norms, on normative attitudes.²² But then, if we accept the reducibility of normative attitudes to behavioural dispositions and formulate the thesis of supervenience of norms as a strong supervenience thesis, we obtain the following:

- (3) *Supervenience of norms on behavioural dispositions*: If two individuals possess the same (correct second-order and consistent first-order) behavioural dispositions, whether in the same or different possible worlds, they can be said to institute the same norms.²³

To make this idea more precise we can define a set of dispositions as a function $D: S \rightarrow B$ from states of affairs to behavioural responses. In other words, a set of dispositions can be conceived as a set of pairs $D \subset S \times B$ where S is a set of states of affairs and B is a set of patterns of behaviour. Given this definition we can say that an individual x possesses a set of dispositions D only if for each $(s, b) \in D$, if x is in the state of affairs s she will act in accordance with the pattern of behaviour b .²⁴ Accordingly, the thesis of the supervenience of norms becomes as follows:

- (4) *Supervenience of norms on behavioural dispositions*: Necessarily, if x , in a certain state of affairs s , institutes a norm N , there exists a set of dispositions $D: S \rightarrow B$ such that $s \in S$, and x possesses D , and everyone that is in s and possesses D institutes N .²⁵

This definition states that two subjects can be said to institute the same norm if they have an identical set of dispositions—and these dispositions are correct, in the sense explained above. Each individual that presents a correct set of dispositions of this kind can be said to institute the related norm. It is also important, however, to make clear what supervenience does not imply. Supervenience entails that individuals having the same correct dispositions institute the same norms, but the converse does not hold. On the contrary, it is a central feature of the concept of supervenience that if a set of properties A supervenes on a set of properties B , a property $a \in A$ can supervene on several different subsets of B . This means that, if norms supervene on behavioural dispositions, it is possible that two individuals institute the same norm even if they possess different dispositions. But how does it happen?

In response to this question, one could maintain that the variability in the dispositional basis is connected with the fact that different individuals may be disposed to react to other practitioners' performances in different manners. So, two individuals can institute the same norm even if, for example, one of them is disposed to apply external sanctions whereas the other applies only internal sanctions. (It is obviously possible to imagine far more ingenious variations in the dispositional basis than these differences in the ways of sanctioning. This, however, would not change the line of reasoning I am considering.)

²² In fact, if settling all the facts concerning normative attitudes settles all the facts concerning normative statuses, two worlds that are alike for what concerns normative attitudes cannot differ in their norms.

²³ This definition is adapted from Loeffler 2005, p. 58.

²⁴ That possess of a disposition entails a subjunctive conditional—at least in *ideal conditions*—is presumably uncontroversial. I will not enter—nor this is relevant for my argument—into the much debated issue of reducibility of dispositions to conditionals—for further details, see Mumford 2003; Fara 2006.

²⁵ For this way of formulating strong supervenience see Kim 2003, p. 561.

But if this is so what follows for the prospect of a dispositionalist naturalism about norms? Well, it remains possible to take the union of the sets of dispositions that institute a given norm and assume that that norm is coextensive—indeed, reducible—to such a set. Thus, we obtain the following:

- (5) *Reduction of norms to behavioural dispositions*: If x_1 , in a certain state of affairs s , institutes a norm N and x_1 possesses a set of dispositions $D_1: S \rightarrow B$ such that $s \in S$, and x_2 , in a certain state of affairs s , institutes a norm N , and x_2 possesses a set of dispositions $D_2: S \rightarrow B$ such that $s \in S$, and ... then N is coextensive, hence reducible to $\cup D_i$.²⁶

This shows that Brandom's account points towards a reduction of norms to dispositions.²⁷ But then we face a complete reduction of normative phenomena to naturalistic facts. In other words, Brandom's account of norms—and Brandom's semantics—prove to constitute a new kind of naturalism about normativity and meaning: accepting Brandom's elucidation of norms means accepting the idea that there may be a story entirely couched in naturalistic terms that explains how individuals, starting from a small set of dispositions to social behaviour and to acquire new dispositions, can institute a whole world of norms.

3. Conclusions

In this paper I raised some doubts about Brandom's pragmatist strategy of explanation of norms. I argued that if we attempt to explain normative statuses through recourse to normative attitudes it is impossible to preserve a hiatus between norms and regularities of behaviour. Since understanding of norms is reducible to possessing the right behavioural first- and second-order dispositions, and since calibrating one's own behaviour as a consequence of normative assessments can be described, at least in principle, in non-normative vocabulary, the upshot of Brandom's pragmatism about norms is an account of normative phenomena—and especially of semantic phenomena—that does not require reference to normative notions.

This result is clearly at odds with some of Brandom's basic assumptions. As a consequence, we are left with two possibilities. Either we abandon the primitiveness of normative notions and accept the reducibility of norms to dispositions and consequently to naturalistic facts, or, if we want to safeguard the issue of non-reducibility, we must revise the pragmatist strategy followed by Brandom, especially for what concerns the aim of offering an explanation of normative phenomena in terms of social activity.²⁸ The first horn of the dilemma indicates a new route towards naturalism about norms; on the contrary, the second proposal seems to point in the direction of a partial separation of intentional vocabulary from the sphere of naturalistic reports. In both cases, however, it should be clear

²⁶ Cf. Kim 1990.

²⁷ A different problem—which I cannot currently pursue—is whether a given disposition may supervene on different categorical properties—for more on this issue see again Mumford 2003. Note, however, that for our present concern, once we grant that normative attitudes are indeed reducible to dispositions, the underlying ontology of dispositions themselves becomes largely irrelevant.

²⁸ This revision could follow several routes: for example, one could argue more directly against the behaviouristic interpretation of attitudes in terms of dispositions. Otherwise, one could reject the transition from global supervenience to strong supervenience, by admitting that facts about norms are instituted by normative attitudes, and hence by behavioural dispositions, merely in the sense that settling all the facts concerning normative attitudes settles all the facts concerning normative statuses, without this condition entailing that if two individuals share the same dispositions they institute the same statuses.

that the original inspiration of Brandom's approach—combining an account of the social institution of norms with the primitiveness of normative notions—cannot be preserved.

References

- Brandom, R.B. (1994) *Making It Explicit: Reasoning, Representing, and Discursive Commitment*, Cambridge, MA, Harvard University Press.
- Brandom, R.B. (2008) *Between Saying & Doing: Toward an Analytic Pragmatism*, Oxford, Oxford University Press.
- Clark, A. (2002) "Local Associations and Global Reason: Fodor's Frame Problem and Second-Order Search", *Cognitive Science Quarterly*, 2, pp. 115-40.
- Davidson, D. (2001) *Subjective, Intersubjective, Objective*, Oxford, Oxford University Press.
- Fara, M. (2006) "Dispositions", in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2006 edn, URL = <http://plato.stanford.edu/archives/fall2006/entries/dispositions/>.
- Grönert, P. (2005) "Brandom's Solution to the Objectivity Problem", *Pragmatics and Cognition*, 13, pp. 161-175.
- Hattiangadi, A. (2003) "Making It Implicit: Brandom on Rule-Following", *Philosophy and Phenomenological Research*, 66, pp. 419-431.
- Kim, J. (1990) "Supervenience as a Philosophical Concept" *Metaphilosophy*, 21, pp. 1-27.
- Kim, J. (2003) "Supervenience, Emergence, Realization, Reduction" in M.J. Loux, D.W. Zimmerman (eds.), *The Oxford Handbook of Metaphysics*, Oxford, Oxford University Press.
- Laurier, D. (2005) "Pragmatics, Pittsburgh Style", *Pragmatics and Cognition*, 13, pp. 141-160.
- Loeffler, R. (2005) "Normative Phenomenalism: On Robert Brandom's Practice-Based Explanation of Meaning", *European Journal of Philosophy*, 13, pp. 32-69.
- Mumford, S. (2003) *Dispositions*, pbk edn., Oxford, Oxford University Press.
- Price, H. (2004) "Naturalism without Representationalism", in M. De Caro, D. Macarthur (eds.), *Naturalism in Question*, Cambridge, MA, Harvard University Press.
- Rosen, G. (1997) "Who Makes the Rules Around Here?", *Philosophy and Phenomenological Research*, 57, pp. 163-171.