



Vol-444

TAP-2009

Towards an Analytic Pragmatism

Workshop on Bob Brandom's Recent Philosophy of Language

Proceedings of the Workshop on Bob Brandom's Recent Philosophy of
Language: Towards an Analytic Pragmatism
Genoa, Italy, April 19-23, 2009

Edited by

**Cristina Amoretti
Carlo Penco
Federico Pitto**

University of Genoa, Department of Philosophy
Via Balbi 4, 16126 Genoa, Italy

Table of Contents

Modality, Intentionality and Discursive Practice

Pragmatism, Inferentialism and Modality in Sellars's Arguments against Empiricism <i>Robert Brandom</i>	p. 6
Brandom's Five-Step Program for Modal Health <i>Fredrik Stjernberg</i>	p. 18
The Status of Intentional Vocabulary in Discursive Practice: Reading Making it Explicit in the Light of "Between Saying and Doing" <i>David Lauer</i>	p. 23

Logic, Semantics and the Theory of Meaning

Inferentialism and the Normativity of Meaning <i>Jaroslav Peregrin</i>	p. 33
Disagreement, Error and Two senses of Incompatibility <i>Tanja Pritzlaff</i>	p. 41
Varieties of Analytic Pragmatisms <i>Danielle Macbeth</i>	p. 49
Is logic Demarcated by Its Expressive Role? <i>Berhard Weiss</i>	p. 58
Logic and Pragmatics <i>Daniele Porello</i>	p. 69
Can Negation be Defined in Terms of Incompatibility? <i>Nils Kurbis</i>	p. 77
A Reconstruction of Hierarchical Relations between Incompatibility-Entailment, Committive, and Permissive Consequences <i>Edgar Andrade Lotero and Catarina Dutilh-Novaes</i>	p. 90

Intentionality and the Philosophy of Mind

Meaning, Disposition and Supervenience

Leonardo Marchettoni

p. 100

Mindful Consequences of Inferentialism and Normativism. Why Mental Episodes Ain't in the Head (at All)

Pierre Steiner

p. 112

How Analytic Philosophy Has Failed Cognitive Science

Robert Brandom

p. 121

Two Aspects of Content: Semantic Inferentialist's Reconciliation of Authoritative Self-Knowledge and Content Externalism

Shuheji Shimamura

p. 134

Explicating as Distancing

Jeremy Wanderer

p. 144

Mind Body Problem and Brandom's Analytic Pragmatism

Francois-Igor Pris

p. 146

Towards an Analytic Pragmatist Account of Folk Psychology

Derek W. Strijbos - Leon C. De Bruin

p. 152

Pragmatism and Metaphysics

Brandom between Anthropology and Metaphysics

Bernd Prien

p. 160

On the Very Idea of Brandom's Pragmatism

Tadeusz Szubka

p. 167

Pragmatism and Metaphysics in Brandom's "Between Saying and Doing"

Elena Ficara

p. 172

Metaphilosophical Reflections on the Idea of Metaphysics

Robert Brandom

p. 177

Appendix: Tutorials on Bob Brandom's Philosophy

p. 187

Assertion and Inference (on: Making It Explicit)

Carlo Penco

Some Pragmatist Themes in Brandom's Reconstructive Rationality (on: Tales of the Mighty Dead)

Italo Testa

Brandom and "The Logician's Dilemma" (on: Ch. II of Between Saying and Doing)

Raffaella Giovagnoli

Pragmatism and Intentionality (on: Ch. VI of Between Saying and Doing)

Giorgio Bertolotti

Modality, Intentionality and Discursive Practice

Pragmatism, Inferentialism, and Modality in Sellars's Arguments against Empiricism

Robert Brandom

Section One: Introduction

In this essay I want to place the arguments of “Empiricism and the Philosophy of Mind” into a slightly less local context, by tracing further, into neighboring works, some strands of argumentation that intersect and are woven together in his critique of empiricism in its two principal then-extant forms: traditional, and 20th century logical empiricism. Sellars always accepted that observation reports resulting non-inferentially from the exercise of perceptual language-entry capacities play both the privileged epistemological role of being the ultimate court of appeal for the justification of empirical knowledge-claims and therefore (given his inferentialist semantics) an essential semantic role in determining the contents of the empirical concepts applied in such judgments. But in accord with his stated aspiration to “move analytic philosophy from its Humean into its Kantian phase,” he was severely and in principle critical of empiricist ambitions and programs in epistemology and (especially) semantics that go beyond this minimal, carefully circumscribed characterization of the cognitive significance of sense experience. Indeed, I think the lasting philosophical interest of Sellars's thought lies primarily in the battery of original considerations and arguments he brings to bear against all weightier forms of empiricism. Some, but not all, of these are deployed in the opening critical portions of “Empiricism and the Philosophy of Mind,” where the ground is cleared and prepared for the constructive theorizing of the last half. But what is on offer there is only part of Sellars's overall critique of empiricism. We accordingly court misunderstanding of what is there if we do not appreciate the shape of the larger enterprise to which it contributes.

In an autobiographical sketch, Sellars dates his break with traditional empiricism to his Oxford days in the thirties. It was, he says, prompted by concern with understanding the sort of conceptual content that ought to be associated with “logical, causal, and deontological modalities.” Already at that point he says that he had the idea that

what was needed was a functional theory of concepts which would make their role in reasoning, rather than supposed origin in experience, their primary feature.¹

This telling passage introduces two of the master ideas that shape Sellars's critique of empiricism. The first is that a key criterion of adequacy with respect to which its semantics will be found wanting concerns its treatment of *modal* concepts. The second is that the remedy for this inadequacy lies in an alternative broadly functional approach to the semantics of these concepts that focuses on their *inferential roles*—as it were, looking *downstream* to their *use*, as well as *upstream* to the circumstances that elicit their application.

This second, inferential-functionalist, semantic idea looms large in “Empiricism and the Philosophy of Mind.” In fact, it provides the raw materials that are assembled and articulated into Sellars's positive account of the semantics of the concepts applied in reporting thoughts and sense-impressions. Concern with the significance of *modality* in the critique of empiricism, however, is almost wholly absent from that work (even though it is evident in articles Sellars wrote even earlier). I do not think that is because it was not, even then, an essential element of the larger picture of empiricism's failings that Sellars was seeking to convey, but rather because it was the result of a hard-won but successful divide-and-conquer expository strategy. That is, I conjecture that what made it possible for Sellars finally to write “Empiricism and the Philosophy of Mind” was figuring out a way to articulate the considerations he advances there *without* having also at the same time to explore the issues raised by empiricism's difficulties with modal concepts. Whether or not that conjecture about the intellectual-biographical significance of finding a narrative path that makes possible the separation of these aspects of his project is correct, I want to claim that it is important to understand what goes on in EPM in the light of the fuller picture of the expressive impoverishment of empiricism that becomes visible when we consider what Sellars says when he *does* turn his attention to the semantics of modality.

¹ In *Action, Knowledge, and Reality*, H. N. Castaneda (ed.) [Indianapolis, Bobbs-Merrill, 1975] p 285.

There is a third strand to the rope with which Sellars first binds and then strangles the excessive ambitions of empiricism. That is his methodological strategy of considering *semantic* relations among the meanings expressed by different sort of vocabulary that result from *pragmatic* dependencies relating the practices one must engage in or the abilities one must exercise in order to count as using those bits of vocabulary to express those meanings. I will call this the ‘pragmatist’ element in Sellars’s multi-front assault on empiricism. It makes a significant contribution to the early, critical portion of EPM, though Sellars does not overtly mark it, as he does the contribution of his inferential functionalism to the later, more constructive portion. The concern with what one must *do* in order to *say* various kinds of things remains implicit in what Sellars *does*, rather than explicit in what he *says about* what he *does*. As we will see, both the pragmatist and the inferentialist ideas are integral to his critique of empiricist approaches to modality and to his constructive suggestions for a more adequate treatment of modal vocabulary.

Section Two: The Inferentialist and Pragmatist Critique of Empiricism in EPM

I think the classical project of analytic philosophy in the twentieth century was to explore how the meanings expressed by some target vocabularies can be exhibited as in some sense a logical elaboration of the meanings already expressed by some base vocabularies. The conception of the desired semantic relation between vocabularies (the sense of ‘analysis’) has varied significantly within this broadly defined semantic project, including definition, paraphrase, translation, reduction in various senses, supervenience, and truth-making, to name just a few prominent candidates. I take it to be integral to the analytic philosophical project during this period that however that semantic relation is conceived, *logical* vocabulary is taken to play a special role in elaborating the base vocabulary into the target vocabulary. The distinctively twentieth-century form of *empiricism* should be understood as one of the *core programs* of this analytic project—not in the sense that every participant in the project endorsed some version of empiricism (Neurath, for instance, rejects empiricism where he sees it clashing with another core semantic program that was dearer to his heart, namely naturalism), but in the sense that even those who rejected it for some target vocabulary or other took the possibility of an empiricist analysis to be an important issue, to set a legitimate philosophical agenda.

Construed in these terms, twentieth century empiricism can be thought of as having proposed three broad kinds of empiricist base vocabularies. The most restrictive kind comprises *phenomenalist* vocabularies: those that specify how things subjectively *appear* as opposed to how they objectively are, or the not-yet-conceptualized perceptual *experiences* subjects have, or the so-far-uninterpreted sensory *given* (the data of sensation: sense data). A somewhat less restrictive genus of empiricist base vocabularies limits them to those that express *secondary qualities*, thought of as what is *directly perceived* in some less demanding sense. And a still more relaxed version of empiricism restricts its base vocabulary to the *observational* vocabulary deployed in non-inferentially elicited perceptual reports of observable states of affairs. Typical target vocabularies for the first, phenomenalist, class of empiricist base vocabularies include those expressing empirical claims about how things really or *objectively* are—that is, those expressing the applicability of any objective empirical concepts. Typical target vocabularies for secondary-quality empiricism include any that specify *primary qualities* or the applicability of concepts that are not response-dependent. And typical target vocabularies for observational vocabulary empiricism include *theoretical* vocabulary. All species of empiricism are concerned with the possibility of underwrite semantics of the modal vocabulary used to express laws of nature, probabilistic vocabulary, normative vocabulary, and others sophisticated vocabularies of independent philosophical interest. The standard empiricist alternatives are either to show how a given target vocabulary can be semantically elaborated from the favored empiricist base vocabulary, on the one hand, or to show how to live with a local skepticism about its ultimate semantic intelligibility, on the other.

At the center of Sellars’s critique of empiricism in EPM is an argument against the weakest, least committive, observational, version of empiricism (a critique that then carries over, *mutatis mutandis*, to the more demanding versions). That argument depends on both his inferential-functionalist semantics and on his pragmatism. Its fundamental strategy is to show that the proposed empiricist base vocabulary is not *pragmatically* autonomous, and hence not *semantically* autonomous. Observational vocabulary is not a vocabulary one could use though one used no other. Non-inferential reports of the results of observation do not form an autonomous stratum of language. In particular, when we look at what one must *do* to count as making a non-inferential report, we see that that is not a practice one could engage in except in the context

of *inferential* practices of using those observations as *premises* from which to draw inferential *conclusions*, as *reasons* for making judgments and undertaking commitments that are *not* themselves observations. The contribution to this argument of Sellars's inferential functionalism about semantics lies in underwriting the claim that for *any* judgment, claim, or belief to be *cognitively*, *conceptually*, or *epistemically* significant, for it to be a potential bit of *knowledge* or *evidence*, to be a *sapient* state or status, it must be able to play a distinctive *role in reasoning*: it must be able to serve as a *reason for* further judgments, claims, or beliefs, hence as a *premise* from which they can be *inferred*. That role in reasoning, in particular, what those judgments, claims, or beliefs can serve as reasons or evidence *for*, is an essential, and not just an accidental component of their having the semantic content that they do. And that means that one cannot count as understanding, grasping, or applying concepts *non-inferentially* in observation unless one can also deploy them at least as premises in *inferences* to conclusions that do *not*, for that very reason, count as *non-inferential* applications of concepts. Nor, for the same reason, can any discursive practice consist entirely of non-inferentially acquiring *premises*, without any corresponding practice of drawing *conclusions*. So non-inferential, observational uses of concepts do not constitute an autonomous discursive practice: a language-game one could play though one played no other. And this conclusion about the pragmatic dependence of observational uses of vocabulary on inferential ones holds no matter what the subject-matter of those observations is: whether it is observable features of the external environment, how things merely appear to a subject, or the current contents of one's own mind.

Here the pragmatist concern with what one must *do* in order to be able to *say* (or think) something combines with semantic inferentialist-functionalism about conceptual content to argue that the proposed empiricist base vocabulary is not pragmatically autonomous—since one must be able to make claims inferentially in order to count as making any non-inferentially. If that is so, then potentially risky inferential moves cannot be seen as an in-principle optional superstructure erected on a semantically autonomous base of things directly known through observation.

Although this is his most general and most powerful argument, Sellars does not limit himself to it in arguing against the substantially more committive forms of empiricism that insist on phenomenalist base vocabularies. In addition, he develops a constructive account of the relations between (at least one principle species of) phenomenalist vocabulary and objective vocabulary that depends on pragmatic dependences between what one must *do* in order to deploy each kind, to argue once again that the proposed empiricist base vocabulary does not form a semantically autonomous stratum of the language. This is his account of the relation between 'looks'-talk and 'is'-talk.

It develops out of his positive account of what one must *do* in order to use vocabulary observationally. To apply the concept green non-inferentially one must be able to do at least two sorts of things. First, one must be able reliably to respond differentially to the visible presence of green things. This is what blind and color-blind language-users lack, but non-language-using pigeons and parrots possess. Second, one must be able to exercise that capacity by reliably responding differentially to the visible presence of green things by applying the *concept green*. So one must possess, grasp, or understand that concept. "Grasp of a concept is mastery of the use of a word," Sellars says, and his inferential functionalism dictates that this must include the *inferential* use of the word: knowing at least something about what follows from and is evidence for or against something's being green. This the blind or color-blind language-user has, and the pigeon and parrot do not. Only the performances of the former can have the pragmatic significance of taking up a stand in the space of reasons, of committing themselves to something that has a *conceptual*, that is, inferentially articulated, content.

The point of Sellars's parable of John in the tie shop is to persuade us that the home language-game of the 'looks' or 'seems' vocabulary that expresses how things merely appear to us, without undertaking any commitment to how they actually are, is one that is pragmatically parasitic on the practice of making in-principle risky reports of how things objectively are. For what one must *do* in order to count as saying how things merely *look*, Sellars claims, is to *evince* the reliable differential disposition to respond to something by claiming that it is green, while *withholding* the endorsement of that claim (because of one's collateral beliefs about the situation and one's reliability in it). If that is what one is doing in making a 'looks'-claim, then one cannot be wrong about it in the same way one can about an 'is'-claim, because one has withheld the principal commitment rather than undertaking it. And it follows that phenomenalist 'looks'-talk, which expresses how things merely appear, without further commitment to how things actually are, is not an autonomous discursive practice—not a language-game one could play though one played no other—but is in fact pragmatically parasitic on objective 'is'-talk.

My point in rehearsing this familiar argument is to emphasize the role played both by Sellars's pragmatist emphasis on what one must be able to *do* in order count as *saying* various kinds of thing—*using* vocabulary so as to express certain kinds of meanings—and by his inferentialist-functionalist insistence that the role some vocabulary plays in *reasoning* makes an essential contribution to its semantic content. Although Sellars does not go on to make this argument, the way these two lines of thought conspire to undermine the semantic autonomy of candidate empiricist base vocabularies provides a template for a parallel objection to secondary-quality empiricism. For at least a necessary condition on anything's being a secondary-quality concept is that it have an observational role that supports the introduction of corresponding 'looks'-talk, so that mastery of that 'looks'-talk can be taken to be essential to mastery of the concept—as 'looks-green' arguably is for mastery of the concept green, but 'looks-square' is *not* for mastery of the concept square. What would be needed to fill in the argument against secondary-quality empiricism via the non-autonomy of its proposed base vocabulary, would be an argument that nothing could count as mastering a vocabulary consisting entirely of expressions of this sort, apart from all inferential connections to primary-quality concepts that did not have this structure.

Section Three: A Tension within Empiricism about Modality

Thus far I have confined myself to offering a general characterization of anti-empiricist arguments that appear in "Empiricism and the Philosophy of Mind." None of them involve empiricism's treatment of modality. Now I want to put those arguments in a somewhat different frame, by conjoining them with one that is presented elsewhere, and which *does* turn on the significance of modal concepts. The previous arguments concerned the suitability of some vocabulary to serve as the *base* vocabulary of an empiricist analysis—since plausible motivations for caring about such an analysis typically require that it be semantically autonomous. This one turns on the criteria of adequacy of the analysis itself. My remarks in this section concern Sellars's arguments in his essay "Phenomenalism," which can be regarded as a kind of companion piece to EPM. (Later I will discuss another contemporary essay that I think should be thought of as yoked together with these two in a troika.) The first, modal, point is one that Sellars registers there, but does not linger on—his principal concern being rather with a second point, concerning another aspect of the vocabulary in which phenomenalist analyses would have to be couched. But given my purposes here, I want to make a bit more of the modal point than he does.

The basic idea of a phenomenalist-empiricist semantic analysis of ordinary objective vocabulary is that the expressive work done by talk of mind-independent objects and their properties and relations can be done by talk of *patterns* in, *regularities* of, or *generalizations* concerning sense experiences characterized in a phenomenalist vocabulary. Saying that the curved red surface I am experiencing is an experience *of* an apple that has parts I am *not* experiencing—a similarly bulgy, red back and a white interior, for instance—is properly understood as saying something about what I *would* experience if I turned it around or cut it open. That it continued to exist in the kitchen when I left the room is a matter of what I *would* have experienced *had* I returned. The first, obvious, observation is that an account of objective reality in terms of the *powers* of circumstances to *produce*, or my *dispositions* to *have*, sensations, experiences, beings-appeared-to and so on essentially involves *modal* concepts. The patterns, regularities, or generalizations in subjective appearances that are supposed to constitute objective realities are modally robust, counterfactual-supporting patterns, regularities, or generalizations. Talk of what I actually *do* experience will not by itself underwrite claims about unexperienced spatial or temporal parts of empirical objects. Twentieth-century logical empiricism promised to advance beyond traditional empiricism because it could call on the full expressive resources of *logical* vocabulary to use as the 'glue' sticking sensory experiences together so as to construct simulacra of external objects. But *extensional* logical vocabulary is not nearly expressively powerful enough for the phenomenalist version of the empiricist project. So the phenomenalist conditional 'terminating judgments' into an infinite set of which C. I. Lewis proposes (in his *Analysis of Knowledge and Valuation*) to translate the 'non-terminating judgments' of ordinary objective empirical discourse have to use his modal notion of *strict* or necessary implication. And similar points could be made about other phenomenalist reductionists such as Ayer. The consequence of this observation to which I want to draw attention is that one cannot use such a strategy in one's phenomenalist empiricist analysis, translation, or reduction of objective talk *and* at the same time be a Humean skeptic about what *modal* vocabulary expresses. Essential features of the only remotely plausible *constructive* strategy of phenomenalist

empiricism are simply incompatible with the most prominent *skeptical* consequences about modal concepts characteristically drawn both by traditional and twentieth-century logicist empiricism.

This is a powerful argument. Sellars's principal concern in his essay "Phenomenalism," however, is with a subsequent point. The conditionals codifying the patterns, regularities, or generalizations concerning sense experience that correspond to judgments about how things objectively are must not only be subjunctive, counterfactually robust conditionals, but in order to have any hope of being materially adequate (getting the truth-conditions even approximately correct) their *antecedents* must themselves be expressed in *objective* vocabulary, *not* in *phenomenalist* vocabulary. What is true (enough) is that if I were *actually* to turn the apple around, cut it open, or return to its vicinity in the kitchen I *would* have certain sense experiences. It is *not* in general true that if I merely *seem* to do those things I am guaranteed to have the corresponding experiences. For, phrased in such phenomenalist terms, the antecedent is satisfied in cases of imagination, visual illusion, dreaming, hallucination and so on that are precisely those *not* bound by the supposedly object-constituting rules and regularities. As Sellars summarizes the point:

To claim that the relationship between the framework of sense contents and that of physical objects can be construed on the [phenomenalist] model is to commit oneself to the idea that there are inductively confirmable generalizations about sense contents which are 'in principle' capable of being formulated without the use of the language of physical things....[T]his idea is a mistake.

It is a mistake because:

[T]he very selection of the complex patterns of actual sense contents in our past experiences which are to serve as the antecedents of the generalizations in question presuppose our common sense knowledge of ourselves as perceivers, of the specific physical environment in which we do our perceiving and of the general principles which correlate the occurrence of sensations with bodily and environmental conditions. We select those patterns which go with our being in a certain perceptual relation to a particular object of a certain quality, where we know that being in this relation to an object of that quality normally eventuates in our having the sense content referred to in the consequent.

This argument then makes evident

the logical dependence of the framework of private sense contents on the public, inter-subjective, logical space of persons and physical things.

So the phenomenalist vocabulary is not autonomous. It is not a language-game one can play though one plays no other. In particular, the uses of it that might plausibly fulfill many of the same pragmatic functions as ordinary objective empirical talk themselves presuppose the ability to deploy such objective vocabulary.

As Sellars points out, the lessons learned from pressing on the phenomenalist version of empiricism apply more generally. In particular, they apply to the more liberal version of empiricism whose base vocabulary is observational, including observations of enduring empirical objects, and whose target vocabulary is theoretical vocabulary. To begin with, if talk of theoretical entities is to be translated into or replaced by talk of patterns in, regularities of, or generalizations about observable entities, they must be *lawlike*, *counterfactual*-supporting regularities and generalizations. They must permit inferences to what one *would* observe if one *were* to find oneself in specified circumstances, or to prepare the apparatus in a certain way. For, once again, the patterns, regularities, or generalizations about observations the assertion of which an instrumentalist empiricist might with some initial plausibility take to have the same pragmatic effect as (to be doing the same thing one is doing in) deploying theoretical vocabulary must reach beyond the parochial, merely autobiographically significant contingencies of what subjects happen actually to observe. The theory is that electrical currents cause magnetic fields regardless of the presence of suitable measuring devices. And that can only be made out in terms of what is *observable*, that is *could* be observed, not just what *is* observed. And that is to say that the instrumentalist-observational form of empiricism is also incompatible with Humean-Quinean skepticism about the intelligibility of what is expressed by alethic modal vocabulary.

And an analogue of the second argument against phenomenalist forms of empiricism also applies to instrumentalist forms. For, once again, the *antecedents* of the counterfactual conditionals specifying what

could or *would* have been observed *if* certain conditions *had* obtained or certain operations *were* performed cannot themselves be formulated in purely observational terms. The meter-needle *would* have been observably displaced if I had connected the terminals of a volt-ohmmeter to the wire, but that something *is* a VOM is *not* itself a fact restatable in purely observational terms. Even leaving apart the fact that it is a *functional* characterization not equivalent to any specification in purely *physical* terms, a description of the construction of some particular kind of VOM is still going to help itself to notions such as being made of copper, or being an electrical insulator (another bit of vocabulary that is both functional and theoretical). To satisfy the semantic ambitions of the instrumentalist it is not enough to associate each theoretical claim with a set of jointly pragmatically equivalent counterfactual-supporting conditionals whose *consequents* are couched wholly in observational vocabulary. All the theoretical terms appearing in the *antecedents* of those conditionals must be similarly replaced. No instrumentalist reduction of any actual theoretical claim has ever been suggested that even attempts to satisfy this condition.

Though Sellars does not, and I will not, pursue the matter, one expects that corresponding arguments will go through, *mutatis mutandis*, also for the kind of empiricism that seeks to understand the use of primary-quality vocabulary wholly in terms of the use of secondary-quality vocabulary. What we mean by talk of primary qualities will have to be cashed out in terms of its *powers* to produce, or our *dispositions* to perceive, secondary qualities—that is, in terms of modally robust, counterfactual-supporting generalizations. And it will be a challenge to specify the antecedents of a materially adequate set of such conditionals wholly in the official secondary-quality vocabulary.

Section Four: A Direct Argument Against Empiricist Skepticism about Modality

The arguments I have considered so far set limits to the semantic ambitions of phenomenalist and instrumentalist forms of analytic empiricism, first by focusing on the *pragmatic* preconditions of the required semantic autonomy of the proposed empiricist base vocabularies, and second by looking in more detail at the specific sorts of *inferential* patterns in the base vocabulary in terms of which it is proposed to reconstruct the circumstances and consequences of application of items in the various target vocabularies. Here it was observed that the material adequacy of such reconstructions seems to require the ineliminable involvement of terms from the target vocabulary, not only on the right side, but also on the left side of any such reconstruction—in the *definiens* as well as in the *definiendum*. Modality plays a role in these arguments only because the material adequacy of the reconstruction also turns out to require appeal to counterfactually robust inferences in the base vocabulary. Insofar as that is so, the *constructive* semantic projects of the phenomenalist, instrumentalist, and secondary-quality forms of empiricism are at odds with the local semantic skepticism about what is expressed by alethic modal vocabulary that has always been a characteristic cardinal *critical* consequence of empiricist approaches to semantics, as epitomized for its traditional phase by Hume and for its logicist phase by Quine.

In another massive, pathbreaking essay of this period, “Counterfactuals, Dispositions, and the Causal Modalities”² (completed in February of 1957), Sellars argues directly against this empiricist treatment of modality, completing what then becomes visible as a two-pronged attack on the principle contentions and projects of empiricism, only the opening salvos of which were fired in “Empiricism and the Philosophy of Mind.”³ His principal target here is the “tendency to assimilate all discourse to describing,” which he takes to be primarily “responsible for the prevalence in the empiricist tradition of ‘nothing-but-ism’ in its various

² Pp. 225-308 of *Minnesota Studies in the Philosophy of Science*, Vol. II, ed. by H. Feigl, M. Scriven, and G. Maxwell, (University of Minnesota Press; Minneapolis, MN: 1957). Henceforth CDCM.

³ As in EPM (and even, though to a lesser extent, in “Phenomenalism”), in this essay Sellars describes himself not as denying empiricism, but rather as correcting it, protecting its core insights from the damage done by their over-extension. But he also makes it clear that the result of such rectification is a Kantian view that gives equal weight to rationalist insights, when they are suitably reconstructed. So for instance he says:

It is my purpose to argue that the core truth of Hume’s philosophy of causation is not only compatible with, but absurd without, *ungrudging* recognition of those features of causal discourse as a mode of rational discourse on which the ‘metaphysical rationalists’ laid such stress, but also mis-assimilated to describing.” [Section 82]

And the final sentence of the essay invokes the “profound truth” of Kant’s conception of reason, “which empiricism has tended to distort.”

forms (emotivism, philosophical behaviorism, phenomenalism)...”⁴ The form Sellars addresses in this essay is the Humean one that can find in statements of laws of nature, expressed in alethic modal vocabulary that lets us say what is and is not necessary and possible, “nothing but” expressions of matter-of-factual regularities or constant conjunctions (though he claims explicitly that considerations corresponding to those he raises for causal modalities are intended to apply to logical and deontological modalities as well⁵). His arguments are directed against the view that holds modal vocabulary semantically unintelligible, on grounds of inability to specify what it is saying about what the world is like, how it is describing things as being, insofar as by using it we are asserting something that goes beyond endorsing the existence of non-modally characterizable universal generalizations.

Hume found that even his best understanding of actual observable empirical *facts* did not yield an understanding of *rules* relating or otherwise governing them. Those facts did not settle which of the things that *actually* happened *had* to happen (given others), that is, were (at least conditionally) *necessary*, and which of the things that did *not* happen nonetheless were *possible* (not ruled out by laws concerning what did happen). The issue here concerns the justifiability and intelligibility of a certain kind of *inference*: modally robust, counterfactual-supporting inferences, of the kind made explicit by the use of modal vocabulary. Hume (and, following him, Quine) took it that epistemologically and semantically fastidious philosophers face a stark choice: either show how to explain modal vocabulary—the circumstances of application that justify the distinctive counterfactual-supporting inferential consequences of application—in nonmodal terms, or show how to live without it, to do what we need to do in science without making such arcane and occult supradescriptive commitments.

This demand was always the greatest source of tension between empiricism and naturalism, especially the scientific naturalism that Sellars epitomized in the slogan: “Science is the measure of all things, of those that are, that they are, and of those that are not, that they are not.” For modern mathematized natural science shorn of concern with laws, counterfactuals, and dispositions—in short of what is expressed by alethic modal vocabulary—is less than an impotent Samson, it is an inert, unrecognizable, fragmentary remnant of a once-vital enterprise. Sellars’s general recommendation for resolving this painful tension (felt particularly acutely by, and one of the principal issues dividing, the members of the Vienna circle) is to relax the exclusivism and rigorism he traces to empiricism’s semantic descriptivism:

[O]nce the tautology ‘The world is described by descriptive concepts’ is freed from the idea that the business of all non-logical concepts is to describe, the way is clear to an *ungrudging* recognition that many expressions which empiricists have relegated to second-class citizenship in discourse are not *inferior*, just *different*.⁶

Sensitized as we now are by Sellars’s diagnoses of *semantic autonomy* claims as essential to various empiricist constructive and reconstructive projects, both in EPM and in the “Phenomenalism” essay, and familiar as we now are with his criticisms of them based on the inferentially articulated *doings* required to use or deploy various candidate base vocabularies, it should come as no surprise that his objections to critical empiricist suspicions of and hostility towards modality follow the same pattern. For the Humean-Quinean empiricist semantic challenge to the legitimacy of modal vocabulary is predicated on the idea of an independently and antecedently intelligible stratum of empirical discourse that is purely descriptive and involves no modal commitments, as a semantically autonomous background and model with which the credentials of modal discourse can then be invidiously compared.

In this case, as in the others, the argument turns both on the *pragmatism* that looks to what one is doing in deploying the candidate base vocabulary—here “purely descriptive” vocabulary—and on the nature of the *inferential* articulation of that vocabulary necessary for such uses to play the expressive role characteristic of that vocabulary. The argument in this case is subtler and more complex than the others however. For one thing, I take it that Sellars does *not* deny the intelligibility-in-principle of purely descriptive discourse that contains no explicitly modal vocabulary.⁷ For another, there are special

⁴ CDCM, Section 103.

⁵ *Ibid.*

⁶ CDCM, Section 79.

⁷ Sellars is, frustratingly but characteristically, not explicit about his attitude towards the pragmatic autonomy in principle of such purely descriptive discourse. He says:

The idea that the world can, in principle, be so described that the description contains no modal expression is of a piece with the idea that the world can, in principle, be so described that the

difficulties involved in, and corresponding delicacies required for, working out the general pragmatist-inferentialist strategy so as to apply it to this case, by specifying the relation between the expressive role distinctive of modal vocabulary, on the one hand, and what one is *doing* (in particular, the inferential commitments one is undertaking) in using ordinary, non-modal, descriptive vocabulary itself, on the other.

The pragmatic dependency relation that lies at the base of Sellars's argument is the fact that:

...although describing and explaining (predicting, retrodicting, understanding) are *distinguishable*, they are also, in an important sense, *inseparable*. It is only because the expressions in terms of which we describe objects, even such basic expressions as words for perceptible characteristics of molar objects, locate these objects in a space of implications, that they describe at all, rather than merely label. The descriptive and explanatory resources of language advance hand in hand....⁸

Descriptive uses of vocabulary presuppose an inferentially articulated "space of implications," within which some descriptions show up as reasons for or explanations of others. Understanding those descriptions requires placing them in such a space. This pragmatist claim about what else one must be able to *do*—namely, *infer*, *explain*, treat one claim as a *reason* for another—in order for what one is doing to count as *describing* connects to the use of *modal* vocabulary via the principle that:

To make first hand use of these [modal] expressions is to be about the business of explaining a state of affairs, or justifying an assertion.⁹

That is, what one is *doing* in *using* modal expressions is explaining, justifying, or endorsing an inference. So what one is doing in saying that As are *necessarily* Bs is endorsing the inference from anything's being an A to its being a B.

The first sort of difficulty I alluded to above stems from the fact that there are other ways of endorsing such a pattern of inference besides *saying that* all As are necessarily Bs. One's endorsement may be *implicit* in other things one *does*, the reasoning one engages in and approves, rather than *explicit* in what one *says*. So from the fact (assuming, as I shall, that it is a fact) that the activity of describing is part of an indissoluble pragmatic package that includes endorsing inferences and the fact that what one is doing in making a modal claim is endorsing an inference, it does not at all follow that there can be no use of descriptive vocabulary apart from the use of modal vocabulary. The second difficulty stems from the fact that although Sellars may be right that what one is *doing* in making a modal claim is endorsing a pattern of inference, it is clear that one is not thereby *saying that* an inference is good. When I say "Pure copper necessarily conducts electricity," and thereby unrestrictedly endorse inferences from anything's being pure copper to its conducting electricity, I have nevertheless *said* nothing about any inferences, explanations, justifications, or implications—indeed, have said something that could be true even if there had never been any inferences or inferencers to endorse them, hence no describers or discursive practitioners at all.¹⁰ These two observations set the principal criteria of adequacy both for Sellars's positive working-out of the pragmatist-inferentialist treatment of modal vocabulary, and for his argument that the purely descriptive

description contains no prescriptive expression. For what is being called to mind is the ideal of statement of 'everything that is the case' which, however, serves *through and through only* the purpose of stating what is the case. And it is a logical truth that such a description, however many modal expressions might properly be used in *arriving at* it or in *justifying* it, or in showing the *relevance* of one of its components to another, could contain no modal expression. [Section 80]

Sellars's view about this ideal is complex: there is sense in which it is intelligible, and a sense in which it is not. Such a discourse would be unreflective and unself-conscious in a way ours is not. For reasons that will emerge, it would belong to what at the end of the essay he calls the stage of human language "when linguistic changes had *causes*, but not *reasons*, [before] man acquired the ability to reason about reasons." [Section 108].

⁸ CDCM Section 108.

⁹ CDCM Section 80.

¹⁰ Sellars connects this obvious fact with the observation that:

Idealism is notorious for the fallacy of concluding that because there must be minds in the world in order for *us* to have reason to make statements about the world, therefore there is no sense to the idea of a world which does not include minds. [CDCM Section 101]

base vocabulary invoked by the empiricist critic of the semantic credentials of modal vocabulary lacks the sort of discursive autonomy the empiricist criticism presupposes and requires.

Sellars's central rhetorical strategy in this essay is to address the issue of what is expressed by modal claims about necessary connections by offering:

...a sympathetic reconstruction of the controversy in the form of a debate between a Mr. C (for Constant Conjunction) and a Mr. E (for Entailment) who develop and qualify their views in such a way as to bring them to the growing edge of the problem.¹¹

Officially, he is even-handed in his treatment of the vices and virtues of the empiricist, who denies that the use of modal vocabulary can express any legitimate semantic content beyond that expressed by a descriptive, extensional universal generalization, and the rationalist, who understands that content in terms of entailments expressing rules of reasoning. In fact, however, as becomes clear when he launches into his own account, he is mainly concerned to develop a version of the rationalist account. As the second half of the essay develops, Sellars's marks his abandonment of the disinterested pose by an uncharacteristically explicit expository shift:

It is now high time that I dropped the persona of Mr. E, and set about replying to the challenge with which Mr. C ended his first critique of the entailment theory.¹²

Doing that requires careful investigation of the differences between and relations among four different sorts of item:

- Practical endorsement of the propriety of an inference from things being A to their being B;
- The explicit statement that one may infer the applicability of 'B' from the applicability of 'A';
- The statement that A physically entails B;
- The statement that As are necessarily Bs.

The first is the sort of thing Sellars takes to be pragmatically presupposed by the activity of describing, that is, deploying descriptive vocabulary. The second fails to capture such practical endorsements, because of the possibility of asserting such statements regarding the *expressions* 'A' and 'B' without understanding what they express.¹³

The third sort of statement expresses Mr. E's initial stab at an analysis of the fourth. It is the answer to the question: what sort of entailment is it that modal statements are supposed to express?:

Mr. E has a ready answer. ...it might...be called 'natural' or 'physical' entailment, for while any entailment is a logical relation, we can distinguish within the broad class of entailments between those which are, and those which are not, a function of the specific empirical contents between which they obtain. The latter are investigated by general or formal logic (and pure mathematics). Empirical science, on the other hand, to the extent that it is a search for *laws*, is the search for entailments of the former kind. (Putative) success in this search finds its expression in statements of the form 'It is (inductively) probable that A physically entails B.'¹⁴

¹¹ CCDM Introduction.

¹² CDCM Section 85. In fact, Sellars's 'defense' of Mr. C (see the passage from Section 82 quoted in note 3 above) consists of showing what concessions he needs to make to Mr. E. This proceeds first by Mr. C's qualification that "'A causes B' says that (x)[Ax→Bx] and implies that he latter is asserted on inductive grounds" [62], followed by the necessity of conceiving "of induction as establishing principles *in accordance with which* we reason, rather than as major premises *from which* we reason." [83] As will appear, the former concession, introducing the notion of what is contextual implied by contrast to what is explicitly said, is then dialectically made available to be pressed into service by Mr. E. This bit of dialectic is a pretty rhetorical flourish on Sellars's part, but I doubt that in the end it reflects any deep feature of the confrontation between the empiricist and rationalist approaches to modality.

¹³ As Sellars says:

But one can know that Turks, for example, ought to withdraw '...' when they commit themselves to '---' without knowing the language, whereas the statement that 'p entails q' contextually implies that the speaker not only knows the language to which 'p' and 'q' belong, but, in particular, knows how to use 'p' and 'q' themselves. [CDCM Section 81]

¹⁴ CDCM Section 56.

The virtue of statements like “A physically entails B” is that they do plausibly codify the practical endorsement of an inference that is implicit in what one does in the form of something one can explicitly say, without bringing in irrelevant commitments concerning particular expressions, the activity of inferring, or discursive practitioners. The remaining difficulty is that they seem plainly not to have the same content, not to say the same thing, as explicitly modal statements of objective necessity.

Sellars’s response to this problem is to acknowledge that modal statements do not *say that* some entailment holds, but to distinguish between what is *said* by using a bit of vocabulary and what is ‘*contextually implied*’ by doing so. Sellars says very little about this latter notion, even though it bears the full weight of his proposed emendation of the rationalist account. It is recognizably the same distinction he had appealed to earlier, in “Inference and Meaning”, as the distinction between what making a statement *says* and what it *conveys*. There his example is that in asserting “The weather is fine today,” I *say* that the weather is fine today, but *convey* that I *believe* that it is fine.¹⁵ That otherwise uninterpreted example suggests to me that what Sellars has in mind is the distinction between *semantic* and *pragmatic* inferences. That is the distinction between inferences underwritten by the *contents* of what is *said* or asserted, on the one hand, and inferences underwritten by what one is *doing* in saying them, on the other. The inference from “The weather is fine,” to “It is not raining,” is of the first sort; the inference from my asserting “The weather is fine,” to “Brandom believes the weather is fine,” is of the second sort. Inferences of these two kinds may generally be distinguished by the Frege-Geach embedding test: look to see whether those who make the inference in question also endorse the corresponding conditional. “If the weather is fine, then it is not raining,” is generally true, while “If the weather is fine, then Brandom believes it is fine,” is not generally true. (Compare the inference from my *saying* “That is an ugly tie you are wearing,” to “Bob is annoyed with me.”)

If that is in fact the distinction Sellars is after, then it seems to me that the view he is expounding and defending can be put less paradoxically if we don’t take a detour through entailment statements, but concern ourselves directly with the relation between the endorsement of patterns of inference and modal statements. The underlying rationalist insight is a pragmatist-inferentialist one: what one is *doing* in making a modal claim is endorsing a pattern of inference. Modal vocabulary makes possible new kinds of *sayings* that have the *pragmatic effect* of endorsing inferences. To say that is not yet to say what they *say*, it is only to say what one is *doing by* saying them. But it does settle the *pragmatic significance* of such modal claims, in the sense of their appropriate circumstances and consequences of application.¹⁶ If one practically endorses the pattern of inference that treats classifying anything at all as an A as sufficient

¹⁵ Sellars, “Inference and Meaning”, p. 280/332 in J. Sicha (ed.) *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars* [Ridgeview Publishing Company, Reseda CA, 1980]—hereafter *PPPW*.

¹⁶ It is the attempt to specify this peculiar and distinctive sort of pragmatically mediated relation between vocabularies that leads Sellars to say things like:

It is sometimes thought that modal statements do not describe states of affairs in the world, because they are *really* metalinguistic. This won’t do at all if it is meant that instead of describing states of affairs in the world, they describe linguistic habits. It is more plausible if it is meant that statements involving modal terms have the force of *prescriptive* statements about the use of certain expressions in the object language. Yet there is more than one way of to ‘*have the force of*’ a statement, and failure to distinguish between them may snowball into a serious confusion as wider implications are drawn. [CDCM Section 81]

and

Shall we say that modal expressions are metalinguistic? Neither a simple ‘yes’ nor a simple ‘no’ will do. As a matter of fact, once the above considerations are given their proper weight, it is possible to acknowledge that the idea that they are metalinguistic in character oversimplifies a fundamental insight. For our present purposes, it is sufficient to say that the claim that modal expressions are ‘in the metalanguage’ is not too misleading if the peculiar force of the expressions which occur alongside them (represented by the ‘p’ and the ‘q’ of our example) is recognized, in particular, that they have ‘straightforward’ translation into other languages, and if it is also recognized that they belong not only ‘in the metalanguage’, but in discourse about *thoughts* and *concepts* as well. [CDCM Section 82]

and

We must here, as elsewhere, draw a distinction between what we are committed to concerning the world by virtue of the fact that we have reason to make a certain assertion, and the force, in a narrower sense, of the assertion itself. [CDCM Section 101]

grounds (“all on its own”, as Sellars says, in order to capture the way the pattern of inferences in question is counterfactually robust) for concluding that it is a B, then one is committed to the claim that all As are necessarily Bs. And commitment to that claim is commitment to practically ratify that pattern of inference. Assuming, as Sellars has claimed, that using ordinary, non-modal, descriptive vocabulary requires practically endorsing such patterns of inference (“situating descriptions in a space of implications”), that means that anyone who has the practical ability to deploy “purely descriptive” vocabulary already knows how to do everything he needs to know how to do to deploy modal vocabulary as well. He need not actually do so, since practically undertaking those inferential commitments does not require that one have available a language with vocabulary permitting one to *do* that by *saying* something. But *all* a practitioner lacks in such a circumstance is the *words* to hook up to discriminative and responsive abilities he already possesses. In this precise sense, the ability to deploy modal vocabulary is *practically implicit* in the ability to deploy non-modal descriptive vocabulary.

Sellars has claimed that the activity of describing is unintelligible except as part of a pragmatic package that includes also not just the making of inferences, but the making of *counterfactually robust* inferences: the sort of inferences involved in *explanation*, and licensed by explicitly modal statements of *laws*. He sums up the claim admirably in the title of another one of his early papers: “Concepts as Involving Laws, and Inconceivable without Them.” Grasp of a concept is mastery of the use of a word, Sellars says. And that use includes not only sorting inferences (however fallibly and incompletely) into materially good and materially bad ones, but also, among the ones one takes to be materially good, to distinguish (however fallibly and incompletely) between counterfactual circumstances under which they do, and counterfactual circumstances under which they do not, *remain* good. Part of taking an inference to be materially good is having a view about which possible additional collateral premises or auxiliary hypotheses would, and which would not, infirm it. Chestnut trees produce chestnuts—unless they are immature, or blighted. Dry, well-made matches strike—unless there is no oxygen. The hungry lioness would still chase the antelope if it were Tuesday or the beetle on the distant tree crawled slightly further up the branch, but not if lioness’s heart were to stop beating. The point is not that there is any particular set of such discriminations that one must be able to make in order to count as deploying the concepts involved. It is that if one can make *no* such practical assessments of the counterfactual robustness of material inferences involving those concepts, one could not count as having mastered them.

Against the background of this pragmatist-inferentialist claim about what is involved in the ordinary descriptive use of concepts, Sellars’s claim, as I am reading him, is that explicitly modal “lawlike” statements are statements that one is committed or entitled to whenever one is committed or entitled to endorse such patterns of counterfactually robust inference, and commitment or entitlement to which in their turn commit or entitle one to the corresponding patterns of inference. Saying that about them settles what one needs to *do* to *use* such modal statements. It does *not* say how one is thereby *describing* the world as being when one does. It does not, in particular, *describe* a pattern of inference as good (though that saying does, in its own distinctive way, *express endorsement* of such a pattern). It does not do those things for the simple reason that the use of modal expressions is *not* in the first instance *descriptive*.¹⁷ It codifies explicitly, in the form of a statement, a feature of the use of descriptive expressions that is indissolubly bound up with, but not identical to, their descriptive use. Nonetheless, in knowing how to use vocabulary descriptively, one knows how to do everything one needs to know how to do in order to use modal vocabulary. And that is enough to show that one cannot actually be in the Humean predicament presupposed by the empiricist challenge to the intelligibility of modal vocabulary. For one cannot know how to use vocabulary in matter-of-factual descriptions (“The cat is on the mat,”) and not have any grip on how to use modal, counterfactual, and dispositional vocabulary (“It is necessary for live cats to breathe,,” “The cat could still be on the mat if the mat were a slightly different shade of blue, but not if it turned into soup,,” “The cat would leave the mat if she saw a mouse,,”). Although *explicitly* modal *vocabulary* is an in-principle optional superstructure on practices of deploying descriptive vocabulary, what it expresses cannot be mysterious in principle to those who can engage in those base-level practices.

In taking this line, Sellars quite properly sees himself as reviving a central idea of Kant’s. The ability to use empirical descriptive terms such as ‘mass’, ‘rigid’, and ‘green’ already presupposes grasp of the kind of

¹⁷ Sellars says:

[Mr. E.] conceives of induction as establishing principles *in accordance with which* we reason, rather than as major premises *from which* we reason. [CDCM Section 83]

properties and relations made explicit by modal vocabulary. It is this insight that leads Kant to the idea of ‘pure’ concepts or ‘categories’, including the alethic modal concepts of necessity and possibility that articulate causal laws, which must be available *a priori* because and in the sense that the ability to deploy them is presupposed by the ability to deploy ordinary empirical descriptive concepts. The categories, including modality, are concepts that make explicit what is implicit in the empirical, descriptive use of any concepts at all. Though the details of *which* laws, the statements of which express counterfactually robust patterns of inference, actually obtain is an empirical one, *that* empirical descriptions are related by *rules* in the form of laws, which do support counterfactually robust inferences, is *not* itself an empirical matter, but a truth about the framework of empirical description. I want to call the underlying insight “the Kant-Sellars thesis about modality.” It is the claim that in being able to use non-modal, empirical-descriptive vocabulary, one already knows how to do everything one needs to know how to do in order to deploy modal vocabulary, which accordingly can be understood as making explicit structural features that are always already implicit in what one *does* in describing.

Section Five: Conclusion

Articulating and justifying his version of the Kant-Sellars thesis about modality is Sellars’s constructive response to the empiricist tradition’s “nothing-but-ism” about modality: its demand that what is expressed by modal claims either be shown to be expressible in non-modal terms, or be dispensed with entirely by semantically fastidious philosophers and scientists. This complements and completes his demonstration, in the “Phenomenalism” essay, that this critical consequence of an over-ambitious empiricism is in any case incompatible with any constructive empiricist effort to reconstruct or replace the use of target vocabularies such as objective-descriptive vocabulary, primary-quality vocabulary, and theoretical vocabulary in terms of the favored empiricist base vocabularies, if that effort is subject to even the most minimal criteria of material adequacy. Together, these arguments show what Sellars eventually made of his early intuition that the soft underbelly of empiricism, in both its traditional and its twentieth-century logistical form, is its treatment of modality.

My overall aim in this essay has been to place the arguments against empiricism presented in the first half of “Empiricism and the Philosophy of Mind” in the larger context opened up by laying them alongside the further battery of arguments aimed at the same target that derive from consideration of that tradition’s views about modality. And I have been concerned to show that the methodological strategies that guide all of these discussions are Sellars’s *pragmatist* insistence on looking at what one must be able to *do* in order to deploy empirical descriptive vocabulary, and his *rationalist* commitment to the necessary *inferential* articulation of the concepts expressed by the use of such vocabulary. I think that even fifty years on, there is still a lot of juice to be squeezed out of these ideas.

But I want to close with another, perhaps more frivolous suggestion. Every sufficiently engaged reading becomes a rewriting, and I have been offering here, *inter alia*, the outline of a different narrative strategy that Sellars could have adopted in the late 1950s. Under some such title as *The Limits of Empiricism*, he could have re-presented the material that in fact appeared first as roughly the first half of “Empiricism and the Philosophy of Mind,” and the second halves of each of “Phenomenalism” and “Counterfactuals, Dispositions, and Causal Modalities,” organized around and introduced in terms of the themes I have traced here. It is interesting to speculate about how his reception might have been different—and about where we would find ourselves today—had this been the shape of Sellars’s first book.

Brandom's five-step program for modal health

Fredrik Stjernberg

fredrik.stjernberg@liu.se

Linköping University, Sweden

Abstract: In Chapter 4 of his (2008), Robert Brandom presents an argument to show how our modal thought and thought about counterfactuals is legitimate and indispensable. I find myself in substantial agreement with much of the reasoning, but there are a few things worth discussing (I hope!). I discuss a few issues regarding the argument: First, how bad is the problem it is supposed to solve? Second, how much does the argument prove? Third, can a thoroughgoing pragmatist theory, cashing out modality in terms of commitments, get off the ground? And finally, is there a problematic circularity in the Kant-Sellars thesis?

As things stand right now, the situation regarding modals is quite unsatisfying. On the one hand, modal (and counterfactual) expressions are regularly used in the explanation of other concepts, and in the development of for instance scientific theories. On the other hand, existing theories of modality and of our knowledge of possibility and necessity (including counterfactuals) leave a lot to be desired. Thus Kit Fine and Christopher Peacocke have voiced their dissatisfaction in the following ways:

It is an oddity of current thinking about modality that it has been heavily influenced, one might even say dominated, by two extreme and highly implausible views. The first of these, associated with the name of Quine, is that modal notions are lacking in sense. ... The second of these two views, associated with the name of David Lewis, is that the possible and the actual are on an ontological par. Other possible worlds and their inhabitants are just as real as the actual world and its inhabitants; and there is no difference between them in regard either to the degree or to the kind of reality that they possess. (Fine 2005, p. 1)

And Peacocke, in his treatment of problems concerning the knowledge of modals:

The metaphysics of necessity seemingly [has] ... this distinction: that there is practically no philosophical view of the matter so extraordinary that it has not been endorsed by someone or other. (Peacocke 1999, p. 119)

Much of the problem with our knowledge of modals has its origins in an empiricist conception of what the basis for knowledge and understanding amounts to. How could we, by observing contingent facts, come to know that something *must* be the case, or that if something *were* the case, then something else *would be* the case? Observation seems only to tell us what is in fact the case, with no counterfactual frills, as it were. William Whewell presented this problem, with its obvious roots in Hume, in a forceful way. He said that experience:

can observe and record what has happened; but she cannot find, in any case, or in any accumulation of cases, any reason for what *must* happen ... To learn a proposition by experience, and to see it to be necessarily true, are two altogether distinct processes of thought ... If anyone does not clearly comprehend this distinction of necessary and contingent truths, he will not be able to go along with us in our researches into the foundations of human knowledge; nor indeed, to pursue with success any speculation on the subject. (Whewell 1840, pp. 59-61)

If our grip on the world exclusively depends on what our observations can tell us, our access to knowledge of necessity and possibility is endangered, at least if we can make good the suppressed empiricist premiss here: that there is such a thing as pure observational knowledge of the world that is untainted by modal knowledge.

In his (2008), Brandom presents an argument designed to show that *all* our knowledge, observational or not, is shot through with modal knowledge.¹ This argument has acknowledged Kantian and Sellarsian roots, but Brandom spells it out in greater detail than his predecessors. If Brandom's argument works, empiricist worries are misplaced, or alternatively, we might say that the empiricist is powerless to meet these worries on her own terms, but that a better view of empirical knowledge will leave room for our knowledge of modal and counterfactual facts. If we know anything at all, then we know at least some modals. So Whewell's problem was never something we should have worried about, since it rests on an incorrect conception of observation. Brandom's argument is interesting, because if it works, we get a new kind of support for appeals to knowledge of counterfactuals, different from attempts to provide a direct rationalist support for such knowledge.

Brandom's central argument for his claim about knowledge of modals is in five steps, with both a preliminary and a more considered conclusion. It is mainly set out in chapter 4 of Brandom (2008). The argument starts with what Brandom calls the *Kant-Sellars thesis about modality*, which is that mastery of ordinary empirical vocabulary requires that we *already* know how to use modal vocabulary (2008, pp. 96ff, see also p. 115). Grasping any claim, modal or not, already presupposes grasping some counterfactual or modal claim. This kind of claim is not only supported in Kantian circles; we can find a recent endorsement of something much like it in Timothy Williamson:

In practice, the only way for us to be cognitively equipped to deal with the actual is by being cognitively equipped to deal with a wide variety of contingencies, most of them counterfactual. (Williamson 2007, p. 137)

So understanding counterfactuals is a necessary precondition of understanding anything at all about the world. Brandom proceeds to develop this idea more systematically, in an argument in five steps. First, the argument in outline:

1. Observationality: Every discursive practice must have some vocabulary that can be used observationally (Brandom 2008, p. 106).
2. Goodness of material inference: Those who engage in discursive practices must distinguish in practice between materially good and materially bad inferences (ibid.)
3. Non-monotonicity: Material inference is in general non-monotonic. It is defeasible, and its defeasibility cannot be cancelled by some exhaustive spelling out of the possible defeaters (ibid.).
4. Justification: Many of a subject's beliefs could only be justified by exhibiting them as conclusions of material implication. A believer is "epistemically responsible" insofar as she acknowledges a commitment to being able to justify many, if not most, of her beliefs (p. 108).
5. Epistemic responsibility: To count as a discursive practitioner, one must be at least minimally epistemically responsible (p. 108).

These five steps yield a preliminary conclusion, the *updating problem*: "Every change of belief ... is *potentially* relevant to the justification of every prior belief" (p. 108). I observe changes in the world around me all the time, and if every such change is potentially relevant for any one of my prior beliefs, I am in trouble.

How are we supposed to be able to hold on to the right set of beliefs, and update successfully? Brandom argues that the only solution to the updating problem is that people who use a vocabulary already from the outset must have an idea of the "counterfactual robustness" for their material inferences. Two speakers, who on the surface agree on a factual claim, may turn out to not be in agreement, if they turn out to disagree about virtually every counterfactual related to the factual claim. In such cases, we can start to wonder if they even agree about the

¹ Arguments of this form are also found in Brandom (1994), see for instance pp. 633-636. In Price (2007), Price considers a different line of interpretation of Brandom's project, concentrating upon the *genealogical* side, the side that is interested in finding out how our modal thought develops, not what justifies it. I am not averse to such a reading, but have chosen to concentrate on understanding Brandom's reasoning as an attempt to find a justification.

basic, factual sentence. Agreement about factials requires some kind of underlying agreement about counterfactuals.

The next step is that such counterfactuals can be used to introduce modal locutions, in the way Ryle suggests his (1950): understand “If p were true, q would be true” as being equivalent with “It is not possible that p and not q ”. Then we can take our use of counterfactuals to account for our knowledge of modal truths. Since we are discursive creatures, we can be granted such knowledge, and hence we have provided a transcendental argument for modal knowledge: modal knowledge is needed for non-modal knowledge, and since we have such knowledge, we should realize that empiricist worries about modal knowledge are without basis.

This picture of how our modal knowledge should best be understood requires further grounding. As it stands, it is little more than a sketch (so my account is a sketch of a sketch), and some of the details may turn out to be problematic. How severe is the Updating Problem? How much epistemic responsibility is required? What kinds of counterfactual are relevant for understanding a given factual statement? How *do* we know the counterfactuals? How much disagreement about counterfactuals can we tolerate? But I think that the general thrust of Brandom’s argument should be clear enough from my brief account, and it is this general strategy for grounding our knowledge of counterfactuals that I will focus on in my talk.

Some of the steps in the reasoning above are not controversial. For instance, (1) appears to be little more than a truism, whereas some other steps may be more problematic. The claim about non-monotonicity seems to me to be substantially correct, but I guess there are people out there, still trying to come up with completely monotonic patterns of reasoning for observational matters. But even if we grant all steps above, a central difficulty remains. What backing does the five-step argument give to our ability to use counterfactuals and modals? From an empiricist point of view, there is something almost miraculous about our supposed knowledge of counterfactuals: how *could* we know these, when there is no basis in observation for them? There may still be room for the empiricist to wonder.

First, how bad is the Updating Problem? There is at least a way of understanding this problem that is a good description of the scientist’s predicament: we simply don’t know in advance which new observations that are relevant for the standing of the beliefs we happen to hold right now. This goes with the non-monotonicity Brandom rightly stresses: if non-monotonicity means that we cannot spell out in advance what kinds of defeaters there will be for our beliefs, then we will be stuck with the Updating Problem, no matter what. Appeals to counterfactual robustness won’t help us here. It is part of scientific reasoning that new evidence, unanticipated or arriving from some completely unexpected source, *can* make us change our firmly held beliefs. So even if the Updating Problem maybe is mitigated by Brandom’s suggested way out, we will never be completely rid of it.

For the second set of questions, it seems that Brandom’s argument gives us a weak conclusion – we are only given a guarantee that we, as responsible subjects, must make use of *some* counterfactuals, but nothing much is said about *which* counterfactuals, and it is not shown that we actually will *know* any counterfactuals, or which counterfactuals that are true. Perhaps all the five-step argument gives us is a weaker conclusion: in order to count as knowing facts, we must be *using* counterfactuals – but there is no real check on our ability to use them; it has not been shown that we must be able to use them correctly, as long as there is some agreement between speakers. The argument will by itself not give us that much check on our uses of counterfactuals. It can perhaps give us a backing for a conclusion that is slightly weaker, but not without bite: in order to count as communicating, we must hold each other responsible to roughly the same counterfactuals. Again, there is one way to understand this that even the staunchest empiricist can go along with. Rough agreement in counterfactuals is just a special case of rough agreement in beliefs, and will so not by itself have to be disallowed by the empiricist, even if she of course will have grave doubts concerning the legitimacy of such beliefs. So we have two readings of the conclusion of the argument:

Weak reading: we must *be in agreement* on some counterfactuals, if we are to see each other as epistemically responsible and hence discursive creatures.

Strong reading: we must *know* roughly the same counterfactuals, if we are to see each other as epistemically responsible and hence discursive creatures.

The weak reading appears to follow from the five steps and the Updating Problem, but does not by itself give us knowledge of counterfactuals. The stronger reading would do just that, but I cannot see how the five steps would give us knowledge. If Brandom's five-step argument is to give us the strong, and desired, conclusion, we need something extra: something showing that the use of a particular modal or counterfactual statement is *justified*, that there is some way to distinguish correct from incorrect use. The bare assertion that our knowledge of facts must rest on our using counterfactuals in some way or other, does not by itself show enough, even if it shows that we must agree (roughly) in our uses of counterfactuals.

But how might we go about to show this? There is a dilemma for Brandom's position here: the transcendental argument doesn't show that our use of counterfactuals is correct, so something more is needed. But if we beef up the transcendental argument with some other argument that shows the correctness of the use of counterfactuals, then no transcendental argument is needed: then we have a direct argument for our use of counterfactuals. Such "beefing up" might for instance be some kind of traditional appeal to intuition of necessary truths. But few put much faith in such intuition, and if we were to have such faith, there would be no need for the five-step argument: in that case we would just have direct modal knowledge. So the ambitious argument to show the indispensability of modal knowledge for observational knowledge is left dangling. There is in fact another indirect way of reasoning, that might appeal to Brandom. This is saying that we *do* have knowledge in general, and some of this knowledge is observational. Then we can argue that since there is no such thing as a purely observational layer of knowledge, we must in fact be granted modal knowledge as well. Perhaps this extra reasoning can be made to work, but it still seems vulnerable to the observation that it by itself doesn't give us knowledge of any specific modals. So the argument as it stands is still incomplete.

The third issue I want to raise concerns Brandom's phenomenalism about modality.² The connections between necessity and our commitments appears to be problematic. The claim is that necessity is in a way a product of our commitments (for reasons of space, I will skip the fancy machinery of *Between Saying and Doing*; not because it's uninteresting, but because I think the points I will be bringing up are unaffected by the exact set-up of the claim). Modality is as it were arising out of incompatibility of commitments (see Brandom 2008, ch. 5 and Appendices). But it seems that even if we spell out the intricate and interesting details of this story *much* better than this, some problems remain.

We can for instance all mistakenly see ourselves as committed to a certain claim, whereas it in fact turns out that we are not – we had missed some subtle aspect of the reasoning. Or things can be the other way round: we think we are not committed to something that in fact is a necessary consequence of other things we *are* committed to. It took a while until people managed to come up with alternatives to Euclidean geometry. The parallel postulate was not necessary, but everyone was holding themselves and everyone else committed to it for a long time. Intuitionists think that some people are mistakenly committed to something that in fact is false. People working within the confines of classical logic think that dialetheists are wrongly seeing themselves as non-committed to the law of non-contradiction.

Perhaps we can try to counter such observations by saying that this doesn't really matter: various communities can be *determined* by their different commitments. So instead of a whole group being wrong about a given commitment, we might see the group as existing as the result of the commitment. But this doesn't really work, since determination is itself a modal notion (if *A* is determined by *B*, the truth of *A* is necessitated by *B*), so commitment, understood in this modified way, cannot be used to explain modality.

A related problem remains. Consider a purported necessary truth *p*. In a sense, consistent with the phenomenalism about deontological status, *p* is a necessary truth because we are committed to treat it like a

² See Brandom (1994, 291ff).

necessary truth. Are we then committed to treat this commitment as creating a necessary truth? Neither alternative is good. If the answer is *yes*, then we need an extra commitment to make our commitment into a commitment that gives us necessary truths. If the answer is *no*, then it is unclear what the first commitment accomplished. No doubt this general sketch of an argument can and should be tidied up in several respects, not least to do better justice to the intricacies of Brandom's developed theory, but I think a workshop on Brandom's analytical pragmatism would be the right setting to at least start examining the issues.

A final point I would like to consider is probably pretty minor, but it seems to indicate something of interest. This is that there may be a kind of lingering circularity in the Kant-Sellars thesis itself. It is not clear from the wording, but the gist of the thesis must be counterfactual in nature: if we *didn't have* counterfactual knowledge, we *would not have* observational knowledge at all. So someone who doubts that we know counterfactuals will not be swayed by a thesis, formulated in a way that assumes that we *do* know counterfactuals. Why would this still be a minor point – isn't circularity always bad? Well, in this case, maybe not. It is not always unfair to assume an argued version of what you think is the most basic means of acquiring knowledge. The alleged circularity here is perhaps unavoidable but harmless – if our basic ways of knowing something really must integrate knowledge of counterfactuals, then it is no great surprise that the statement of such knowledge in itself involves counterfactuals, just as we have to lean on some kind of logic when explaining logic. We have to start somewhere, and starting in the middle of our ability to know stuff, just like any naturalist would, seems like a good place to start.

Bibliography

- Brandom, R. (1994) *Making it Explicit*, Harvard UP, Cambr., Mass.
- Brandom, R. (2001) "Modality, Normativity, and Intentionality", *Philosophy and Phenomenological Research* vol 63:3
- Brandom, R. (2008) *Between Saying and Doing. Towards an analytic pragmatism*, Oxford UP, Oxford
- Fine, K. (2005) *Modality and Tense*, Oxford UP, Oxford
- Peacocke, C. (1999) *Being Known*, Oxford UP, Oxford
- Price, H. (2007) "Brandom and Hume on the Genealogy of Modals", forthcoming in *Philosophical Topics*, currently available at: <http://www.usyd.edu.au/time/price/publications.html>
- Ryle, G. (1950) "'If', 'So', and 'Because'", in M. Black (ed.), *Philosophical Analysis*, Cornell UP, Ithaca
- Whewell, W. (1840) *Philosophy of the Inductive Sciences Founded upon their History*, J.W. Parker and Son, London
- Williamson, T. (2007) *The Philosophy of Philosophy*, Blackwells, Oxford

The Status of Intentional Vocabulary in Discursive Practice. (Reading *Making It Explicit* in the Light of *Between Saying and Doing*)

David Lauer

dlauer@zedat.fu-berlin.de

<http://www.geisteswissenschaften.fu-berlin.de/we01/mitarbeiter/dlauer/index.html>

Freie Universität Berlin, Institut für Philosophie, Habelschwerdter Allee 30, 14195 Berlin

Abstract. The topic of my paper is an aspect of the relation between Robert Brandom's *Between Saying and Doing* (henceforth *BSD*) and his *Making It Explicit* (henceforth *MIE*), and the philosophical projects developed in them respectively. One of the many merits of *BSD* is that it provides a metaperspective on what is going on in *MIE*. One could almost jokingly say that the theory of *BSD* can be made to serve a goal we might call *Making 'Making It Explicit' Explicit*. My aim is to critically reassess (in the light of *BSD*) what *MIE* has to say about the status of a particular kind of pragmatic explicative vocabulary, namely vocabulary for intentional interpretation (propositional attitude ascription), or *intentional vocabulary* for short. Thus I will offer an exercise in what the dead mighty Germans called thinking critically *mit X gegen X*, and which they regarded, rightly I think, as a most philosophical form of praise.

1 What Is *Making It Explicit* Explicit?

I want to begin with a very short recapitulation of what the project of *MIE*, as I understand it, actually is. Its central goal is to understand genuine intentionality (to borrow John Haugeland's term), which is the capacity of systems to have non-derived mental states that have *content* expressible by *that*-clauses, i.e. propositional attitudes. Another way of putting the point would be to say that having genuine intentionality in this sense just means *having thought*. This is what we ascribe to systems whose practices we can only make sense of by taking the intentional stance to them. At least that is what Brandom seems to be saying when he writes that a system counts as exhibiting this specific type of intentionality »insofar as the differential responsiveness of the system to the results of its own performances is *essentially* mediated by states whose functional role in the feedback process can be understood *only* by taking them to be propositionally contentful, that is, by specifying them in an intentional vocabulary« (*BSD*, p. 183; italics mine). Now I know that it would not be literally correct to say this, but there seems to be a peculiar kind of VP-*necessity* claim involved here:¹ In

¹ In the present context, I assume acquaintance with the vocabulary of meaning-use-analysis that Brandom introduces in *BSD*. »VP-necessity« is not an official element of this vocabulary.

order to count as genuinely intentional, one might be tempted to say, it is not enough that intentional vocabulary be *sufficient* to specify a system or the practices it is involved in, for intentional vocabulary is sufficient (*can* be used) to specify almost anything, including the behaviour of thermostats and iron bars. It rather seems to be the case that a system exhibits genuine intentionality if and only if it is necessarily specified by intentional vocabulary.²

What are the essential features of genuine intentionality which a pragmatist account would have to make intelligible? It seems to me to be in line with Brandom's thinking to say that there are (at least) three such features:

(a) *Rationality*: Philosophers like Dennett and Davidson have stressed that to ascribe intentional states to a system is to describe it in the light of the constitutive ideal of rationality. To explain and predict a system's behaviour in the light of propositionally contentful intentional ascriptions is to *rationalize* its behaviour, to *make sense* of it by ascribing states to it that would count as *reasons* for the system to behave in the way it does. Having genuine intentionality is therefore the same as being at home in the space of reasons, as Sellars put it. What constitutes the identity of a contentful state or expression is not its causal position in the natural world, but its *inferential position* in a rationally connected web of contentful states.

(b) *Objectivity*: Propositional content can be characterized in terms of truth conditions, and even if a semantics does not use the notion of truth as its starting point, in the end it has to arrive at the point of being able to explain what it is that makes propositional contents represent states of affairs. This amounts to the same thing as being able to explain the representational character of propositional content. What is propositionally contentful necessarily has a representational aspect, it represents things as being a certain way. Nothing that does not display that aspect would be recognizable as expressing a proposition.

(c) *Reflexivity*: A system having thought, minimally, would be one whose behaviour can be interpreted as displaying states with content. But that is not enough. Not every understandable behaviour is understanding behaviour. Being minded demands not only displaying, but *understanding* content (meaning). The rationality and objectivity of the contents that are exhibited in the practices of the system, that is, must not exist only in the eye of a beholder. They must exist, be there exhibited in the practice, *for* the systems engaging in those practice itself.

Now the most fundamental theoretical commitment of *MIE* is a commitment to *pragmatism*. For a pragmatist, the question »What is genuine intentionality?« is transformed into the question »What is it that systems we treat as genuinely intentional are capable of *doing* that allows and necessitates specifying their practices in intentional terms?« That is, what are the necessary and sufficient conditions for a practice to count as instituting or exhibiting genuine intentionality? At this point, pragmatism aligns itself with the linguistic paradigm in the philosophy of mind, which claims that the practices on the basis of which we call some beings genuinely intentional are *discursive* practices. Having language, according to this tradition, is constitutive of having thought. Having language itself, in turn, is to be understood in terms of being

² This would not have to be taken to mean that it would be impossible to specify a genuinely intentional practice in non-intentional ways, only that the non-intentional ways of specifying it would be conceptually dependent on the intentional one.

capable of *using* linguistic symbols. Thus, the project of *MIE* is to specify what one has to do in order to count as speaking a language. I will call any set of discursive practices which suffices to confer on its practitioners genuine intentionality (including its central features specified above) an *Autonomous Discursive Practice* (ADP). Thus it is a condition of adequacy on a normative-pragmatist account of intentionality that it be able to say what it means for a practice to institute inferential relations, the objectivity of content, and the reflexivity of its practitioners, and thus makes it an ADP.

2 What Does One Do If One Makes Something Explicit?

The aim of *MIE*, we can now say, is to make explicit what having language is. Now what does *making it explicit* mean? Brandom's standard characterization of the relation between the implicit and the explicit is to say that, by making it explicit, we become able to explicitly *say* what before we could only implicitly *do*. This corresponds to the idea that making it explicit is codifying a piece of implicit *knowing-how* in the form of a piece of propositional *knowing-that*. But certainly this way of putting it gives only a very broad and unspecific idea of this central notion of the theory. E.g., is every act of *talking about* what we do an act of making it explicit, or is it being able to *say what counts as doing* what we do, or even becoming able to do whatever it is that we do *by saying something*? I suggest to understand *BSD* as offering the conceptual tools make this a little clearer in the following way:

(a) The fundamental idea, as I take it, is to define *specification* as a relation between a vocabulary V1 and a practice P2. V1 is said to be sufficient to specify P2 (to be VP-sufficient) if and only if it allows one to say what one has to do in order to count as engaging in P2. Thus, the instructions a tennis coach gives to a novice to the game, telling him what counts as a proper service, what counts as winning a set, and so on, might be said to specify the practice of playing tennis.

(b) Building on the notion of specification, we can then introduce a precise notion of making it explicit as follows: Let us introduce the practice P1 that is PV-sufficient to deploy the vocabulary V1 (that is, engaging in P1 counts as using V1). Further, let us assume that P1 can be elaborated (by algorithm or by training) from P2. Thus PP-sufficiency holds between P2 and P1: PP-sufficiency holds »when the capacity to engage in one sort of practice or to exercise one sort of ability is *in principle sufficient* for the capacity to engage in other practices, or to exercise other abilities« (*BSD*, p. 33). If these conditions obtain, the vocabulary V1 used to specify P2 isn't just any old vocabulary. It is dignified in a certain way, because everything one needs to do in order to deploy and understand V1 is in principle contained in P2 (in the sense that it can be elaborated from it). Thus I suggest that we should say that V1 makes P2 explicit if V1 is elaborated from and explicative of P2. I will call vocabulary of this sort explicative vocabulary.

(c) We arrive at a special case of the relation of making it explicit if we introduce the further assumption that P2 (the practice being made explicit) is itself a discursive practice, that is a practice sufficient to deploy a vocabulary V2. In that case, the resultant of V1 being VP-sufficient to specify P2, and of P2 being PV-sufficient to deploy V2, is a VV-sufficiency relation between V1 and V2, which means that using V1 allows one to say what one must do in order to count as saying the things expressed by using V2. In other words, V1 is a pragmatic metavocabulary for V2. We might call

this *metadiscursive specification*: saying what counts as *saying* something else. Furthermore, since P1 is elaborated from P2, we can say that – in a way – we always already understood everything we need to understand V1 if only we understood V2. Thus, V1 is »semantically transparent« towards V2. Therefore it »can legitimately be appealed to as an auxiliary elaborating vocabulary in semantic analysis« (BSD, p. 49) of V2. In this case, V1 is not just a metavocabulary to V2, it is a metavocabulary that is elaborated from and explicative of P2. It stands in a pragmatically mediated VV-sufficiency relation to V2 Brandom calls »LX«. ³ Thus it is *LX explicative* vocabulary. If V1 is an LX explicative vocabulary for any discursive practice, it is *universally LX explicative vocabulary*.

Now, let us apply these specifications to the project of *MIE*. We can then say that its aim is to develop a universal semantic metavocabulary, powerful enough to specify discursive practices as such. It must satisfy the following conditions:

(1) In order to be a *universal* semantic metavocabulary (one for discursive practice *as such*), it has to be VP-sufficient to specify any language whatsoever, that is, any Autonomous Discursive Practice (ADP).

(2) In order to be a *legitimate* semantic metavocabulary, it must be shown to be possible to extend, by elaboration, *any* ADP into a set of practices that are PV-sufficient to count as deploying the semantic metavocabulary.

Conditions (1) and (2) guarantee that the metavocabulary is elaborated-explicating for any language, or *universally LX explicative vocabulary*.

(3) In order to be a philosophically *illuminating* semantic metavocabulary, it must be *conceptually independent* of the vocabulary of intentional interpretation that it is supposed to explain. It must not, on pain of circularity, make use of intentional locutions, e.g. specify certain performances within the practices as acts of expressing thoughts or as acts of saying *that p*, for it is precisely the question what it *means* for a practitioner to even *have* a thought, and therefore what it means for any of his performances to be *expressive* of a thought. What is needed, therefore, is an account that says what it is that *makes* a certain practice discursive in the first place, without presupposing that the practitioners are genuinely intentional beings. Thus, all in all, the project is, in Brandom's words, »specifying in a non-intentional, non-semantic vocabulary what it is one must do in order to count as deploying some vocabulary to say something, hence as making intentional and semantic vocabulary applicable to the performances one produces (a kind of pragmatic expressive bootstrapping)« (BSD, p. 78 f).

3 The Layer Cake Picture of Autonomous Discursive Practice and the Status of Intentional Vocabulary

In *MIE*, Brandom's general term for universally LX explicative vocabulary – vocabulary being *explicative* of *any* discursive practice – is *logical vocabulary*. Logical vocabulary in this broad sense does not only comprise logical vocabulary proper (logical operators, quantifiers), but also normative, modal, and intentional vocabulary. In the terminology of *BSD*, the term »logical vocabulary« is reserved for logical vo-

³ See *BSD*, p. 47.

cabulary in the narrow sense, which is just one species of the genus of universally LX explicative vocabulary. I will stick to the broad use of *MIE* unless otherwise indicated. Now, if logical vocabulary is explicative vocabulary for any ADP, it follows that, for any ADP, the practical capacities sufficient to deploy logical vocabulary are not PP-necessary for the capacity to engage in the ADP. In other words, an ADP (a language) does not necessarily include logical vocabulary (although by definition any ADP includes practices that are PP-sufficient for the potential *elaboration* of logical vocabulary). This is what Brandom calls the »two-stage or layer cake picture of the relation of logical to non-logical vocabulary« (Brandom 1997, p. 206).

Let's have a closer look at this: Practitioners in a language game (in the sense of an ADP) have a certain know-how that consists in their being able to calculate the consequences of each move in the game for the normative statuses of every player, licensing or ruling out certain further moves. These calculations are what their *understanding* of the moves in the practice consists in. Brandom's by now famous metaphor for this process is *to keep the score* of the game. Now scorekeeping itself is *implicit* in the ADP: The way the score is kept by a player just manifests itself in what normative attitudes any player adopts toward her fellow players, what she takes them to be entitled and committed to. No logical vocabulary, no device of *explicit* scorekeeping is needed for this practice to work. The introduction of logical vocabulary allows practitioners not only to *treat* others as being committed to certain norms, but to treat them as committed by *saying* that they are. But this, according to the official picture, only makes explicit the norms – hence, in the appropriate practices, the contents – that were, so to speak, always already there, *implicitly* conferred on performances by the roles they play in an appropriately structured normative practice. According to the Layer Cake Picture, ascribing intentional states to practitioners and propositional contents to their performances should be understood as *making explicit* the norms that *were already implicitly instituted* in a practice of the suitable sort.

Now I want to focus on a special kind of logical, i.e. universally LX explicative vocabulary, namely the locutions we use to engage in explicit intentional interpretation of speakers and thinkers, i.e. propositional attitude ascription vocabulary like »says that«, »believes that« and so on. I will call this type of vocabulary *intentional vocabulary* for short. Brandom's claim is that this type of vocabulary is a type of logical, i.e. universally LX explicative vocabulary. More precisely, it is a type of *pragmatic* (as opposed to semantic) logical vocabulary, i.e. it belongs in a box with other types of normative vocabulary (as opposed to modal vocabulary and logical vocabulary in the narrow sense). It is explicative of the pragmatic interpersonal relations between normative attitudes and normative statuses that constitute any ADP, and »[i]n virtue of this explicating expressive role, propositional-attitude-ascribing locutions deserve to count as logical vocabulary« (*MIE*, p. 499). It follows from this that what was said about logical vocabulary in general must also be true of intentional vocabulary. Thus, an ADP must be interpretable as being discursive even before intentional locutions are introduced into (elaborated from) it. The practitioners must be interpretable as saying things even before they can *say* that they or others say things. Brandom is very clear about that:

»[S]corekeeping (and so linguistic practice generally) does not require that one be able explicitly to attribute deontic statuses – to say (assert) that someone is committed

to the claim that *p*. The logical locutions whose expressive role is to make the adoption of such pragmatic attitudes explicit in the form of claimable contents – propositional-attitude-ascribing vocabulary such as the regimented ‘...is committed to the claim that ...’ or its vernacular correlate ‘...believes that...’ – form an optional superstratum whose expressive role can be understood in terms of what is implicit in ground-level linguistic practice, but which is not required for, or presupposed by, such practice.« (*MIE*, p. 629)

4 Going Intentional Demands Going Explicit

If what I have said so far is correct, Brandom is committed to the following theses:

(1) Something is a piece of logical vocabulary if and only if it is a piece of universally LX explicative vocabulary.

(2) If something is a piece of intentional vocabulary, it is a piece of logical vocabulary.

(3) An ADP does not necessarily include logical vocabulary.

(4) Therefore, an ADP does not necessarily include intentional vocabulary.

I will now try to argue that claim (4) is to be rejected. What I want to claim is that the use of intentional vocabulary is a necessary element of any autonomous discursive practice itself. *Pace MIE*, these locutions do not form an optional superstratum whose expressive role can be understood in terms of making explicit discursive contents that were already implicit in ground-level discursive practice. Rather, these locutions are not *expressive*, they are *co-constitutive* of discursive commitments, and hence of any autonomous discursive practice. In other words, practices sufficient to deploy intentional vocabulary are PP-necessary for any ADP, just as, e.g., inferential practices.

The argument for this claim runs as follows: The Layer Cake Picture implies that it must be possible to account for the institution of objective deontic statuses and conceptual contents in terms of practical deontic attitudes of practitioners who do not yet have the capacity for higher-order normative attitudes. But understanding a practice as an *autonomous discursive* practice – attributing genuine discursive intentionality to the practice and its practitioners – is specifying a set of practices which confer propositional contents on performances *for the practitioners*. It is attributing a *grasp*, on part of the scorekeepers, of objective discursive commitments with representational contents. This means attributing to a scorekeeper a grasp of the *difference* between commitments another scorekeeper or she himself has *acknowledged* and those she has consequentially but unknowingly *undertaken*, that is, the capacity to attribute differing *attitudes* (acknowledging / undertaking) toward statuses (commitments). Without the capacity to attribute normative attitudes toward commitments on part of a scorekeeper, there is no genuine intentionality, because being able to attribute this – to others or to oneself – is what manifesting a grasp of the difference between *objective* normative statuses – what someone has bound herself by – and normative attitudes – what she merely takes herself to be bound by – consists in. Therefore, without concepts and expressions of these attributions and statuses (that is, without logical concepts in Brandom’s sense of the term) on part of the scorekeeper, genuine intentionality could not be ascribed. But possession of these concepts is manifested in explicit practices of rationally engineering, updating, and rectifying discursive commitments either socially (this is highlighted in *MIE*, Chapter 8) or self-reflectively (this is high-

lighted in *BSD*, Lecture 6). But either of these practices depends on the availability of practical performances that are in themselves acts of self- or other-*ascribing* discursive commitments. The capacity to explicitly *ascribe* normative statuses therefore is a necessary condition for there to *be* such statuses instituted *in* the practice. Thus, a scorekeeper can only be credited with genuine intentionality if she is engaged in logical practices of explicitly keeping the score, that is using explicit ascriptions of attitudes and commitments. But, since objective contents and propositional attitudes can only be said to be conferred on performances in a practice that institutes conceptual norms, norms that are objective for the practitioners who are bound by them, it follows that we can attribute the conferral of *propositionally contentful* normative statuses only to practices whose practitioners are full-scale interpreters, that is, *fully logical* at least in the sense of having the capacity to use intentional vocabulary.

Therefore, there could not be a grounding layer of pre-logical autonomous linguistic practice, instituting objective norms with propositional contents, while the second explicative layer merely made explicit the norms implicitly constituted on ground level. Rather, before the advent of intentional vocabulary, the norms on ground level could not be understood as instituting contents at all. *No practice in which the norms are purely implicit could institute objective, i.e. propositionally contentful normative statuses* and thereby confer genuine intentionality on its practitioners and propositional contents on their performances. Language in the full sense of the term *cannot* be understood in layer-cake-style. There may have been *norms* on ground level, but no conceptual norms (not even implicit). Before we made »it« explicit, »it« wasn't genuinely intentional, that is propositionally contentful. What makes the norms conceptual, so to speak, is their explicitation. *Going intentional, in other words, demands going explicit*. This means that the attribution of genuine intentionality is justified if and only if the interpretees are full-scale interpreters too – interpreters who can attribute genuine intentionality themselves, who can, that is, *explicitly* keep the score. The answer to the question, then, what warrants – indeed, necessitates – taking and interpreting a practice as an autonomous linguistic practice, conferring conceptual states and contents on its practitioners and their performances, is that the practice be PV-sufficient to deploy the very vocabulary that is used to specify it – intentional vocabulary.

Another, much more traditional way of making this point would be to say that if you want to account for full-scale objectivity, you need to account for full-scale subjectivity too. Put in Kantian terms, there can be no propositional thought, no thinking of objects, no genuine intentionality, without the conceptual capacities needed to attach the »I think« – and, of course, as Davidsonians would add, the »she thinks« – to one's thoughts. Brandom, of course, being a Kantian and a Davidsonian, is the last person on Earth who needs to be lectured about this. But I wonder if pointing this out shouldn't in fact make the conclusion I reached above more than palatable to him.

5 Conclusions

If this kind of reasoning is accepted, it follows that one of the premises of the argument ascribed to Brandom above has to give. I suggest that we should give up (3) and admit that logical vocabulary at least of the pragmatically explicative type is PP-necessary for any ADP.

Now, in reaction to a brilliant paper by Daniel Laurier, who makes more or less the argument I have rehearsed here, Brandom is close to admitting that (3) should be given up,⁴ but adds: »The claim that ascriptional locutions are in-principle late-coming expressions that make explicit features of practices intelligible as autonomously discursive in advance of their introduction could be acknowledge to be mistaken without upsetting the general methodology or architectonic of the project.« (Brandom 2005, p. 244)⁵

I want to cast a little doubt on the entitlement for this nonchalance. Let me admit right away that certainly, even if what I have said were correct, the walls of inferentialism would not come tumbling down. But I think the view presented here does have significant consequences, first for the general architectonic of *MIE* as it is presented in that book itself and also in some very recent of Brandom's articles, which, if I am right, must count as serious self-misrepresentations; secondly for some basic claims of *MIE* and *BSD*. For reasons of space limitations, I will only mention the latter ones:

Obviously, the view presented here quite dramatically changes the conception of an Autonomous Discursive Practice. I have argued that any ADP must be PV-sufficient to deploy intentional vocabulary, and that it follows from this that it must contain practices PV-sufficient to deploy logical – in the broad sense, i.e. universally LX explicative – vocabulary.

It also requires changing our minds about what explicative vocabularies of the pragmatic type, i.e. normative and intentional vocabularies, are. Contrary to the official picture, I would claim, they turn out *not to be metavocabularies* of any ADP, but PP-necessary *sub-practices* of any ADP. Note that they can still be called *explicative* vocabularies. For it can still be maintained that there is a set of basic practices they specify from which the practice of using that type of vocabulary may be elaborated. But before that, to put it simply, that set of basic practices was *not* an ADP. Thus vocabulary of that type makes something explicit. We may even say that it makes commitments and entitlements explicit. But before the advent of that type of vocabulary, these commitments did not qualify as discursive commitments (thus, the »it« in »making it explicit« here does not refer to something conceptual). Elaborating practices sufficient to deploy intentional vocabulary is not introducing a metavocabulary for an ADP, it is instituting one. Making it explicit, in this case, is therefore not, as Brandom used to say, »a process of transformation of what in virtue of its role in that process becomes visible as a *content* that appears in two *forms*, as implicit and then as explicit« (Brandom 2000, p. 16), it is rather the process of instituting conceptual content in the first place.

Finally, let me add a speculation of what all this might mean for our picture of semantic (modal/logical in the narrow sense) explicative vocabulary. The argument I have presented only applies to normative and intentional vocabulary. Thus it would be unwarranted to generalize the claim that normative and intentional vocabularies are necessarily a subset of any ADP to *all* types of logical vocabulary. However, it is

⁴ In fact, he seems to be admitting it – as Laurier notes – in *MIE*, e.g. on p. 640.

⁵ Maybe this is behind his uncharacteristically defensive (but still, if I am right, unwarranted) claim in *BSD* that intentional vocabulary »may or may not be a sub-vocabulary of the autonomous vocabulary« (*BSD*, p. 183).

certainly not implausible to assume, as John McDowell does,⁶ that the ideas assembled here, if correct, cannot leave untouched the account of that other type of universally LX explicative vocabulary, given how narrowly entwined they are even in Brandom's own account of their relation in Chapter 6 of *BSD*. However, more arguments would have to be rehearsed in order to say something sensible about this, so it must wait for another occasion.

References

- Brandom, Robert (1994): *Making It Explicit*. Cambridge/MA: Harvard UP.
- Brandom, Robert (1997): »Replies«, in: *Philosophy and Phenomenological Research* 57, 189-204.
- Brandom, Robert (2000): *Articulating Reasons*. Cambridge/MA: Harvard UP.
- Brandom, Robert (2005): »Responses«, in: *Pragmatics and Cognition* 13, 227-249.
- Brandom R. B. (2008) *Between Saying and Doing*. Oxford, Oxford UP.
- Laurier, Daniel (2005): »Pragmatics, Pittsburgh Style«, in: *Pragmatics and Cognition* 13, 141-160.
- McDowell, John (2005): »Motivating Inferentialism. Comments on Making It Explicit (Ch. 2)«, in: *Pragmatics and Cognition* 13, 121-140.

⁶ E.g. in McDowell 2005, p. 133 ff., and on some earlier occasions.

Logic, Semantics and the Theory of Meaning

Inferentialism and the Normativity of Meaning

Jaroslav Peregrin

peregrin@ff.cuni.cz webpage: <http://jarda.peregrin.cz>

Institute of Philosophy, Academy of Sciences of the Czech Republic, Jilská 1, 110 00 Praha 1, Czech Republic

Abstract. The paper addresses some frequent objections to the claim that meaning is normative, thus defending the inferentialist construal of meaning that does entail this claim. The objections we discuss are (i) that there is no norm that assertions should aim at the truth, (ii) that there are no norms commanding us how to speak and (iii) that a normative account is bound to collapse into a naturalistic one. We conclude that the way in which normativity is built into the inferentialist framework is not vulnerable to any of these objections. Along the way, we try to clear up some misunderstandings surrounding the status of normativity within semantics.

1 Is Meaning Normative?

Inferentialist view of meaning (based on the assumptions that (i) a meaning is not an object labeled - stood for, represented ... - by an expression; and that (ii) meaning is normative in the sense that to say that an expression means thus and so is to say that it should be used so and so¹) faces two kinds of objections. First, there are general objections to any normative construal of meaning, and then there are more specific objections targeted specifically at inferentialism. In this paper I will address the objections of the former kind (I address some of those of the latter kind elsewhere²).

To give an example of the way a general objection to the claim that meaning is normative usually goes, let me quote one of its most influential exponents, Paul Boghossian (2005, 212):

To put the matter concisely, the linguistic version of the normativity thesis, in contrast with its mentalist version, has no plausibility whatever; and the reason is that it is not a norm on assertion that it should aim at the truth, in the way in which it is a norm on belief that it do so.

¹ See Peregrin (2008) for a more detailed exposition of this aspect of inferentialism.

² See Peregrin (2001); and especially Peregrin (t.a.) (All my papers as well as a longer draft of this paper are available from my homepage at jarda.peregrin.cz/mybibl.php.)

Sellars (1992, p. 101), whose lead I follow here, thinks otherwise: he understands truth as correct assertability, which entails that an assertion is *correct* if what is asserted is true; and hence, in this sense, that we *ought to* assert the truth³.

Meaning, according to the inferentialist, is normative in the sense that when I say that an expression means thus and so, then what I say does not amount to stating a fact, but rather invoking a propriety: it is stating that the expression is *correctly* used thus and so. True, on one of its readings, this may still be read as stating a kind of fact, namely that an activity within a community is - as a matter of fact - governed by certain rules; however, there is a second and crucial reading on which this is not the case, for the claim does not amount to a declarative statement at all, it is rather an *endorsement*. As Sellars puts it in his letter to Chisholm,

My solution is that

'...' means ---

is the core of a unique mode of discourse which is as distinct from the *description* and *explanation* of empirical fact, as is the language of *prescription* and *justification*. (Chisholm and Sellars, 1958, p. 527)

I am fond of describing the situation in terms of an 'inner space' that some systems of rules have the ability to constitute. From outside of the space we can only report the fact that such or other rules are in force for the insiders; but once we join the insiders, the rules start to be in force *for us* and hence be in force (full stop); and claiming this does not amount to stating a fact, it is a different speech act. Hence, let me call the former reading the 'outsider' reading and the latter the 'insider' one.

Thus, claims to the effect that something is correct or that something ought to be done (I will call them *normatives*, for short), on the insider reading, are something different from assertions or reports. They do not report that something is the case, they point out that something *ought to be* the case; hence they always involve the utterer's taking a rule for being in force, her endorsing it. In this respect, they are similar to oaths of loyalty: they always involve one's decision to assume a certain status, namely to bind oneself by a rule, and in this sense they *institute* something (namely a certain social link) rather than *report* it. However, the case when the institution happens in a single instant (like in the case of signing an oath) is only a very special case; more generally, binding oneself with a rule is more like the case of loyalty that is not formally established with an instant oath, but is continuously testified by one's performances and declarations. Normatives of this kind involve the instituting and upholding of a rule.

This, of course, is not the only thing that the normatives do: besides this, they may express that the rule in question, as applied to a particular case, renders the case right or wrong. If I tell you "You should not kill this cat", I am claiming that (given a certain rule to which I, and presumably you, submit) killing this cat would be wrong. Concentrating on this, we might say that the normatives, even on their 'outsider' read-

³ I also think that there *cannot be* a norm that we should *believe* the truth, as it seems to me obvious that one is not free to decide what to believe. We *can* say that it is a norm that we *ought to interpret our peers as believing* the truth. This is the celebrated Davidsonian principle of charity; but this is a far cry from claiming that we *ought to believe* the truth.

ing, report a specific kind of facts, *viz.* normative facts - but the possibility of taking this characterization literally is compromised by the fact that the alleged fact is a fact only as a result of a rule that does not exist quite independently of the statement, for the statement takes part in its constitution.

Hence, compare the following two claims:

- (1) Killing this cat would be easy
- (2) Killing this cat would be wrong

Both claims can be seen as classifications: they classify a certain hypothetical action from a certain viewpoint. But whereas (1) uses a classificatory criterion that is wholly independent of the classification and can thus read as objective in the most straightforward sense, things are different with (2). Though it can, perhaps, be read in the same objective and hence disengaged way (the 'outsider' reading), the important point is that it can also be read in a rather different way, where its aim is not only to classify, but at the same time to uphold the criterion that is employed, to declare one's allegiance for it (the 'insider' reading). Hence two sentences, though their grammatical structure is the same, may be used to accomplish dissimilar speech acts.

2 Do inferential rules guide us?

Another kind of objection to a construal of meaning as a product of rules was most clearly articulated by Glüer and Pagin (1999), who argue that "rules that can be regarded as *constitutive of* meaning, are not capable of guiding speakers in the ordinary performances of speech acts." I think that here we must distinguish between two senses of "guiding": on the most straightforward sense, in which "guiding" amounts to directly telling what to do, Glüer and Pagin are undoubtedly right. But there is a different sense of "guiding", on which "guiding" involves telling what *not* to do, and in this sense rules that are constitutive of meaning *do* guide us. These rules erect limits to what we can assert if we are to use expressions as meaning what they do⁴.

However, the rules in question come to be buried under the visible surface of our discursive practices. If we see the practices as employing expressions with specific meanings, then the rules become invisible for their being in place is part and parcel of the very meaningfulness of the discourse. Taking the expressions as having their meanings necessarily, is to take the constitutive rules as incapable of being violated, and hence as not being rules worthy of their name. However, as a matter of fact, expressions always have their meanings *contingently*, and what holds the meanings fixed is precisely our holding on the rules that constitute them, our staying within the corresponding space of meaningfulness.

Hence if Glüer and Pagin claim that "the problem with constitutive rules ... is that there simply isn't anything that they require", I see this as an illusion arising from mistaking the inside of the space of meaningfulness for a reality that was once estab-

⁴ As I think it is the actual term "rule *following*" which has given rise to the underlying confusion, I would suggest replacing it by something less confusing - perhaps "bouncing off rules"?

lished and is independent of us since. Contrary to this, I maintain that the existence of this very space is secured by the fact that the rules of language *do* require something - it is only while we accept them that the space is here in the first place.

A more radical version of Glüer's and Pagin's objection is voiced by Hattiangadi (2006). She writes (238):

For a rule to be prescriptive, it must tell me what I ought to do. According to MP [the assumption that if *t* means *F* then it applies correctly to something iff it is an *F*], the meaning constituting rule for 'horse' must imply that 'horse' applies correctly to all and only horses. However, it is not the case that I ought to apply 'horse' to all and only horses—I am not obligated to apply 'horse' to all horses because I cannot do so, and 'ought' implies 'can'. The weaker rule, stating that I should apply 'horse' only to horses cannot constitute the meaning of 'horse'. The rule that tells me to apply 'horse' only to horses does not distinguish between my meaning horse by 'horse' and something else, such as brown horse or black horse.

As follows from the above discussion, the first sentence of this argument is true only if "telling me what I ought to do" encompasses also "telling me what I ought *not* to do". And it is clear that if English is to involve a rule describable as "'horse' applies correctly to all and only horses", this cannot be construed as stating that I am to say "(lo, a) horse (!)" whenever I am confronted with a horse. Rather, such a rule would tell us that pointing at a horse is incompatible with denying "This is a horse" – i.e. that if we do the former, we ought to avoid the latter.

What, then, about Hattiangadi's claim that "the rule that tells me to apply 'horse' only to horses does not distinguish between my meaning horse by 'horse' and something else, such as brown horse or black horse"? Indeed it does not, and if this were the only rule governing 'horse', then it would indeed be indistinguishable from the more specific terms. But fortunately it is not. The inferential pattern governing 'horse' and thus constituting its meaning involves the rules that we can infer (*This is a*) horse from (*This is a*) brown horse, but not vice versa.

Hence once again, I do not disagree with Hattiangadi insofar as she claims that there are no rules of language which function as commands guiding our linguistic activities. The disagreement comes when she comes to claim that *meaning is not normative*. The trek from the claim about linguistic rules not being commands, which I take to be established, to this general thesis is long and, I am convinced, unassailable. And I do not see that Hattiangadi has even set out on it in her paper.

3 What do the normatives 'really' say?

Let us now consider an objection Hattiangadi (2003) launches specifically against Brandom's concept of normative attitudes:

[I]t is unclear how Brandom's view differs from a straightforwardly naturalistic one. The starting point is supposed to be a proto-hominid community in which there are norms, but no concepts or contents - i.e., neither propositional atti-

tudes, nor explicit thoughts. Brandom says, 'the account of norm-instituting social practices must appeal to capacities that are plausibly available in primitive prelinguistic cases, and yet provide raw materials adequate for the specification of sophisticated linguistic practices, including logical ones.' The key, according to Brandom, is to look at '*assessments* of propriety', at 'attitudes of taking or treating performances *as* correct or incorrect'. And although Brandom uses normative vocabulary to say that the proto-hominids treat each other's performances as 'correct' or 'incorrect', he suggests that they do so by way of their purely *physical* behavior and abilities.

The last sentence is odd: it suggests that besides "*physical* behavior and abilities" there would exist some other (*nonphysical? metaphysical?*) kind of behavior. However, the claim of the inferentialist is not that there would be two kinds of behavior, one physical and one not, but rather that certain very complex patterns of behavior are not usefully describable in other than normative terms, which we are unable to translate into non-normative vocabulary, i.e. especially not into the vocabulary of physics. (This is to say neither that the behavior is not physical, nor that it is in no way describable in non-normative terms - it is to say merely that no its non-normative description possesses the explanatory power which we need and which we get from the normative description.)

Hattiangadi continues that "it appears as though Brandom is offering a dispositionalist account of the determination of correctness - since the starting point includes nothing more than behavioral dispositions", based merely on "the capacity for responsive discrimination". It is true that from the 'outsider perspective' we can describe the practices of a community of norm-endorsing individuals in non-normative terms. After all, it is clear that an external observer can observe nothing else than (very complex) behavioral patterns.

But we should pay attention to what does *not* follow from this. Firstly, it does *not* follow that the 'outsider' would be able to make do with a naturalistic vocabulary in some narrow sense of the word (like the vocabulary of physics). I think that an analogue of Davidson's (1973, p. 154) argument to the effect that a 'vocabulary of agency' (Ramberg's, 2000, term) is irreducible to the naturalistic vocabulary because it serves a different aim, applies here. (Even somebody who has no idea what chess is can come to understand and explain what chess players do - but can he do so without the employment of such terms as *rules*, *error* etc. - i.e. using only a straightforwardly naturalistic vocabulary?) However, this is a problem we can safely waive here.

Secondly and more importantly, it does *not* follow that we can translate normative idiom, as employed by the insiders, into a non-normative one. When I claim that killing this cat would be wrong, I am not reporting a behavioral pattern which I instantiate: what I am doing - perhaps among other things - is endorsing a rule, and endorsing is not describing. True, an observer from outside of my community may report on this: *Peregrin endorses a rule according to which killing the cat he is pointing at would be wrong*, which can be perhaps further rendered as *Peregrin instantiates such and such kind of behavioral pattern centered around killing animals (involving not killing them, diverting others from doing so, ostracizing those who do kill one, ...)*, and now he emits a sound aiming at making his peers display the same pattern and

consequently not kill the cat he is pointing at, but nothing like this yields us a *translation* of my utterance.

And thirdly and most importantly, I can assume an 'outsider perspective' w.r.t. a set of rules and report on them from this perspective only via acquiescing in some other set of rules. (I can report on a chess game without endorsing the rules of chess; distancing myself from the rules of chess would merely make it impossible for me myself to check one of the players or to castle. But the very practice of *reporting* presupposes a language game, that is, according to the inferentialist, essentially rule-governed.) And even if we waived this problem, we must realize that we live in a world whose inhabitants are very much constructed in terms of our normative attitudes (other people are taken as *thinking, rational and responsible persons*; their antics, as *intentional deeds*; the sounds they emit, as *meaningful utterances*; ...) and as a result, using Sellars' often quoted phrase, we are in a world "fraught with ought". Trying to describe this with a purely naturalistic vocabulary would leave us with a drastically impoverished *simulacrum* of our world. (This has to do with the dialectics of what Sellars, 1962, called the *manifest image* and the *scientific image* of the world.)

The upshot is that even if we were able to give a naturalistic account of some kind of normative attitudes (from the 'outsider perspective' - a possibility I do not wish to dismiss though I have tried to indicate that it is far from straightforward), we can do so only by falling back on other normative attitudes, that keep us in the business of *giving accounts*, and indeed living within our human world.

Are normative attitudes simply cases of "responsive discrimination"? Of course they *presuppose* some abilities of responsive discrimination, and indeed they *can* be seen as cases of such discrimination. However, and this is crucial, they do more than this. We have earlier noticed how the discrimination between the *correct* and the *incorrect* differs from a more ordinary discrimination (such as the discrimination of the difficult from the easy): in the latter case, the criterion of the discrimination is independent of the person making the discrimination and of the process of discrimination, whereas in the former the process itself takes part in constituting and upholding the criterion. We have stressed that by judging something to be correct or incorrect we are not only stating the fact that it is such, we are also endorsing the rules underlying such a verdict. By every act of this normative kind of classification, we are not only classifying, but also sustaining the classificatory criterion.

Suppose that members of a tribe start to use certain kinds of sounds in certain specific ways and that they start to try to stop anybody from using the sound in different ways (say by beating her with sticks). Of course, at this point we do not yet have genuine rules, but merely regularities, i.e. something describable exhaustively in the behaviorist idiom and incapable of conferring anything like a real content on the sound. (And if we decide to call the acts of diverting members of the tribe from using the sound in certain ways 'punishments', then we are only using the word as a metaphor.)

What may happen then, however, is that what can only be called a *rule* or a *punishment* metaphorically, develops into something that *is* a rule or a punishment, hence into something that is, as such, describable only in terms of a normative vocabulary. How does this transition happen? Well, it is a kind of a bootstrapping process involving the development of language, during which the participants are gaining the ability

of classifying situations in certain ways and thereby articulating and grasping, to use Sellars' (1969) phrase, various *ought-to-be's*, while at the same time gaining the ability to engage within genuinely normative practices. (The two abilities are inextricable, for to be able to articulate situations we need meaningful language, which is a matter of advanced normative practices, whereas to have such advanced normative practices, we need language. However, bootstrapping is not circularity.) Members exhibiting truly *normative* attitudes not only divert others from doing certain things, but do it because they see it as violation of rules that, according to them, should be accepted.

4 Conclusion: why meaning is normative

Let me now summarize the sense in which meaning, from my inferentialist perspective, is normative:

1. It is inferential rules which constitute semantics. They constitute it similarly to how the rules of chess constitute *kings*, *pawns* and *bishops* – by instituting certain roles for tokens governed by them.

2. The fact that the rules constitute the meanings does not rob them of their normativity – they are genuine *constraints* that guide our behavior, preventing us from doing certain things. We must not be misled by the fact that if we decide to move 'inside' the 'space' constituted by the rules, they move out of our sight.

3. Normativity is ultimately grounded in normative attitudes; which, however, does not allow us to reduce the normative to the non-normative, but merely to reduce complex normativity to its elementary forms⁵. The impossibility of the reduction is not a matter of the fact that there would be two incommensurable strata of reality, but rather that the normative and the non-normative idioms constitute two different kinds of speech acts, and that the enterprise of accounting for what is correct, for commitments and entitlements, or for persons being responsible to each other, is incommensurably different from accounting for causal relationships.

4. The ultimate effect of the rules consists in their constrictive import opening up a brand new space for actions not previously available. Rules institute roles; and the actions that are regulated by the rules essentially involve the role-bearing items as such; they cannot be seen as some old actions in new guises.

The fact that the entities populating the space thus opened up by the system of rules and being the proper subjects of the rules are themselves *constituted* by the rules does not mean that they are unreal and can only parasite on something that is real. This is analyzed in detail by Haugeland (1998): constitution, as he puts it, must not be seen as a creation *ex nihilo*, but rather as 'letting be' – as an acknowledgment of something that, despite not being able to be what it is without us, is nevertheless independent of us.

These considerations may sound rather esoteric; but I think that chess provides us with a vivid illustration. The rules of chess that make up the game obviously open up a space of new actions and enable the players to engage in exciting activities not previously available to them. And the inferentialist idea is, analogously, that the rules of

⁵ In this respect it is a reduction similar to Davidson's reduction of his 'vocabulary of agency' to the relationship of *holding true*.

language open up the space of meaningfulness which enables us to engage in practices not previously available to us – practices of distinctively human communication, rational deliberation, building theories etc.

Hence meaning, for the inferentialist, is normative because any claim to the effect that an expression means whatever it does mean is a *normative* claim; it tells us how the expression is *correctly* used, or which *rules* it is governed by. From this viewpoint even the very claim that meaning is normative may be misleading: semantic claims do not talk about 'normative entities' attached to expressions, but instead they delimit how the given expression ought properly to be handled.

References

- Boghossian, P. A. (2005): 'Is Meaning Normative?', in Ch. Nimtz & A. Beckermann (eds.): *Philosophie und/als Wissenschaft*, Mentis, Paderborn, 205-218.
- Brandom, R. (1994): *Making It Explicit*, Harvard University Press, Cambridge (Mass.).
- Chisholm, R. M. & W. Sellars (1958): 'Intentionality and the Mental: Chisholm-Sellars Correspondence on Intentionality', in H. Feigl, M. Scriven, and G. Maxwell (eds.): *Minnesota Studies in the Philosophy of Science*, vol. II, University of Minnesota Press, Minneapolis, 521-539.
- Davidson, D. (1973): 'Radical Interpretation', *Dialectica* 27, 1973; reprinted in and quoted from *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford, 1984, 125-140.
- Glüer, K. & P. Pagin (1999): 'Rules of Meaning and Practical Reasoning', *Synthese* 117, 207–227.
- Hattiangadi, A. (2003): 'Making it Implicit: Brandom on Rule Following', *Philosophy and Phenomenological Research* 66, 419-431.
- Hattiangadi, A. (2006): 'Is Meaning Normative?', *Mind & Language* 21, 220–240.
- Haugeland, J. (1998): 'Truth and Rule Following' in *Having Thought*, Harvard University Press, Cambridge (Mass.).
- Peregrin, J. (2001): *Meaning and Structure*, Ashgate, Aldershot.
- Peregrin, J. (2008): 'Inferentialist Approach to Semantics', *Philosophy Compass* 3, 1208-1223.
- Peregrin, J. (t.a): 'Inferentialism and the Compositionality of Meaning', to appear.
- Ramberg, B. (2000): 'Post-ontological Philosophy of Mind: Rorty versus Davidson', R. Brandom (ed.): *Reading Rorty*, Blackwell, Oxford, 351-369.
- Sellars, W. (1962): 'Philosophy and the Scientific Image of Man', in R. Colodny (ed.): *Frontiers of Science and Philosophy*, University of Pittsburgh, Pittsburgh, 35-78.
- Sellars, W. (1969): 'Language as Thought and as Communication', *Philosophy and Phenomenological Research* 29, 506-527.
- Sellars, W. (1992): *Science and Metaphysics*, Ridgeview, Atascadero.

Disagreement, Error and Two Senses of Incompatibility – The Relational Function of Discursive Updating

Tanja Pritzlaff

email: t.pritzlaff@zes.uni-bremen.de

webpage: <http://www.zes.uni-bremen.de/homepages/pritzlaff/index.php>

University of Bremen, Centre for Social Policy Research, Division 'Theory and Constitution of the Welfare State', Parkallee 39, 28209 Bremen, Germany

Abstract. In *Between Saying and Doing: Towards an Analytic Pragmatism*, Robert B. Brandom puts forward a general method of formally representing relations between meaning and use (between vocabularies and practices-or-abilities) and shows how discursive intentionality can be understood as a pragmatically mediated semantic relation. In this context, the activity that pragmatically mediates the semantic relations characteristic of discursive intentionality is specified as a practice of *discursive updating* – a practice of rectifying commitments and removing incompatibilities. The aim of the paper is to take a closer look at the practice of discursive updating and to show that the inconsistencies, errors and failures in discursive practice that form the basis for the described update function can only be fully understood against the background of an agent's membership in the discursive community – i.e. if one looks at the explicitly *social* role of discursive updating.

1 Introduction

The aim of the following remarks is to take a closer look at the functional role of discursive updating – as presented especially in chapter 6 of *Between Saying and Doing* – and to further elaborate Robert Brandom's conception of discursive updating by assigning an explicitly *social* role in interactions with other agents to it.

As Brandom outlines in *Between Saying and Doing* and in a recent interview as well, inconsistencies, errors and failures in discursive practice can be described as a basis for an *update function* in the evolving structure of discursive practices. As Brandom puts it, the “role of disagreement and error, in particular the role of finding ourselves with commitments that are incompatible by our own lights, is an absolutely essential feature of the intelligibility of what we are doing, constraining ourselves by norms, that is, making ourselves subject to normative appraisal as to the goodness of our reasons for believing what we believe and for doing what we do. The principal motor of conceptual development is finding ourselves with incompatible commitments and acknowledging in practice the obligation to change something, so as to remove that incompatibility” (Pritzlaff/Brandom 2008, p. 375).

Following Brandom's conception, cases of disagreement and error involve two different kinds incompatibility. They involve incompatibility in the modal sense, i.e. that

it is impossible for one and the same object in the world to have two materially incompatible properties, and they involve incompatibilities in the normative sense, i.e. that one cannot be entitled to two commitments that are incompatible in the light of the practices and attitudes of an agent, of the norms implicit in an agent's behavior. In Brandom's conception, these two kinds of incompatibilities are answered by practices of commitment revision that integrate the subjective perspective (that focuses on the knowing and acting subjects) and the objective perspective (that focuses on the objects and state of affairs in the world), corresponding to "the subjective and objective poles of the intentional nexus between what discursive practitioners *do*, their activity of claiming, and the objects, properties, and facts that they thereby count as *saying* something *about*" (Brandom 2008, p. 200).

In expanding Brandom's conception of discursive updating by assigning an explicitly *social* role – in interactions with other agents and by explicit references to external authorities and social norms – to it, the paper also aims at a critical view on Brandom's claim that "the *activity* of taking or treating two *commitments* to be incompatible in the subjective normative sense just *is* what it is to take or treat two properties or states of affairs as incompatible in the objective modal sense" (ibid.).

2 The Practice of Discursive Updating

In chapter three of *Between Saying and Doing*, Brandom discusses arguments against two versions of AI-functionalism: Against the plausibility of the claims of the intellectualist program of classic symbolic AI, i.e. to understand knowing *how* to do something in terms of *that* something is true; and against the pragmatist thesis of a pragmatic version of AI, a thesis about understanding knowing or believing *that* in terms of knowing *how*. His argument against this second version of AI – what he calls the "substantive practical algorithmic decomposability version of AI" (Brandom 2008, p. 78) – aims to show that all autonomous discursive practices exhibit some aspect that is "not algorithmically decomposable into non-discursive practices-or-abilities" (ibid., 79).¹ In this context, the functional role of discursive updating is introduced.

For Brandom, the aspect of autonomous discursive practices that is not algorithmically decomposable into non-discursive practices-or-abilities is "the practice of doxastic updating – of adjusting one's other beliefs in response to a change of belief, paradigmatically the addition of a new belief" (ibid.). The reason why the decomposition is not possible is to be found in a characteristic of discursive updating that can be described as a kind of context sensitivity. The updating process is, as Brandom puts it, "highly sensitive to collateral commitments or beliefs" (ibid., 80). The significance of undertaking a new commitment depends not only on the content of that particular commitment. The significance of a new commitment stems from its interrelations within a network of other commitments an agent has already undertaken. To understand what a bit of vocabulary means can be characterized as knowing what difference the undertaking of that commitment would make "to what else the one using it is committed and entitled to" (ibid., 79). The addition of a new belief to an already ex-

¹ "That would be something that is PV-necessary for deploying any autonomous vocabulary (or equivalently, PP-necessary for any ADP) that cannot be algorithmically decomposed into practices for which no ADP is PP-necessary" (ibid.).

isting web or network of related commitments and entitlements requires the mastery of a practice that updates the whole set, a practice that reassigns the meaning and significance of the other, already existing elements in the light of the new belief as well. Being able to perform this practice includes being able to distinguish which further commitments would and which would not in firm or defeat an undertaken commitment, i.e. the ability to “associate with each commitment a range of counterfactual robustness”. This means that one must not only be able to identify claims that are incompatible with the new additional commitment, but claims that are incompatible with it only in the context of one’s other collateral beliefs, i.e. claims that are “contextually incompatible” with it (ibid., 80).

For Brandom, the global updating ability exhibited in the performance of this practice is an ability that cannot be assigned to non-discursive creatures. Doxastic updating requires that in the light of new information, language users are able to distinguish between information that is and information that is not relevant to the claims and inferences one endorses. Since any new information about an object carries with it new information of some kind about every other object, and any change in any property of one object changes some of the relational properties of all other objects, language users have to be able to separate contextually relevant from contextually irrelevant information. Or, as Brandom puts it: The crucially important cognitive skill that is needed to perform the practice of doxastic updating is the capacity “to *ignore* some factors one is capable of attending to” (ibid., p. 81); and the defining feature of that skill is displayed by the ability to decide *what* to ignore. As competent language users we are able to decide which aspects of a new bit of information are relevant or significant in the context of claims about objects we are concerned with and to decide which complex relational properties we should ignore in our reasoning (ibid.).

3 Two Senses of Incompatibility

In chapter 6 of *Between Saying and Doing*, Brandom focuses on the semantic relations between words and the world. The use of words is not limited to constituting relations between vocabularies. To say something consists of talking *about* something in the world, about the objects or states of affairs that the words and sentences refer to or represent.

The argument developed in chapter 6 of *Between Saying and Doing* further elaborates the complex, pragmatically mediated semantic relations between normative and modal vocabularies introduced in chapter 4 and 5. Brandom claims in chapter 6 that the intimately related features of normative vocabulary and modal vocabulary correspond to the subjective and objective poles of intentional relations, “between what discursive practitioners *do*, their activity of claiming, and the objects, properties, and facts that they thereby count as *saying* something *about*” (ibid., 200). While normative vocabulary “makes explicit important features of what knowing and acting subjects *do* when they deploy a vocabulary, when they *use* expressions so as to *say* something”, modal vocabulary “makes explicit important features both of what is *said* and of the objective world that is talked *about*” (ibid., p. 181). Normative vocabulary and modal vocabulary both articulate discursive commitments. But while normative vocabulary “addresses in the first instance *acts* of committing oneself”, modal vocabulary addresses “the *contents* one thereby commits oneself to” (ibid.), in the sense of

“how one has committed oneself to the world being, how one has represented it as being” (ibid.).

In the course of this argument, Brandom differentiates between two – related – senses of incompatibility: an objective modal sense of incompatibility (“a matter of what states of affairs and properties of *objects* actually are incompatible with what others” (Brandom 2008, p. 191)) and a subjective normative sense (that concerns *commitments* on the part of knowing-and-acting *subjects*). Brandom’s argument leads to the key point that by “engaging in the practice of rectifying commitments, subjects are at once *both* taking or treating the *commitments* involved as incompatible in the *normative* sense of obliging them to do something about that collision, *and* taking or treating two *states of affairs* regarding *objects* as incompatible in the *modal* sense that it is impossible for both to obtain” (ibid., p. 193). What is made explicit in the objective sense of incompatibility by modal vocabulary and in the subjective sense by deontic normative vocabulary are, as Brandom puts it, “essentially complementary aspects”, “connecting knowing and acting *subjects* with the *objects* they know about and act on” (ibid., p. 196). Acknowledging material inferential and incompatibility relations among commitments, therefore, essentially involves a *representational* relation to *objects*, i.e. to facts in the world that one is talking *about*, and to laws relating possible facts.

In this context, practical intentionality, identified as the most fundamental kind of intentionality, is characterized as a directedness towards objects and a practical involvement with those objects exhibited by creatures dealing skillfully with the world. The most basic form of such activity consists of “an open-ended sequence of feedback-governed performances” (ibid., p. 178). And these feedback-governed processes function only insofar as they refer in some way to changes in the world – changes one responds to or changes that are induced by responses. Feedback loops of perception-and-performance essentially involve “objects, events, and worldly states of affairs” (ibid., p. 178). *Discursive* intentionality is to be understood as a species of such feedback-governed practical engagement, as “a development of and a special case of” basic *practical* intentionality (ibid., p. 179). A creature exhibits specifically discursive intentionality insofar as its performances and ways of responding are mediated by relations of material inference and incompatibility. The directedness at objects characteristic of practical intentionality turns, as Brandom puts it, “into something intelligible as *representation* of those objects when the process of practical engagement takes the form of deontic updating structured by material *inferential* and *incompatibility* relations, that is, when it becomes *discursive* intentionality” (ibid., p. 184). And the two poles of discursive intentionality, “knowing and acting subjects and the objects they know of and act on, their representing activities and the objects and objective states of affairs they represent” (ibid., p. 179), can only be understood in the light of the semantic intentional relations they stand in one to another. For doing that, one must, as Brandom puts it, “start with an understanding of the thick, essentially world-involving practices engaged in and abilities exercised, and abstract from or dissect out of that an understanding of the two poles of the semantic intentional relations those practices and abilities institute or establish” (ibid., p. 179-180). This way to proceed is in accord with what Brandom refers to as the pragmatist order of semantic explanation.

Discursive activity involves practical engagements with things, but also the “rational critical responsibility implicit in taking incompatible commitments to oblige one to *do* something” (ibid., p. 189). And this doing, required by the incompatible observational commitments in the world as well as by the normative obligation to do something about the incompatibility of one’s own commitments, is the *practice of discursive updating*. While the first aspect becomes apparent in a modal notion, i.e. that it is impossible for an object to be made of pure copper and to be an electrical insulator at the same time, the second aspect is associated with the practices and attitudes of the subjects engaged in discursive practices. It is a matter of the “norms implicit in their behavior, what they in practice *take* or *treat* as incompatible in acknowledging and attributing the deontic statuses of commitment and entitlement” (ibid., p. 191). The practice of discursive updating, therefore, functions in two senses as a way of noting and repairing incompatibilities. It is a response to two different kinds of incompatibilities, although these two senses or kinds are, as Brandom puts it, “related in a surprising and revealing way” (ibid., 190-191).

In an objective modal sense, it responds to “what states of affairs and properties of *objects* actually are incompatible with what others, in the world as it is independent of the attitudes of the knowing-and-acting subjects of practical, feedback-governed transactional engagements” (ibid., p.191). In the normative sense, discursive updating aims, as Brandom puts it, at “the material inferential completeness and compatibility of one’s commitments, in the normative sense that insofar as one falls short of those ideals, one is normatively obliged to *do* something about it, to repair the failure” (ibid., p. 187).

4 The Social Role of Discursive Updating

Within the process of commitment revision as described by Brandom, the relational function of discursive updating is of vital importance. But while it seems clear that the need to update one’s commitment in the light of the two senses of incompatibility as described above is not a merely individual, but also a *social* requirement, Brandom doesn’t go into details on how a practice of discursive updating actually functions within the social sphere. And although a parallel between the way in which an individual subject deals with incompatible commitments and the way a community deals with incompatible commitments seems to exist, it doesn’t become clear how an updating practice may function in cases in which *different* subjects within a community contradict each other, *disagree* about properties of facts in the world or about how to proceed when faced with, for example, diverging external norms or authorities they refer to in trying to resolve a conflict. The kind of normativity exhibited in these cases seems to be not only a kind of in-process, but a kind of interpersonally established normativity. The question that follows from this is: In which sense is finding ourselves with commitments that are not incompatible by our own lights, but are incompatible in the light of other members of our community an essential feature of what we are doing? If finding ourselves with incompatible commitments and acknowledging in practice the obligation to change something, so as to remove that incompatibility, is the principal motor of conceptual development, wouldn’t it be possible to say that finding ourselves and our own commitments in conflict with the commitments of

others is a principal motor of, for example, the development of social or political norms?²

As Raffaella Giovagnoli points out, discursive updating functions in cases and contexts of conflict – not only in the subjective, but also in the social sphere. In interaction with other agents, we get into contact with points of view different from our own (Giovagnoli 2007, p. 93). Following Giovagnoli’s line of argumentation, the relational function of discursive updating as characterized by Brandom can be further elaborated, focusing on the critical dimension of its social role in practices that can be understood as *social* practices in the narrower sense of the word, i.e. as interactions with other agents. Social practices in this sense can be described as feedback-governed, as „complex patterns of mutual responsiveness“ (Rouse 2007, p. 52), embedded in a structure of “default” and “challenge” of the commitments and entitlements of different agents (Giovagnoli 2007, p. 85). In the social sphere of interaction with other agents, *performances respond to one another* through acts of correction and repair, through acts of translation, feedback loops, reward or punishment of a performer, by trying to replicate an act in different circumstances, mimicking it, and so on (Rouse 2007, p. 49). In the game of giving and asking for reasons, agents appeal to external authorities (Giovagnoli 2007, p. 83) to defend their claims, and update their beliefs not only about objective facts in the world (for example about the properties of copper (Brandom 2004, p. 250)), but also about normative facts (Brandom 2000a), for example about legal terms (Klatt 2008).

Inconsistencies, errors and failures in discursive practice that form the basis for the described update function can only be fully understood against the background of an agent’s membership in the discursive community (Kukla/Lance 2009, p. 190-195), i.e. if one looks at the explicitly *social* role of discursive updating.

Discursive practices are characterized by Brandom as special cases and developments of feedback-governed, ‘thick’ practices – ‘thick’ in the sense of “essentially involving objects, events, and worldly states of affairs” (ibid., p. 178). They function successfully if objective facts about what actually follows from and is incompatible with what are incorporated in the “material inferences and incompatibilities that articulate the concepts expressed by the vocabulary deployed according to the practical norms implicit in that practice. This essentially holistic process involves getting on to how things objectively are not just by making true claims, but also by acknowledging the right concepts” (ibid., p. 186). Even if one assumes that “the how things are is allowed to have normative significance for the correctness of someone’s sayings and believings only in the context of someone else’s *attitudes* towards how things are” (Brandom 2000b, p. 165), that is that they are filtered through the takings-true of the one assessing a claim, the central feature of a successful discursive practice seems to be its relation to facts in the world. And even if one assumes that the facts are “caught up in social practices by being endorsed by the one attributing knowledge”, so that the picture doesn’t contain a kind of “contact between naked, unconceptualized reality and someone’s application of concepts” (ibid.), skepticism about the normative word-

² Brandom himself draws an analogy between conceptual norms and political norms in Brandom 1979.

world relations implied in this conception of discursive practices seems to remain.³ At least, it seems to be a crucial point how a kind of “triangulation” (Brandom 2008: 188) on an object in the world would look like that involved the differing commitments of different subjects. The missing link in this picture of a social version of the practice of discursive updating seems to be an elaborate conception of how performances of different subjects respond to one another in a dynamic context of interaction.

5 Conclusion

The aim of the paper has been to take a closer look at the practice of discursive updating and to present some – although very preliminary – ideas about how the functional role of this practice might be interpreted in the light of inconsistencies, errors and failures occurring in the interaction of different agents. If one looks at the explicitly *social* role of discursive updating, the centrality of the relation to objects in the world – the feedback-governed aspect of the practice, characterized by essentially involving objects, events, and worldly states of affairs – seems to be problematic. The relation that needs to be focused on when thinking about discursive updating in a social sense seems to be the relation between the differing, conflicting commitments undertaken by different subjects within a community and the ways in which these differing, conflicting commitments are exhibited by performances that respond to one another in a dynamic contexts.

References

- Brandom, R. B. (1979) “Freedom and Constraint by Norms”, *American Philosophical Quarterly*, 16 (3), pp. 187-196.
- Brandom, R. B. (1994) *Making it Explicit. Reasoning, Representing, and Discursive Commitment*, Cambridge, MA/London: Harvard University Press.
- Brandom, R. B. (2000a) “Facts, Norms, and Normative Facts: A Reply to Habermas”, *European Journal of Philosophy*, 8 (3), pp. 356-374.
- Brandom, R. B. (2000b) “Vocabularies of Pragmatism: Synthesizing Naturalism and Historicism”, in: *Rorty and his Critics*, Robert B. Brandom (ed.), Malden, MA: Blackwell, pp. 156-183.
- Brandom, Robert B. 2004: “From a Critique of Cognitive Internalism to a Conception of Objective Spirit: Reflections on Descombes’ Anthropological Holism”, *Inquiry*, 47 (3), pp. 236-253.
- Brandom, R. B. (2008) *Between Saying and Doing. Towards an Analytic Pragmatism*, Oxford/New York: Oxford University Press.
- Giovagnoli, R. (2007) *Autonomy: A Matter of Content*, Firenze: Firenze University Press.
- Klatt, M. (2008) *Making the Law Explicit. The Normativity of Legal Argumentation*, Oxford/Portland, OR: Hart Publishing.

³ This skepticism also applies to the conception of *representation* as presented already in the second part of *Making it Explicit* (Brandom 1994).

- Kukla, R./Lance, M. (2009) *'Yo!' and 'Lo!'. The Pragmatic Topography of the Space of Reasons*, Cambridge, MA/London: Harvard University Press.
- Pritzlaff, T./Brandom, R. B. (2008) "Freedom is a Matter of Responsibility and Authority. An Interview with Robert B. Brandom", *European Journal of Political Theory*, 7 (3), pp. 365-381.
- Rouse, J. (2007) "Social Practices and Normativity", *Philosophy of the Social Sciences*, 37 (1), pp. 46-56.

That is, BD is extended to G, where DG = DN. Taking now the circle whose diameter is BG, it can be shown (by a chain of reasoning through the diagram that need not concern us) that the point E that is wanted is at the intersection of that circle and AC extended as needed. Already in the *Regulae*, Descartes reports that such a practice of demonstrating through diagrams “did not seem to make it sufficiently clear to my mind why these things should be so and how they were discovered” (AT X 375; CSM I 18). As he remarks in his *Geometry* of this solution in particular, “those not familiar with this construction would not be likely to discover it” (AT I 387; G 188). His own approach, we will see, is a methodical, stepwise construction that eventuates in an equation the root of which solves the problem.

We know that BD equals CD because they are sides of one and the same square. Call that length, which is given by the terms of the problem, a . Let c be the length EF, and take the length DF as that which is sought, the unknown x . CF, then, is equal to $a - x$ because, by compositionality, $CF + FD = CD$. Because the triangles BDF and ECF are similar, we know that $BF : DF :: EF : CF$, that is, that $a - x : c :: x : BF$. So, transforming this proportion into an equation, we know that $BF = cx/(a - x)$. But because BDF is right at D, we also know that $BF^2 = BD^2 + FD^2$; that is, $BF^2 = a^2 + x^2$. Combining the two equations thus gives us $(cx/(a - x))^2 = a^2 + x^2$, which we can transform in a sequence of steps licensed by familiar algorithms of algebra into the equation $x^4 - 2ax^3 + (2a^2 - c^2)x^2 - 2a^3x + a^4 = 0$. The root of this equation can then be constructed and the problem is solved.

In solving this problem, Descartes appeals to four relationships that can be expressed both graphically in a diagram and symbolically in the language of algebra: equality (of sides of a square), compositionality (of lines from their parts), proportionality (of triangles), and the relationship that is expressed in the Pythagorean theorem.⁴ And as we have just seen, the symbolic expression of these relations furthermore enables one to combine the given information systematically into a single equation and thereby to solve the problem without any need for the additional constructions that are involved in the Euclidean demonstration. In Descartes’ geometry one does not need to discover the diagram that is the medium of Euclidean demonstration but only to translate the given information into symbolic form and then combine it all into a single solvable equation. It is just this that enables Descartes to claim that he has a *method* for the discovery of truths in mathematics. Although there is no method by which to discover the diagram that is needed in a Euclidean demonstration, once we conceive the problem symbolically—the diagram not as an iconic display of objects of various sorts in relations but simply as a presentation of relations and proportions that can equally well be expressed algebraically—Descartes can show us how combine the given information symbolically into a single equation and thereby to solve the problem. As Poincaré would later remark, “before Descartes, only luck or genius allowed one to solve a geometrical problem. After Descartes, one has infallible rules to obtain the result; to be a geometer, it suffices to be patient . . .”⁵ Translation of a problem into a symbolic language the signs of which are everywhere governed by rules yields what we might well think of as an algebraic understanding of that problem.

Descartes’ mathematics is systematic, algorithmic, and symbolic. It is in these respects quite unlike the sort of diagrammatic demonstration one finds in Euclid. And yet, Descartes at first thought, mathematics needs images. As he explains in the *Regulae*, mathematics needs images because the intellect reflecting in the absence of an image can discover logical possibilities that are not *real* possibilities, that are shown to be unreal by an image one forms: “even if the intellect attends solely and precisely to what the word denotes, the imagination nonetheless ought to form a real idea of the thing, so that the intellect, when required, can be directed towards the other features of the thing which are not conveyed by the term in question, so that it may never injudiciously take these features to be excluded” (AT x 335; CSM i 61). For example, using the intellect alone one might determine that extension is not body and on that basis mistakenly conclude that there can be extension without body. Such a mistake is avoided, on the *Regulae* account, by one’s forming a real idea of an extension in the imagination and discovering on that basis that it is impossible to form an image of extension that is not also an image of body. The distinction between extension and body is *only* a distinction of the intellect; there cannot actually be extension without body.⁶ That extension requires a body

⁴ I owe this observation to Kenneth Manders. See his “Euclid or Descartes?”, ms.

⁵ Henri Poincaré, Preface to his *Oeuvre de Laguerre*, vol. I, quoted in Manders, “Euclid or Descartes?”.

⁶ This is why Descartes thinks that there can be no vacuum: “The impossibility of a vacuum, in the philosophical sense of that in which there is no substance whatsoever, is clear from the fact that there is no difference between the extension of a space, or internal place, and the extension of a body. For a body’s being extended in length, breadth and depth in itself warrants the conclusion that it is a substance, since it is a complete contradiction that a particular extension should belong to nothing; and the same conclusion must be drawn with respect to a space that is supposed to

is not logically necessary, but as the imagination reveals it is necessary nonetheless. It is, as the point might be put, not formally necessary but instead materially necessary.

According to Descartes' early account the necessary non-logical relations that are needed in mathematics, without which it would be paralyzed, are to be grounded in the faculty of imagination. When Descartes came to realize that his new mathematical practice enabled him to discover truths even in cases in which no corresponding image could be formed, he needed another account. What he came up with was the idea that God creates these necessary but non-logical truths and implants in us the ideas by which to discover them using only the pure intellect. By reflecting on our innate ideas of, for instance, substance, duration, order, and even thinking itself, one can discover by reason alone both the essence and fundamental laws of nature and the essence and fundamental laws of the mind, including the method by which inquiry ought to be conducted. Mathematics in particular thus came to be seen as at once ampliative, that is, a science within which to discover new and significant truths, and also strictly logical, that is, by means of reason alone.⁷ The pure science of the intellect that Plato had first envisaged had finally been realized.

Mathematical reasoning, according to Descartes, is at once deductive, each thing following of necessity from what came before, and also ampliative. By reflecting on one's God-given ideas and assenting only to what is clear and distinct regarding them, that is, what is clearly necessary in them, whether or not logically necessary, one discovers new truths. This, Kant would come to think, is sheer dogmatism. Although logically necessary relations among concepts are unproblematic because grounded in the law of identity, one cannot simply assume as given the non-logical necessary relations that are needed in mathematics.⁸ Of course, if the existence of God could be proven, then one could perhaps talk of (divinely) implanted ideas; but the existence of God cannot be proven and hence one cannot talk of such ideas.⁹ To talk of such ideas is to appeal to nothing more than a brute Given, as unfounded as it is unquestionable. It is this rejection of any Given, of the very idea of an indubitable foundation for our knowledge of necessary but non-logical relations, that entitles Kant to be recognized as the first pragmatist. And because his pragmatism is furthermore systematic, rigorous, and theoretical (in something like the way that the physical sciences are theoretical), it deserves to be called the first analytic pragmatism.¹⁰

But as is invariably the case with pragmatism, Kant's pragmatism can be interpreted in either of two ways. The first, and less radical, way is to take the rejection of the Given to impair in some fundamental way our capacity for knowledge of fully objective truth. On this interpretation, one leaves in place the *need* for a foundation, if knowledge of fully objective truth is to be possible, all the while denying that that need can be met. This is the pragmatism of James, Dewey, Rorty, and, we will see, Brandom. For all such

be a vacuum, namely that since there is extension in it, there must necessarily be substance in it as well" (AT VIII A 49; CSM I 229-230).

⁷ Of course it is not strictly logical in the Scholastic sense of logic, but it is, Descartes thinks, strictly logical in the true sense of logic, in the kind of logic that provides the rules that govern the direction of the mind, "which teaches us to direct our reason with a view to discovering the truths of which we are ignorant" (AT IX B 14; CSM I 186). Both Locke and Leibniz similarly held that by logic alone (as they understood logic) one might extend one's knowledge.

⁸ Thus Kant, unlike Descartes, limits what he calls general logic to what is, as we would say, formally necessary. It is, he thinks, psychologistic to include among purely logical relations those that are necessary but not logically necessary. As I argue below, necessary but not logically necessary relations are to be understood by appeal to the practices and activities of the special sciences, mathematics and the natural sciences. The transcendental aesthetic and transcendental logic serve to set out the necessary conditions of such practices and activities.

⁹ See Kant's remarks in "On a discovery whereby any new critique of pure reason is to be made superfluous by an older one" (AK 8:222).

¹⁰ That Kant's philosophy is systematic and rigorous is obvious enough. That it is theoretical is indicated, on the one hand, by the familiar contrast Kant draws between the practice of mathematics and that of philosophy, and on the other, by the fact that he thinks that the philosopher must follow the practice of the physicist rather than the mathematician. As is made clear in the B Preface, the philosopher, like the natural scientist, puts forth hypotheses and then tests them. The *Critique* itself is such a test, and a very successful one Kant thinks. But he nonetheless cautions us in the Method that "in its transcendental efforts . . . reason cannot look ahead so confidently [as in mathematics it can], as if the path on which it has traveled leads quite directly to the goal, and it must not count so boldly on the premises that ground it [as mathematics can] as if it were unnecessary for it frequently to look back and consider whether there might not be errors in the progress of its inferences to be discovered that were overlooked in its principles and that make it necessary either to determine them further or else to alter them entirely" (A735/B763-A736/B764). As we will see in more detail below, natural science, unlike mathematics on Kant's view, is constitutively self-correcting, and hence philosophy is as well.

pragmatists there is a kind of bedrock, for instance, our language and interests, or our agreements, or our social practices, that is to provide such foundations as there are for knowing, or at least our talk of knowing. To read Kant as a pragmatist of this variety is to take the forms of sensibility and understanding likewise to stand as surrogates for the foundation that is not otherwise to be had. So read Kant is, I think, clearly guilty of psychologism.

The second, more radical and I think more insightful, response to the rejection of the Given is to see not only that the Given is a myth but also that knowledge of fully objective truths does not *require* a Given, and indeed, could not be *helped* by a Given insofar as what is Given is merely a kind of blind prejudice, merely, as Peirce would say, what one is inclined to think. But if the Given, assuming there were one, could not possibly provide a foundation for knowing then it must be the foundationalist picture itself that is a myth; it must be that it is not the *products* of inquiry that are the key to knowing but instead the *process*, the dynamic *activity* of inquiry, the actual practices and processes of science. This is the pragmatism of Peirce and Sellars. To read Kant as such a pragmatist—and hence as a transcendental idealist not *because* of his pragmatism (as on the first reading) but *in spite* of it—is to take the forms of sensibility and understanding not as standing surrogate for some mythic foundation but instead as conditions of the possibility of the *activity* of inquiry. I want very briefly to indicate how such a reading of Kant would go.

According to Kant the practices and processes of science involve two very different sorts of activities, on the one hand, the constructive practices of mathematics and on the other, the theoretical, postulational, and self-correcting practices of the empirical sciences. His key insight on the side of mathematics is that although the new mathematical practice inaugurated by Descartes does not use images or Euclidean diagrams, but instead symbols and equations, it is nonetheless essentially diagrammatic. Indeed, it is the algebraic practice of successive writings and rewritings that first concretely displays the inherently temporal and dynamic character of all mathematical practice, and so even of Euclidean demonstration, which, one might think, essentially involves space for the drawing of the diagram but not time. Although the inherently diagrammatic, spatial, character of reasoning in mathematics is best illustrated in Euclidean geometry, its essentially temporal nature, the fact that it constitutively involves an organic process of unfolding over time is better illustrated in Descartes' algebra. This was Kant's deepest insight: that it is not the product but the process of mathematical proof, that is, the activity of constructing, that is the key to understanding the nature and possibility of our knowledge of the necessary but non-logical relations in mathematics; for it is this process that forges those necessary bonds among concepts that constitutes knowledge in mathematics, that "[brings] forth the truth together with the proof" (A734/B762; cf. B15-16). Space and time serve on this account not as a Given foundation but instead as the conditions of the possibility of such constructive activity.

The practice of natural science is very different according to Kant. It achieves knowledge of its laws, which like the truths of mathematics involve necessary but non-logical relations, not through constructions but through postulation and experimentation: "reason, in order to be taught by nature, must approach nature with its principles in one hand, according to which alone the agreement among appearances can count as laws, and, in the other, the experiments thought in accordance with these principles" (Bxiii). On Kant's account, our inquiries into the nature of the empirical world become *science* properly speaking only when we learn to approach nature not as a student awaiting instruction (in effect, a Given) but "like an appointed judge who compels witnesses to answer the questions he puts to them" (ibid.). Empirical inquiry is in this way constitutively self-correcting. And much as space and time are the preconditions of the constructive activities of mathematics so, Kant argues, the pure concepts of the understanding, and in particular those under the title of relation, are the precondition of self-correction in natural science.¹¹

There are three sorts of judgments under the title of relation: categoricals, hypotheticals, and disjunctives. That both categoricals and hypotheticals are needed if self-correction is to be possible is a familiar theme from Sellars, and Brandom following him. Concepts must be inferentially related if judgment, as opposed to mere classificatory behavior, is to be possible; and such relations, if they are to provide grounds for judgments rather than merely causes of them, must themselves be available as the contents of judgments. There can be categorical judgments that describe only if there are also hypothetical ones.

¹¹ The forms of judgment falling under the titles of quantity and quality are preconditions of all science, mathematics as much as physics; only those under the title of relation (and modality, but these do not concern the content of judgment) are peculiar to the practice of the natural, empirical sciences.

But although hypothetical judgments are necessary for judgment, they are not sufficient according to Kant. One needs also the capacity to judge disjunctively. We can begin to understand why when we consider the fact that not only do we correct our categorical judgments in light of our hypothetical judgments, we also can correct our hypothetical judgments in light of our categorical judgments. That is, any conflict between the two can be resolved either by giving up the relevant categorical judgment or by giving up the hypothetical one. And in most cases of conflict it will be reasonably clear which makes more sense. *But there is no guarantee of this.* In some cases it will seem that *none* of the conflicting claims can reasonably be jettisoned. In such cases, what needs to be revised is not merely a categorical or a hypothetical but one's conception of what is possible, what makes sense at all.

The ancient problem of change provides a simple illustration of the point. Change is pervasive in our lives, but before Aristotle it seemed that change is impossible insofar as it cannot come from what is or from what is not. We have an inconsistent triad: (1) there is change; (2) change cannot come from what is (because it already is); (3) change cannot come from what is not (because nothing comes from what is not). Each of the three claims taken alone seems compelling. That there is change seems manifest in experience; and that nothing can come either from what is (because it already is) or from what is not (because nothing comes from nothing) also seem to be obviously true. The conflict cannot be resolved, so it seems, by rejecting *any* of our three claims.

Hume's skeptical challenge is essentially similar. We have (we think) knowledge of causal relations. Hume shows that this knowledge cannot be either a relation of ideas, known by reason alone (because it is not logically necessary), or a matter of fact based on experience (because though not logically necessary, it is nonetheless necessary). But surely, Hume thinks, such knowledge could only be based on the one or the other. And the same is true of Kant's antinomies. Indeed, antinomy is at the heart of the whole of Kant's critical enterprise insofar as the reality on which thought aims to bear both *must* rationally constrain judgment, if judgment is to be rational at all, and *cannot* rationally constrain judgment insofar as reality has Kant thinks no inherently normative significance.

In cases involving only categoricals and hypotheticals either experience contradicts some purported inferential connection among concepts or some conceptual connection leads to a contradiction among putative experiences. In the cases of concern here we end up contradicting *ourselves* insofar as none of the available options seem to make any sense. We need, then, a third conception of judgment, and it must be just the sort of disjunctive judgment Kant provides. Suppose that categorical and hypothetical judgments were the only two forms of judgment involved in inquiry, that they were together sufficient to ground our second thoughts. There would in that case be no guarantee that there would not be a stalemate between them, that is, a case in which we find ourselves wanting to affirm that A and that not-B and that A entails B. And yet there *cannot* be an irresolvable stalemate for reason. Although it would be dogmatism to assume that we have already in hand the conceptual resources that are needed to understand reality as a whole, it would be skepticism to deny that we can achieve such resources as are needed. To avoid dogmatism we need to recognize that we may well find ourselves with an inconsistent triad that we simply do not know how to resolve by the usual means; to avoid skepticism we need a third form of judgment enabling us to express the difficulty in the form of a judgment and so reason about it. That third form of judgment is precisely what Kant gives us, a disjunctive judgment of the form A is (B or C), where the sphere of possibilities is exhaustive. By making our current understanding of all the possibilities explicit in this way, we are at the same time put in a position to come to understand the new possibilities that will resolve our difficulties. Kant's table of the pure concepts of the understanding establishes in this way the preconditions for our second thoughts and thereby the science of nature.

According to the reading just outlined, Kant's pragmatism involves rejecting not only any Given but also the whole foundationalist picture on which it depends. What is set in the place of the foundationalist picture is the practice of scientific inquiry, specifically, the activity of construction in mathematics and of hypothesis, experiment, and self-correction in natural science. Two developments in the sciences after Kant would prove critical to the development of Peirce's Kantian analytic pragmatism, first, the emergence over the course of the nineteenth century of yet another form of mathematical practice according to which theorems are to be proved by strictly deductive, logical reasoning from definitions alone, and also a new sort of physics that does not merely use mathematics but to a surprising extent just is mathematics. For our purposes it is the first of these, the new mathematical practice of deducing theorems on the basis of definitions, that is of primary importance. What it shows is either that Kant was wrong about the constructive nature of mathematics or that he was wrong about the non-constructive character of purely logical reasoning. Most, including Brandom, assume that the first is correct; they jettison Kant's philosophy

of mathematics while embracing his conception of logic as formal. Peirce argues that it is the second that is correct, and he does so on the basis of his analyses of reasoning, analyses that, he claims, “surpass in thoroughness all that has ever been done in print, whether in words or in symbols—all that De Morgan, Dedekind, Schroeder, Peano, Russell, and others have ever done [notice that Frege is not mentioned],—to such a degree as to remind one of the difference between a pencil sketch of a scene and a photograph of it”.¹² What these analyses reveal, according to Peirce, is that “all necessary reasoning . . . is mathematical reasoning,” that it is diagrammatic (*ibid.*). It follows that a good logical notation, like a good mathematical one, enables one to reason in the notation itself; just as on Kant’s own account of the notations of mathematics, a good logical notation ought to enable one to exhibit logical relations in a way facilitating the discovery of new truths. (I have argued elsewhere that Frege’s notation is good in just this sense.¹³) If, now, we combine that insight with Kant’s insight into the practice of the natural sciences, in particular, his insight into the role of self-correction in that practice, we thereby give up the last semblance of foundationalism in mathematics. Although not the whole point, one constitutive role of axiomatizing, rigorously defining, and deductively reasoning in mathematics, and indeed in logic as well, is to reveal errors in our conceptions and to facilitate thereby the process of self-correction. It is just as Peirce says: “one of the most wonderful features of reasoning” is the fact that “it not only corrects its conclusions, it even corrects its premises”.¹⁴

For Peirce as for Kant “the data for the generalizations of logic are the special methods in the different sciences”; “to penetrate these methods the logician has to study various sciences rather profoundly”.¹⁵ Their analyses are undertaken in order to achieve an understanding of the actual practice of science, the striving for truth. And because they are, the variety of analytic pragmatism Kant and Peirce espouse is quite unlike that we find in Brandom’s Locke Lectures. First, and most obviously, Brandom develops and pursues his version of analytic pragmatism not as a logician but as an analytic philosopher of language; his aim is not so much to *understand* the practice of modern mathematics and the modern mathematical sciences as it is to *emulate* these sciences in order to achieve thereby the sort of understanding that he thinks they achieve. But there are other differences as well. I will consider three that relate to their various understandings of the nature of algebraic understanding, and also a fourth that is more global and perhaps the most important of all.

According to Brandom, algebraic understanding is the understanding that is achieved by “*constructing* the conceptual contents expressed by a target vocabulary” and “it does that by exhibiting them as complexes formed as the products of applying explicit algorithms to the conceptual contents expressed by a base vocabulary” (p. 213). This, he thinks, is not merely a good thing but the “gold standard” of understanding insofar as “it takes the issue of what one means . . . out of the hands, out from under the authority, of the one making the claims. It establishes a fact of the matter about the inferential relations that articulate the contents of the concepts expressed by the target vocabulary that swings free of the beliefs and preferences of the concept user” (p. 214). The idea finds an echo in Peirce’s claim that “the very first lesson that we have a right to demand that logic shall teach us is, how to make our ideas clear . . . to know what we think, to be masters of our own meaning”.¹⁶ But for Peirce making one’s ideas clear, being master of one’s own meaning, is not the intrinsic good that it is for Brandom. Whereas Brandom takes clarity, in particular, conceptual clarity, to be a cardinal virtue, for Peirce such clarity is important primarily as a means: to make one’s ideas clear is the first step on the way to improving them. Brandom does envisage a use for clarity insofar as it more easily enables us to settle disputes as they arise; nowhere does he suggest that such disputes might require us to revise and correct the ideas, however clear, with which we began.

A second, closely related difference is that Brandom’s account, unlike Peirce’s, is essentially static. For Brandom algebraic understanding consists in being clear about some *thing*; his focus is on the products of understanding. Peirce is concerned rather to understand the processes of understanding, and as we have seen, this focus on the activity of inquiry is a direct consequence of his rejecting, with Kant, not only a

¹² Charles Sanders Peirce, “The Three Normative Sciences” (1903), in *The Essential Peirce: Selected Philosophical Writings, II: 1893-1913*, ed. N. Houser and C. Kloesel (Bloomington and Indianapolis: Indiana University Press, 1998), p. 206.

¹³ See my *Frege’s Logic* (Cambridge, Mass.: Harvard University Press, 2005).

¹⁴ Charles Sanders Peirce, *Reasoning and the Logic of Things: The Cambridge Conference Lectures of 1898*, ed. Kenneth Laine Ketner (Cambridge, Mass.: Harvard University Press, 1992), p. 165.

¹⁵ Letter to Daniel C. Gilman quoted in the Introduction to *The Essential Peirce: Selected Philosophical Writings, I: 1867-1893*, ed. N. Houser and C. Kloesel (Bloomington and Indianapolis: Indiana University Press, 1992), p. xxix.

¹⁶ Charles Sanders Peirce, “How to Make Our Ideas Clear” (1878), in *Essential Writings I*, p. 126.

Given foundation but also the whole foundationalist conception of knowing. The difference between Peirce and Brandom on this point is, furthermore, reflected in the notation each aspires to. Brandom's diagrams display various relationships that can obtain among vocabularies and practices-or-abilities. Like the diagrams developed by Euler and Venn, Brandom's diagrams utilize the two-dimensional space of the page graphically to present relations. What Brandom's graphs do not enable one to do is to reason *in* the diagram. Of course one can, given a particular meaning-use diagram, draw inferences about the relations depicted; what one does not, and cannot, do is draw those inferences graphically. And the reason one cannot is that, as Kant's study of mathematical practice led him to see, a notation within which to reason must involve three levels of articulation: primitive signs, wholes of those signs in relation, which in some way present the entities of interest, and finally wholes of these (intermediate) wholes of primitives.

Think, for instance, of calculation in Arabic numeration. The primitives are the ten digits and the intermediate wholes are the numerals we form by concatenations of those primitives. These intermediate wholes are then arranged in a particular way on the page depending on the calculation one wishes to perform, say, addition or subtraction, multiplication or division. Having set up one's numerals appropriately, one then (perceptually) reconfigures the display in various ways: considering now one primitive sign from one numeral and one from the other, one performs a calculation, writes the result in the appropriate place, and goes on in a familiar stepwise fashion until the desired result is obtained. Similarly in Euclidean geometry, there are, first, the primitives (points, lines, angles, and areas), then the wholes composed of those primitives, that is, the familiar figures (circles, triangles, and so on) that make up the subject matter of Euclidean geometry, and finally there are the diagrams themselves within which can be discerned various intermediate wholes depending on how the diagram is regarded. I have argued elsewhere that it is just this feature of Euclidean diagrams that enables them to constitute ampliative demonstrations.¹⁷ And it is true in general that the extraordinary power and fruitfulness of such notations as vehicles of reasoning rests on just this feature. Because Brandom's notation has only two levels of articulation, only primitive parts and wholes of those parts, it is not and cannot be a medium of reasoning. Nor, of course, is it intended to be. What matters for Brandom's purposes is the display of relations, not any activity an inquirer might engage in.

Perhaps it will be objected that Brandom *does* concern himself with activities and processes; after all, his *is* a form of pragmatism. He is concerned with use as much as with meaning. The crucial question, however, is *how* he is concerned with use. As Brandom reads Wittgenstein, that use underlies and so is prior to meaning is shown by the fact that usage and consequently meaning can and does change (see p. 6). Thus if we want to codify meaning, say, in order to create a dictionary, we must first look to how words are actually used by (native) speakers of the language. As the point is developed in *Making It Explicit*, explicit rules codifying meanings essentially depend, on pain of a vicious regress, on rules implicit in practice.¹⁸ The regress of interpretations is to be stopped by one's implicit grasp of a rule as manifested in one's practice. As should be clear, this is the version of pragmatism that jettisons the mythic Given while keeping in place the foundationalist picture overall. The *need* for a Given remains and is here met by appeal to practices.

But one can also read Wittgenstein differently, as rejecting not only the Given but also the whole static foundationalist picture on which it depends.¹⁹ Much as an animal is an instance of a form of life, where a form of life is characterized by particular capacities, perceptual and motor, and a narrative of growth, development, and characteristic behavior, so, on this reading, a person going by a signpost (say) is an instance of a (now socially rather than biologically evolved) form of life characterized by various activities and abilities, among them the ability to see and so to follow the way a signpost points. Much as mere stuff can acquire, in the course of the evolutionary emergence of animals, the significance of food, that is, of being (appropriately or normally) nourishing for an animal of a certain sort, so what is otherwise a mere stuff can acquire, in the course of the social evolutionary emergence of rational animals, the significance of being a reason for us. The signpost itself, in such a context, can tell us the way to go on. In place of the foundationalist picture, we have on this reading the picture of a kind of animal with its characteristic capacities and activities, as an instance of a form of life that is itself a narrative of birth, growth, and finally

¹⁷ See my "Diagrammatic Reasoning in Euclid's *Elements*", forthcoming.

¹⁸ Robert B. Brandom, *Making It Explicit: Reasoning, Representing, and Discursive Commitment* (Cambridge, Mass.: Harvard University Press, 1994).

¹⁹ See John McDowell, "How Not to Read *Philosophical Investigations*: Brandom's Wittgenstein", reprinted in *The Engaged Intellect: Philosophical Essays* (Cambridge, Mass.: Harvard University Press, 2009).

death. The activities of life itself, and in particular our own form of life as the rational animals we are, thus play a constitutive role in this account that is altogether different from that they play on Brandom's reading.

This difference between the two varieties of pragmatism it is possible to read into Wittgenstein is further reflected in the fact that Brandom's analytic pragmatism, by contrast with that of Peirce, inevitably needs a given vocabulary that provides, in effect, the bricks and stones with which to build its structures. As Brandom himself puts the point, "every context in which it [algebraic understanding] is available contains an appeal to a base vocabulary whose use is *not* held in place algebraically, but depends on *another* sort of practical mastery and understanding" (p. 215-6), one that Brandom thinks of as hermeneutic. Because Brandom conceives analysis as a matter of constructing contents by "exhibiting them as complexes formed as the products of applying explicit algorithms to the conceptual contents expressed by a base vocabulary (treated for this purpose and relative to this construction, as simple)" (p. 213), he simply must have some materials given at the outset. The task is that of conceptual engineering, of building structures out of given materials. On a more Kantian and Peircean version of analytic pragmatism what analysis reveals is not structure or form for an antecedently given content, but instead a dynamic, organic process of intellectual inquiry and growth mediated, at least in many instances, by a symbolic, or as Peirce would say, diagrammatic, language, one that is fully meaningful in its own right and has not two but three levels of articulation.

There is, finally, a very striking difference in the overall orientation of Peirce and Brandom that is not, I think, accidental to the varieties of analytic pragmatism they espouse. Because for Peirce the central concern of the philosopher is actual scientific inquiry, which is first and foremost constitutively self-correcting, Peirce's first, and in a sense only, rule of reason is "that in order to learn you must desire to learn and in so desiring not to be satisfied with what you already incline to think".²⁰ To learn, and so to be a true scientist in Peirce's sense, one must be devoted to "the pursuit of truth for truth's sake", where this devotion is, Peirce thinks, a kind of existential commitment, "a mode of life", "the devoted, well-considered, life-pursuit of knowledge".²¹ Peirce describes his own philosophical ideas as the fruits of just such a life, as growing "out of a contrite fallibilism, combined with a high faith in the reality of knowledge, and an intense desire to find things out".²² Brandom's orientation, inspired by the example of David Lewis, is very different. On this approach "what philosophers should do is lay down a set of premises concerning some topic of interest as clearly as possible, and extract consequences from them as rigorously as possible. Having done that, one should lay down another, perhaps quite different set of premises, and extract consequences from them as rigorously as possible" (p. 225). What Brandom finds "liberating and exhilarating about this metaphilosophical attitude is that Lewis . . . didn't care much what reasons one had for starting with one set of premises rather than another. He was entirely open to, and indeed eager to, turn his awe-inspiring intellect to following out the consequences of even the wackiest of claims" (p. 226). It is in just this spirit that Brandom would have us take up his work in the Locke Lectures (ibid.). For Peirce, although "there is no positive sin against logic in *trying* any theory which may come into our hands", nevertheless "it is better to be methodical in our investigations"²³; and this is better because "the more voraciously truth is desired at the outset, the shorter by centuries will the road to it be".²⁴ If one desires truth above all else then Lewis cannot be one's model.

Brandom's variety of analytic pragmatism aims to extend the project of analysis that is exemplified in such work as Russell's theory of descriptions to embrace not only meanings but also use. Because pragmatism, in particular the pragmatism of the later Wittgenstein, is for him simply the idea that there is no Given foundation for meaning, his variety of analytic pragmatism is and must be limited to feats of construction between bits of vocabulary. But perhaps a further step is needed as well. If, as I have argued, the Given could not possibly found either meaning or truth and knowledge insofar as it would be utterly unquestionable and hence merely what we are inclined to think (which is a matter for psychology, not

²⁰ *Reason and the Logic of Things*, p. 178.

²¹ Charles Sanders Peirce, "The Century's Great Men in Science" (1901), originally appearing in the *New York Evening Post*, quoted in the Introduction to *The Essential Peirce*, II, pp. xxii-xxiii.

²² Charles Sanders Peirce, "Concerning the Author" (1897), in *Philosophical Writings of Peirce*, ed. Justus Buchler (New York: Dover, 1955), p. 4.

²³ *Reasoning and the Logic of Things*, p. 178.

²⁴ *Reasoning and the Logic of Things*, p. 170.

philosophy), then it is the whole static foundationalist picture that must be given up, replaced with just the focus on the activities of inquiry in the sciences that we find in Peirce and Kant.²⁵

But although these two varieties of pragmatism are in this regard very different, they are nonetheless alike in embracing the need for systematic theorizing in philosophy, in eschewing the sort of quietism that Wittgenstein can seem to espouse. As the analytic pragmatism of Kant and Peirce furthermore shows, the idea of an analytic pragmatism, which can easily seem oxymoronic, is not inherently problematic. Kant, the first pragmatist, certainly did not eschew philosophical theorizing, the forming and testing of hypotheses, and nor did Peirce; for both, such a “scientific”, and so inherently self-correcting, attitude is constitutive of the practice of philosophy. Indeed, one could say that, for them, to be a pragmatist *just is* to be an analyst insofar as the method of their variety of pragmatism is critique, the task of self-scrutiny that reason can and must take up in the aftermath of dogmatism and the inevitable skepticism that follows it. Brandom’s variety of analytic pragmatism is not so easily seen as a real and organic unity. It may well be that in the end his variety is just as John McDowell has described it, the project of (in Brandom’s words) “perversely transplanting perfectly healthy pragmatist organs into the rotting corpse of analytic philosophy” (p. 202). I leave this question for others to decide.

²⁵ As Sellars puts it, “science is rational, not because it has a *foundation* but because it is a self-correcting enterprise which can put *any* claim in jeopardy, though not *all* at once.” Wilfrid Sellars, “Empiricism and the Philosophy of Mind”, in *Science, Perception and Reality* (London: Routledge and Kegan Paul, 1963), §38, p. 170.

Is Logic Demarcated by its Expressive Role?

Bernard Weiss

1. How and Anti-realist might read Brandom

Michael Dummett and Robert Brandom, though sharing a good deal in their approaches to language and to logic, also differ markedly in their respective views. Although that is not an observation that is likely to strain one's capacity for philosophical insight, the differences are worth noting and understanding.

Brandom distinguishes logic by means of its expressive role. Logic plays a role in enabling the expression of the inferential commitments of any practice as claimings. Or in Brandomese: since inferential practice is PP-necessary for any autonomous discursive practice, the role of logic is in articulating those practices which are necessarily involved in being able to say anything. Logical vocabulary is deployed by practices that are algorithmically elaborated from practices that are PV-necessary for deploying every autonomous vocabulary and suffices to specify those PV-necessary practices. In the earlier work the role of logic is linked with the project of making explicit, in the later work it is linked with the project of analytic pragmatism. In *Making it Explicit* logic enables the expression of inferential commitments as claimings and thus as subject to the business of asking for and giving reasons. It thus exposes those commitments to the glare of reflective rationality. In the later *Between Saying and Doing* the interest shifts to resolving what Brandom calls the logicist's dilemma, namely, explaining logical vocabulary in a way that ensures it does not contribute any significant *content* to the analyses in which it is caught up—that it is semantically transparent—while also ensuring that it is analytically efficacious—it must make a distinctive contribution to the *process* of analysis. The resolution is that logical vocabulary adds nothing to the, that is, any, target vocabulary because the capacities involved in deploying it emerge by algorithmic elaboration from the capacities required to deploy *any* vocabulary. It is analytically efficacious because it fulfils an expressive function enabling one to talk about the analytic articulation of any vocabulary and this—the ability to articulate what follows from what—is an essential part of being able to find expressions in one vocabulary which express the same contents as expressions in another. Though there are interesting differences between the accounts¹ the upshot is similar: in order to perform its expressive function logic is required to be semantically transparent; however logic still has a purpose because that expressive role feeds into the business of analysis or the business of reflecting on one's inferential practice.

Michael Dummett poses a different, though in some respects similar, dilemma. His focus is not so much logical vocabulary as logical inference. He notes that our attempts to justify deduction pull in two directions: in an effort to see deductive inference as valid we tend to think that, in some sense, whatever is required to recognise the conclusion as true is already accomplished in recognising the premises as true; conversely we are tempted to think that there must be some gap here else we will have no way to account for the epistemic usefulness of deduction. As in the Brandomian dilemma the tension arises from both wanting to see logic as, in a sense, vacuous and as having a purpose. But unlike Brandom's way of dealing with the tension, which promises complete resolution, Dummett sees the tension as irresolvable: we can only conclude that deductive inference forces us to admit a gap between the truth of a sentence and its capacity to be recognised as such (at least by direct—non-inferential means). So on Dummett's view, the admission of deductive inference into a practice necessitates a conception of contents expressible in that practice which would not have been required but for deductive inference. Thus whereas Brandom sees logic as being semantically transparent Dummett thinks that deductive inference places a metaphysical demand on content. The demand is metaphysical since the reconception of content derives from the very nature of deductive inference and involves construing content in terms that tend to be favoured by realists, though it doesn't demand a *fully* realist reconstrual of content.

I want to spend some time here simply exploring logic in the framework provided by these two dilemmas.

¹ The latter envisages as distinctively theoretical role for logic and along with this is concerned to develop different sorts of metavocabulary. The former sees things from a much less theoretical perspective and thus emphasises object language extensions; rather than metalanguages.

Conservative Extension

The notion of conservative extension is crucial to both philosophers. For Dummett the requirement is that logic conservatively extend the non-logical practice in the sense that the meanings of terms in the original non-logical fragment are unaffected by the extension of the language to include logic. Thus if we take meaning to be determined by, say, assertion conditions then the insistence will be that no expression in the extended language becomes assertible, if it had not been so before. The requirement is, as we have seen, part of Dummett's conception of how one validates deductive inference but it is also of a piece with his view of the molecularity of language: in order to see mastery of language as accruing in stages we need to be able to think of each stage as establishing a stable set of meanings which survive the introduction of more complex reaches of language.

On Brandom's account the expressive role of logic demands that its introduction constitutes a conservative extension. Since the content of an expression is determined by its inferential role and since we want, in the logical language, to be able to express inferential relations that that expression bears to other expressions we had better not alter the content of the expression, and thus its inferential powers in relation to the old vocabulary, by introducing logical vocabulary. [In addition, if our aim is to articulate just those inferential patterns which are taken to be good in the original vocabulary we had better be sure that logical vocabulary does not forge any new such connections.] Brandom thus achieves a kind of local molecular view in the context of a global holism about content.

I'm not certain that the feat is pulled off quite so easily—not certain, that is, that logical expressivism requires inferential conservativeness. After all, if content is determined by inferential role conceived of holistically then it is not clear that expressions in the logical language would bear *the same* content as they do in the pre-logical language, although obviously the translation of expressions of the original language into the extended language would be homophonic; perhaps the requirement should rather be that the introduction of logic be such as to preserve the homophonic scheme of translation. [And, on the second sort of argument, which aims to ensure that new inferential connections are not forged, so that we aren't confused about which connections we are articulating—it is not clear that there is a problem that needs to be avoided by insisting on conservative extension. Given that the same material inferential connections survive into the extended language we simply need to distinguish between those inferences which explicitate pre-logical inferences and those which do not. Insisting on conservative extensions simply obviates the need for marking this distinction.]

A further point is that the expressive role might be accomplished not through an extension of the language but through a metalanguage. In this case the preservation of content would be accomplished very simply indeed, namely, by using a metalinguistic expression to refer to expressions in the object language. Of course the metalanguage would contain logical vocabulary enabling the expression of entailment relations but I see no reason why and little way to comprehend the requirement of conservative extension; rather we would face a choice about our logic justified presumably in orthodox fashion by appealing to the semantics of the object language. Well perhaps one might simply concede that that metalinguistic project is legitimate but simply not Brandom's and perhaps too we should allow him to choose his project: the distinctive feature of logical vocabulary is thus in facilitating the expression of those inferential relations which obtain in any language *through an extension of that language*. In order to achieve this the insertion of logical vocabulary must be conservative relative to inferential relations in the language.

Let us return to Dummett's account. Matters there were left considerably vague since we simply noted that for him, the extension of the language to include logical vocabulary needs to be conservative relative to the content of terms in the original language. As is quite obvious, the notion of conservative extension is relative and, in taking the relativisation simply to be relative to *a* conception of content, we haven't succeeded in making the notion precise. But it can be made so by specifying a conception of content. Dummett's thought is that deductive inference demands a conception of content which is not a product merely of the use of the original language. If we look at that language we will be unable to discern any reason to justify taking a sentence to be assertible other than when a warrant is actually available for its assertion. As soon as deductive inference is admitted we shall need to think of sentences as being assertible merely when a warrant is, *in principle*, available. Thus deductive inference demands a certain conception of content. Dummett writes,

The relation of truth to the recognition of truth is the fundamental problem of the theory of meaning, or, what is the same thing, of metaphysics... What I am affirming here is that the justifiability of deductive

inference—the possibility of displaying it as both valid and useful—requires *some* gap between truth and its recognition; that is, it requires us to travel some distance, however small, along the path to realism, by allowing that a statement may be true when things are such as to make it *possible* [my emphasis] for us to recognise it as true, even though we have not accorded it such recognition.^{2,3}

An example: in order to see disjunction as having an epistemic function we need to see a disjunction as assertible when we have a method, effective in principle, for determining which disjunct is true. Thus, if the truth of the disjunction requires the truth of one or other disjunct, a sentence may be true when it is only in principle possible to know its truth.

Another example: the technical appendices to *Between Saying and Doing* contain a deduction of classical logic. The basis for the deduction is an incompatibility relation defined on finite sets of sentences which is determinate on the finite power set of sentences in the language⁴. Clearly detection of these incompatibility relations will be something we are only *in principle* capable of doing. It is, of course, open to Brandom to allow the incompatibility relation to be less than fully determinate and thus to achieve a weaker logic. But Dummett's point is that if the logic is to be epistemically useful we will need an incompatibility relation that is determinate even when it is only, in some sense, *in principle possible* for us to detect its obtaining.

Thus, on Dummett's view, we are forced by the need to validate deductive inference—and its various locutions—to admit a notion of content according to which a sentence is true just in case it is in principle possible for one to obtain a direct warrant for it. But, off course, *relative to that conception*, deduction will be conservative. So is there a tension here with Brandom's view?

There are actually two worries one might have with this juxtaposition of Dummett and Brandom. The one just alluded to is that the Brandomian view of logic providing a conservative extension is restored on the final Dummettian view. The second—one that may have been troubling my audience for some time now—is that Dummett's account focuses on deductive inference not directly on logic. Brandom refuses to consider inference-free regions of language.

I don't think that the second worry ought to detain us long—the sorts of inference which concern Dummett are inferences which essentially involve logical vocabulary. Indeed it seems that the very phenomenon we are concerned with requires the formality of logic which, precisely in view of that formality, extends beyond our mere parochial doings; to see an inferential scheme as formally valid is to see it as having some generality of application and capturing that generality requires logical machinery. Put in more Dummettian terms, we are interested in the possibility of achieving indirect warrants for assertion of a statement by means of logical inference. We can distinguish between these indirect warrants and direct warrants, where the latter may include *both* non-inferential warrants and warrants accruing through materially good inferences.

² T&OE 314

³ Later Dummett castigates realism for forcing a conception of truth on us which compromises molecularity in that it is entirely unjustified relative to the use of the pre-logical fragment of language. A realist notion of truth is thus implausibly imported simply to justify classical modes of inferring. Just how is the present position disanalogous? Two points are worth noting: (i) the gap between truth and its recognition is a product of the need to validate deductive inference, not specific modes of inferring, or it's a requirement on seeing a locution such as disjunction as having any role; (ii) the conception of content is built upon the use of the sentence in relation to its direct warrants: the indirect warrant is explained in terms of the in principle availability of its direct warrant.

⁴ The deduction of classical logic is impressive, at first sight, given such slim (intuitionistically acceptable) assumptions but shouldn't really be surprising. What we have is a definition of the conditional in terms of incompatibility as follows:

$$p \rightarrow q \text{ iff } (\forall x)(\text{ if } q/x \text{ then } p/x)$$

But a more orthodox intuitionistic reading of the conditional would be:

$$p \supset q \text{ iff } (\forall x)(\text{ if } W(x,p) \text{ then } W(x,q)), \text{ ' } W(x,p) \text{ ' is ' } x \text{ warrants } p \text{ ' .}$$

What we then have is:

$$\begin{aligned} & p \rightarrow q \\ & \text{ iff } (\forall x)(\text{ if } q/x \text{ then } p/x) \\ & \text{ iff } (\forall x)(\text{ if } \text{In}\{x,q\} \text{ then } \text{In}\{x,p\}) \\ & \text{ iff } (\forall x)(\text{ if } W(x,\sim q) \text{ then } W(x,\sim p)) \\ & \text{ iff } \sim q \supset \sim p \end{aligned}$$

The seemingly distinctively classical ' $\sim p \rightarrow p$ ' becomes the intuitionistically acceptable ' $\sim p \supset \sim p$ ' [and ' $p \rightarrow \sim p$ ' becomes the intuitionistically acceptable ' $\sim \sim p \supset \sim p$ '].

The first worry should also be dismissed. Insisting on conservative extension is indeed a very weak requirement—see Field (*Science Without Numbers*)—one which seemingly any logic that is capable of being seen as good will satisfy. (To see this simply note that if one takes content to be determined by truth conditions then conservative extension relative to content so construed simply amounts to soundness.) Brandom's point should thus be seen as making a more substantial claim: logic is introduced subject to the constraint of conservative extension relative to a conception of content that can be substantiated independently of the requirements of logic. The difference between Dummett and Brandom lies precisely here. According to Brandom logic makes no demands on content; according to Dummett it does.

The difference is important and takes us to a deep difference in the respective conceptions of logic and its philosophical importance. In Brandom's view the fact that logic makes no demands on content entails that it can function as a neutral medium for various programmes of analysis: its very neutrality, its lack of metaphysical substance is what renders semantic logicism plausible. Logic is seen as a tool for analysis. In contrast, on Dummett's view logic just is the crystallisation of metaphysics. Indeed it wouldn't be unfair to say that for Dummett the question of one's choice of logic gives the operational content of the metaphysical question. Logic enables the expression of certain sorts of distinctive complex contents and the space of those possible complex contents encapsulates a metaphysical view about the contents which form the base.

Perhaps the most promising line for a Brandomian to pursue is to argue that we have quite independent reason for discerning content which satisfies logic's requirements. So its demands on content would be vacuous. The place to look is the various objectivity proofs to which our attributions of content are subject.

Objectivity Proofs

Deep into *Making it Explicit* we find Brandom attempting to validate his pragmatic conception of content by showing that content, so conceived, is objective. The objectivity of content emerges from structural elements of the social deontic score-keeping practice. Not only can we *distinguish* a content both from the claim that I know that content and from the claim that everyone knows it, but we can also refute any conditional linking these contents. The business of keeping track of perspectives through such means as *de re* and *de dicto* ascriptions of propositional attitudes enforces a conception of another's perspective which may differ from the truth, as one takes it to be. So, in general, the articulation of the sociality of content ascription into an *I-Thou* sociality allows for the distinction we want, without appealing to an extra-perspectival reality.

The first objectivity proof allows for communal ignorance and communal error and is based on the articulation of the practice of ascribing commitment and entitlement. One can both be committed and entitled to assert a content without seeing anyone else as committed and entitled to assert that content. The second objectivity proof seemingly has a more difficult task because one has to make a similar distinction in relation to oneself. But here the articulation of deontic statuses into commitments and entitlements plays a role: one may be *committed* both to 'p' and 'I believe p' yet not be *entitled* to both and this difference in commitment can be revealed by considering a third person's evaluation of the my situation, such a character may be both committed and entitled to:

S claims that I believe p and S claims that not p.

The objectivity proofs are impressive in their ability to reconstruct features of objectivity that one typically associates with representationalism but to do so by means of the structural relations between perspectives.

What I don't see is how such proofs can be made to work productively in the face of Dummett's dilemma. How can a claim about the structural features of keeping score on perspectives entitle any view about the determinacy of incompatibility relations or the truth of claims where these outrun our actual ability to determine them? Because our actual abilities are exceeded no license for a claim of determinacy can be arrived at from looking at an exercise of those abilities. It is one thing to deny a connection between truth and one's own or our communal recognition of truth but quite another to make the positive claim that truth outruns what we are actually capable of recognising.

If there is a way out of this bind I think it will lie in the notion of algorithmic elaboration since the very notion of elaboration incorporates an idea of that certain capacities *in principle* suffice for development of

others. Speakers will have the ability to determine whether certain sets of sentences are incompatible with one another and we might algorithmically elaborate this ability to yield an ability to tell whether or not any two finite sets of sentences are incompatible. But how could we do this? As Brandom points out, a sentence may be incompatible with a set of two sentences without being incompatible with either: take ‘This is a blackberry’ and {‘This is red’, ‘This is ripe’}. The result is surely general: one set of sentences may be incompatible with another without there being any incompatibility between subsets. In fact that ability to detect incompatibilities between *sets of sentences* is not algorithmically elaborated from any more basic set of abilities, such as the ability to detect incompatibilities between *sentences* (not anyway, unless we help ourselves to logic). In addition, the idea is that we algorithmically elaborate those capacities that are PV-necessary for deploying a given vocabulary. So the business of algorithmic elaboration kicks in after one has settled on those PV-necessary capacities—it cannot therefore be used to beef them up. Clearly though there would be a tension in allowing them to play a role in elaborating capacities PV-sufficient for deploying the explicating vocabulary but to disallow them to feature in the specification of the nature of the capacities PV-necessary for deploying the original vocabulary. Perhaps that speaks against Brandom’s conception of algorithmic elaboration.

2. Multi-premise Inferences

Brandom offers the following account of the algorithmic elaboration of the abilities required to use the conditional. The abilities arise by response substitution of abilities involved in any ADP.

Circumstances:

The response of finding good the inference from p to q is replaced by the response of being prepared to assert ‘ $p \rightarrow q$ ’.

Consequences:

The response of being prepared to assert ‘ $p \rightarrow q$ ’ is replaced by the response of finding good the inference from p to q .

Finding good the inference from p to q is explained as being disposed to assert q , if disposed to assert p .

From the algorithmic elaboration in this simple case one can read off the expressive relation: the sentence ‘ $p \rightarrow q$ ’ expresses the inference which it is related to by response substitution. But this simple case is exceptional. I shall point out first that when we move over to consider cases involving more than one premise we cannot read the relation of expression off the algorithmic elaboration because the expressive relation is the upshot of recursions based on the elaboration.

Conjunction

Brandom presents the following MUD for the conditional—see figure 1. The ADP will, in general, include multi-premise inferences. One way of coping with this is to add conjunction in order to factor these in as conjuncts in a conjunctive antecedent, which then expresses the multi-premise inference. So we arrive at Figure 2. I simply want to probe the role of conjunction here. My hold on these matters is less sure than I’d like and not sure enough to know whether there is a genuine difficulty here. But I can’t see my way clearly through the MUD.

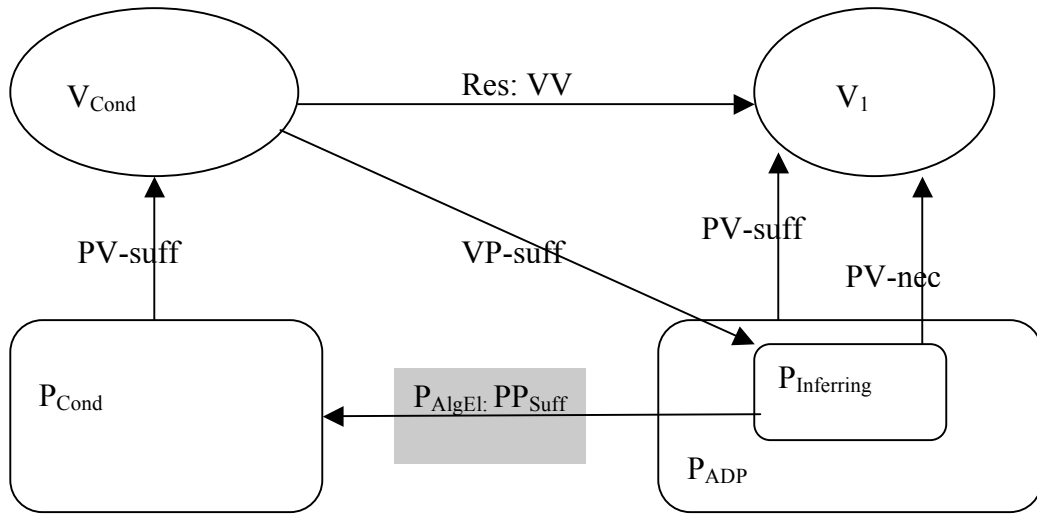
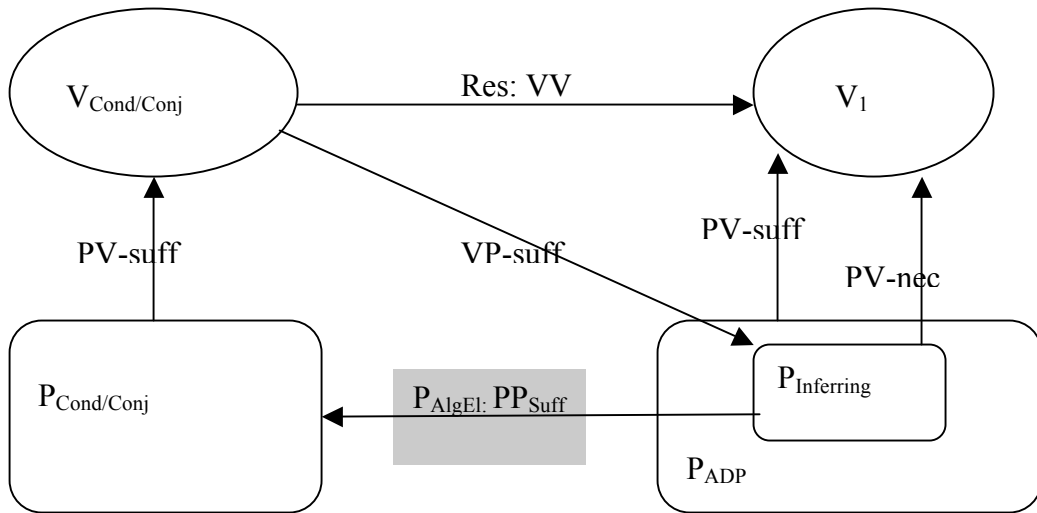


Figure 1



Figure

One might then suppose we could have the following MUD:

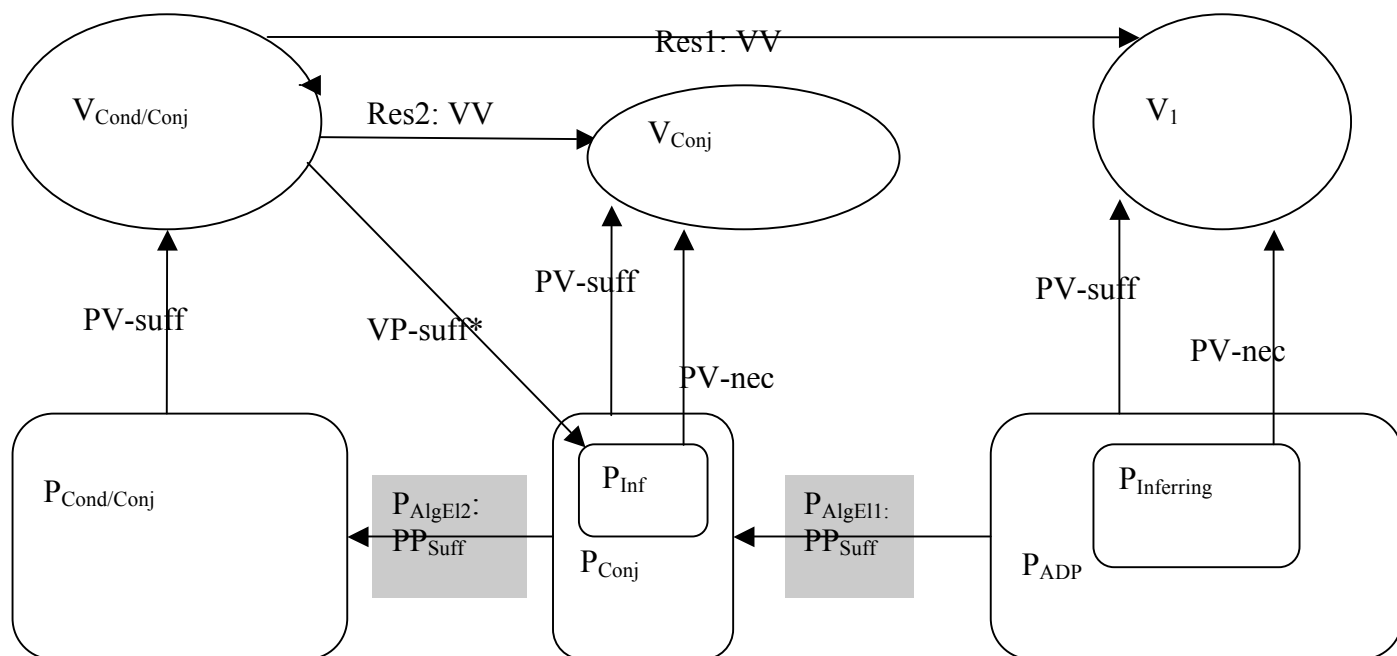


Figure 3

Now the expressive relation involved in VP-suff* is not itself given by the algorithmic elaboration; rather it is based on it. Take the following inferences involving three sentences:

$$p, q \mid - r \qquad p \& q \mid - r$$

Both of these inferences will be expressed by the conditional $(p \& q) \rightarrow r$. Now, in itself, that may be a cause for concern but I shan't treat it as such—perhaps there's no need for a conditional to be uniquely expressive of an inference. Rather my worry emerges when we think of how this will be algorithmically elaborated. Presumably what we shall have is something like the following transitions:

Circumstances:

$$\begin{aligned} \text{Preparedness to infer: } p, q \mid - r &\Rightarrow \text{Preparedness to assert: } (p \& q) \rightarrow r \\ \text{Preparedness to infer: } p \& q \mid - r &\Rightarrow \text{Preparedness to assert: } (p \& q) \rightarrow r \end{aligned}$$

Consequences:

$$\text{Preparedness to assert: } (p \& q) \rightarrow r \Rightarrow \begin{aligned} &\text{Preparedness to infer: } p, q \mid - r \\ &\text{and preparedness to infer: } p \& q \mid - r \end{aligned}$$

The number of inferences expressed by a conditional with conjunctive antecedent will depend on the number of conjuncts in the antecedent. So this transduction from the inferential to the assertive practice will require a distinct elaboration in each case. There is no algorithmic elaboration which *directly* specifies every case. Thus we cannot legitimate the MUR VP-suff*, that is, the relevant expressive relation by giving a single expressive algorithm. The expressive relation cannot, in general, simply be read off the algorithmic elaboration of the relevant capacities.

The problem is highlighted in the relation between $P_{\text{Cond/Conj}}$ and P_{Conj} because the conditional with conjoined antecedent expresses a number—which number is determined by the number of conjunctions in the antecedent—of inferences. But the problem might seem to be present, in a sense, in the relation

between $P_{\text{Cond/Conj}}$ and P_{ADP} since we seem to need separate clauses for algorithmically elaborating inferences involving n premises, for each n . In effect this is, of course, to treat each

$(_ \& _ \& _ \dots _ \& _) \rightarrow _$ as a distinct connective. Of course we don't attempt to do anything quite this silly; rather we give a recursive account of conjunction (replacing preparedness to assert P and to assert Q with preparedness to assert $P \& Q$, and vice versa) and then offer the straightforward algorithmic elaboration of the capacities involved in the conditional by treating these as expressive of single premise (though possibly conjoined) inferences. So we might have the following MUD:

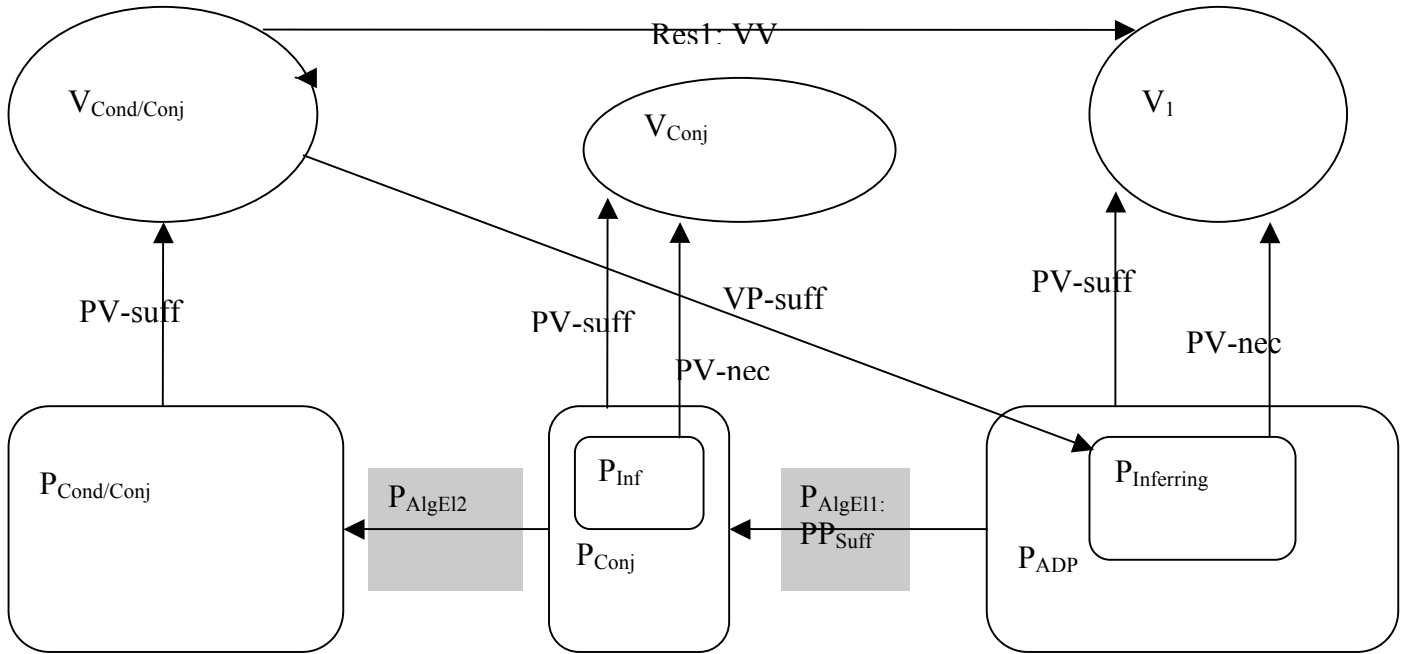


Figure 4

Two worries:

(i) The facility of the conditional to express inferences accepted in the autonomous practice is now built on a role for conjunction which is not, in this sense, expressive. Conjunction is here used to express certain forms of logically complex content—which are *then* placed as the disposal of the expressive project.

(ii) The account is intended to apply to any ADP but, if so then it should apply to a practice that includes both multi-premise inferences *and* conjunction. But we have argued that there are strains in seeing how this can be the case—there is no algorithmic elaboration of the requisite abilities that is itself a transduction of the inferential into the expressive abilities. If this is a failure of VP-suff* it is a counter-example to the general claim of VP-suff in figure 4.

The Conditional and Multi-premises

We are not obliged to express multi-premise inferences by means of a conditional with conjoined antecedent; rather we might instead nest conditionals. In this case we would have:

Circumstances:

The response of finding good the inference from p_1, \dots, p_n to q is replaced by the response of being prepared to find good the inference from p_1, \dots, p_{n-1} to $p_n \rightarrow q$.

Consequences:

The response of finding good the inference from p_1, \dots, p_{n-1} to $p_n \rightarrow q$ is replaced by the response of finding good the inference from p_1, \dots, p_n to q .

As in the case of conjunction there is no reading off the expressive function of the conditional directly from this algorithmic elaboration of the abilities requisite for its deployment. The reason is even more obvious. *Expression* is a relation between a *sentence* and an *inference*; here we directly establish a relation between one *inference* and another. This, of course, need not be fatal to the expressive *project*⁵ provided we can see the expressive role of logical vocabulary emerging from the manner in which those capacities necessary for its deployment are algorithmically elaborated. But note that we must concede here—as we did with conjunction—that the conditional has a role independent of expression, on which its expressive role is based: the conditional enables the expression of conditional contents as conclusions of inferences.

Although the account is not itself expressive, we explained inference in terms of being disposed to assert one sentence on condition that one is disposed to assert another (or some others). So it yields an expressive relation because, at a certain point, one will achieve a sentence that one is prepared to assert unconditionally. That sentence can be taken as expressive of the original inference.

What worries me in this account is that one's preparedness to assert a sentence unconditionally may be a consequence of interference from other inferences one is disposed to find good. Take it that the following inferences are good:

$$\begin{array}{l} q|r \\ p,q|r \end{array}$$

The first inference is unproblematically expressed by the conditional $q \rightarrow r$. When we turn to the second inference we shall perhaps arrive at the following:

$$p|q \rightarrow r$$

Which is explained in terms of abilities as follows: one is disposed to assert ' $q \rightarrow r$ ', if disposed to assert p . But the condition here fails to impose any substantial constraint on assertion of ' $q \rightarrow r$ ' because one is already disposed to assert it. Thus it seems we have arrived at a sentence we are prepared unconditionally to assert and thus will be expressing the inference $p,q|r$, counter-intuitively, by means of the conditional ' $q \rightarrow r$ '. That is the picture when we focus on circumstances of assertion. In order to consider consequences of assertion take the following two dispositions:

$$\begin{array}{l} \text{Disposed to assert } (p \rightarrow (q \rightarrow r)) \\ \text{Disposed to assert } (q \rightarrow r) \end{array}$$

The first is unpacked inferentially as the disposition to find good the inference from p and q to r ; the second as the disposition to find good the inference from q to r . But, given that the second holds, the first will clearly and vacuously hold. Thus, in these circumstances, the inferential abilities associated with the preparedness to make either assertion are the same. So the two assertions have the same expressive power.

Guess there are two ways of making my complaint. The first is to argue that there is no reason to think that the conditional taken to express the inference will be the 'right' one—so the upshot promises to be counter-intuitive. The other is to point out that, in the circumstances, the two assertions have the same grounds and consequences of assertion. Were different expressive powers to accrue to each this would be thoroughly mysterious.

That they have the same expressive power emerges from the manner in which the abilities governing the assertion of conditionals have been algorithmically elaborated and, importantly, explained in terms of simply being disposed to make conditional and unconditional assertions. That seems to be the nub of the problem. For once we take this extensional reading of the nature of the relevant capacities there is nothing to aid us in distinguishing a capacity that is genuinely unconditional from one whose conditions impose no

⁵ Note that I say 'fatal to the expressive *project*'; I've just argued that there may be grounds here for questioning the completeness of the expressive conception of logic.

constraints. Of course it is not, in general true, that the abilities (in terms of conditional dispositions to assert) associated with being prepared to assert a sentence of the *form* $(p \rightarrow (q \rightarrow r))$ and being prepared to assert a sentence of the *form* $(q \rightarrow r)$ can be identified, nor true that the assertive dispositions associated with an inference of the *form* $q|r$ and one of the *form* $p,q|r$ can be identified. But that does nothing to refute the claim that, *in the envisaged circumstances*, there is no pulling them apart. Insisting on the right sort of generality in the way we associate inferential and assertional abilities here is of course what we should be aiming at. The worry is that it is hard to see how we can achieve the right sort of generality without helping ourselves to the notion of formal validity. After all, what seems to go wrong is that neither the inference $q|r$ nor $p,q|r$ is *formally* valid but the movement from the former to the latter inference *is* formally valid. The consequence of this is that there is no distinguishing the former from the latter inference in terms of (conditional) dispositions to assert, given that one accepts the former inference.

Another way of making this point is to note that the argument proceeds under the assumption that we are in conditions in which the inference $q|r$ is taken to be good. So one might hope to respond to it by saying that we need to consider the inference $p,q|r$ in all circumstances, including those in which this inference cannot be assumed to hold good. But then the question is what we mean by ‘all circumstances’: *all possible circumstances* may not include any in which the inference fails (depending on the nature of the inference) and *all logically possible circumstances* just invites in the notion of formality through the back door.

3. The Formality of Logic

Logical vocabulary is that vocabulary required to express as claimings the inferential relations of any ADP. But how do we recognise this feature of logic? Is the claiming an explicitation when it transforms the original inference into a formally valid inference or do we have an independent handle on explicitation which yields an explanation of formality?

Consider the following ‘conditional’ defined in terms of the logical conditional as follows:

$$p \Rightarrow q \text{ iff } (p \& \text{water is H}_2\text{O}) \rightarrow q$$

Now our question is whether ‘ \Rightarrow ’ deserves to be called a piece of logical vocabulary. One might suppose that we can surely rule that out because ‘ $p \Rightarrow \text{water is H}_2\text{O}$ ’ will always hold. But, in general, this need not affect the explicating powers of ‘ \Rightarrow ’ since the locution will still be conservative relative to the vocabulary from whence p is drawn. So we rule out its credentials by showing that it fails its explicating function when p is drawn from the same vocabulary as ‘water is H_2O ’—say talk of physical stuffs.

Mercury is an element \Rightarrow Water is H_2O ,

will hold although the following inference is not accepted,

Mercury is an element $|$ - water is H_2O

Thus here conservativeness fails. So the account has the resources to stave off the counter-example because logic must fulfil its expressive role with respect to *any* vocabulary. The generality of this requirement plays an important role in delivering the formality—thought of as the topic neutrality—of logic. But now consider an arithmetic sentence so obviously true that it will be inferred from any other arithmetic sentence, e.g., ‘ $1=0+1$ ’⁶.

If we now define ‘ \Rightarrow ’ as:

$$p \Rightarrow q \text{ iff } (p \& 1=0+1) \rightarrow q$$

then the counterexample will go through provided that practitioners accept:

⁶ An obvious analytic truth would do equally well.

$r \vdash 1=0+1$, where r is any arithmetic proposition.

To stave off the problem it would seem the expressivist will have to argue that arithmetic vocabulary is itself logical, in that case we could simply accept ' \Rightarrow ' as a piece of logical vocabulary. But even if logicism about arithmetic is true one would not expect it to be a consequence of expressivism about logic⁷.

So what is wrong with ' \Rightarrow '? My sense is that we reject this connective as being logical not because of its expressive role but because the inference:

$(p \ \& \ 1=0+1) \rightarrow q, p \vdash q$

is not (unlike MP) *formally* valid.

Conclusion

According to the first argument here if expressivism entails inferential conservativeness then the expressive function of a piece of vocabulary is not necessary to its logical role; logical vocabulary, in general, fails to have such a role (if logical inference is epistemically useful). According to the second and third arguments the notion of expression needs to be explained in terms of formal validity. Thus even if logical vocabulary is distinctively expressive we cannot use that role in order to distinguish it; since comprehending it as expressive presupposes an ability to distinguish it. Finally I've also suggested that the expressive function of logical locutions is built on a more fundamental (perhaps also more general) role in expressing complex contents—though I haven't attempted to argue the point, it may well be that that more fundamental role is the site in which to seek logic's distinctive character.

How to decide?

- Multi-premises and conjunction
- Objectivity of content: Do Brandom's objectivity proofs give us enough?
Algorithmic elaborations which function in principle?
 - Should one's philosophy of logic apply to those who have an erroneous conception of content? Dummett allows this: realist content/classical logic; anti-realist content/intuitionistic logic
 - VP sufficiency: does this appeal to formality or is the application to any practice enough? [Probably nothing here.]

⁷ This defence would face more severe difficulties had we chosen an analytic truth.

Logic and Pragmatics: linear logic for inferential practice

Daniele Porello

danieleporello@gmail.com

Institute for Logic, Language & Computation (ILLC)
University of Amsterdam,
Plantage Muidergracht 24
1018 TV Amsterdam (NL)

Abstract. In this paper I discuss logic in the pragmatic approach of (Brandom, 2008). I consider different logical consequence relations (classical, intuitionistic and linear) and I will argue that the formal treatment proposed by Brandom, even if I believe it provides powerful intuitions and an interesting framework on logic in general, doesn't allow to state properly the relationship between different logics. I propose an alternative account of the elaboration of logical vocabularies not based on incompatibility semantics, rather on a particular notion of interaction, which I claim is implicit in the practice of giving and asking for reasons, which allows to state the relationship between different logics in terms of different aspects of the inferential practice.

1 Introduction

The analytic pragmatism proposed by Brandom, which states that we should “look at what it is to use locutions *as* expressing meanings – that is, at what one must *do* in order to count as *saying* what the vocabulary lets the practitioners express” (Brandom 2008, p. 9) opens an interesting point of view on foundational issues in logic.

In this paper, I will analyze the relationship between the *use* of a certain logical vocabulary and the inferring (pre-logical) practice-or-abilities from which, following Brandom's approach, the logical vocabulary may be elaborated.

My thesis is that different logical vocabularies are related to different aspects of inferential practice and that inferences codified by different logics, such as classical, intuitionistic and linear logic, say something important on the pre-logical practice itself. In Section 2, I briefly present the approach to logic in (Brandom, 2008): I discuss incompatibility semantics and some consequence of the fact that incompatibility is not apt to represent properly intuitionistic (and linear) consequence relation. In Section 3, after presenting some reasons to consider linear logic a good framework to place the comparison between inferential practices, I propose an alternative approach to incompatibility relations. This treatment provides an interpretation of linear logic in terms of discursive practice and, since in linear logic one can express classical and

intuitionistic logic, it will provide a framework to define an articulation of the notion of inferential practice that can account also for classical and intuitionistic consequence relations, besides linear consequence.

2 Logic and inferential practice

As Brandom summarizes (p. 136), we can see how logic is related to discursive practice, the practice of giving and asking for reasons, and in which sense the use of logical vocabulary can be justified.

Starting from the practice of giving and asking for reasons, one argues that they are sufficient for the practice of deploying basic *normative* vocabulary, in particular the deontic modal vocabulary of 'commitment' and 'entitlement'; then one uses that as pragmatic metavocabulary that specifies how to deploy the concept of *incompatibility*, which is interpreted as constitutively *modal* notion; then one can use this as semantic metavocabulary in which to define a *consequence* relation of incompatibility-entailment; on the basis of the relation of incompatibility-entailment, one then defines a logical vocabulary.

I will focus on the relationship between incompatibility relations and (logical) consequence relations one can define in this setting.

If we take a closer look at the formal theory Brandom develops, we see that it is committed with the assumption that classical logic, at propositional level, is *the* logic of incompatibility: “we have seen that any standard incompatibility relation has a logic whose non-modal vocabulary behaves classically” (p. 139).

Moreover, it turns out in general that all the inferential practice the notion of incompatibility can express or justify are more or less those that can be explicated by means of classical consequence relation.

The reason is that incompatibility relations can define *only* “standard” consequence relations, where a standard consequence relation is defined by two properties: *general transitivity* and *defeasibility*. Consider intuitionistic consequence relation.

The first condition is equivalent to cut rule in sequent calculus, and it is of course satisfied by intuitionistic logic.

This is not the case for defeasibility, which states intuitively that if a proposition *B* is *not* a consequence of a proposition *A*, then there is something that yields an absurdity, when added to *B* but not when added to *A*.

The reason why intuitionistic logic doesn't satisfy defeasibility is that it requires to be able to find a witness also for the badness of an inference (see p. 137 and 165-173). In intuitionistic logic, a witness of good inference from *A* to *B* is always given in a natural way, it is the proof of *B* given *A*. But the fact that *A* doesn't follow from *B*, in intuitionistic logic means in general that there is *no* witness, no proofs of *B* given *A*. This is the constructive, or epistemic, character of intuitionistic logic: we don't have good reasons for what we don't know.

Since intuitionistic inference cannot be fully represented by incompatibility relations, the relationship between classical and intuitionistic logic cannot be stated in terms of pragmatically mediated semantic relations.

I claim that their relation has a special interest for semantics since, as Dummett points out, classical and intuitionistic logic provide two different theories of meaning, with different key concepts: the first one defines meaning in terms of truth conditions, while the second one gives a characterization of meaning in terms of proof, or reasons.

Assuming classical logic as the vocabulary related to propositional inferential practice, we are implicitly assuming that practical inferences represented by conditionals which differs from classical logic conditional are in some sense derived. If one consider what I may call *intuitionistic practice* of inferring, according to which we reject arguments by contradiction, the only way we have to explicate those inferential practices is by saying that intuitionistic inference doesn't behave like classical one and find reasons for this divergence (for example arguing that there are not enough defeasors, see p. 173).

Since, as Brandom proves, standard consequence relations are precisely those that can be obtained by means of incompatibility relations (see p. 138) and that no incompatibility relation can define a non-standard consequence relation, we are able to justify only those inferential practices which can be stated more or less in terms of classical logic.

Moreover, even if non-classical inferences could cleverly be explicated by means of some complicated modal logic construction (which is also justified in Brandom's approach) that would not be grounded in any inferential practice defined by Brandom.

So for example there is no way to justify causal inferences in a pragmatic way.

Let's consider a toy example of causality. Assuming the notion of incompatibility Brandom axiomatizes, one can prove the following (see p. 128):

$$\text{If } A \text{ entails } B \text{ and } A \text{ entails } C, \text{ then } A \text{ entails } B \text{ and } C. \quad (1)$$

This is a famous example proposed by Girard, in order to explicate the meaning of linear logic connectives. As an example of (1), we can consider a drinks dispenser: "if I insert a coin, I get a coffee", "If I insert a coin, I get a tee", then "if I insert a coin, I get a coffee and a tee". As Girard argues, assuming (1) as inference pattern amount to forget any causal relations between premises and conclusions of an inference since, briefly speaking, interpreting propositions as events, we do not make distinction between one occurrence of an event or any number of occurrences¹

Therefore, since the notion of incompatibility is not suitable to represent consequence relations which are well codified as intuitionistic reasoning, I claim that the notion of incompatibility as stated by means of general transitivity and defeasibility, is not adequate to produce a logical vocabulary which explicates good reasoning that are performed in the inferential practice.

¹ See (Girard 2006), pp. 217-218. The reason why classical logic doesn't account for causality can be seen, considering classical sequent calculus, looking at the structural rules of weakening and contraction. In particular, considering weakening, we have that if B is an effect of A , then we should admit that B is an effect of A and any other event.

In the next section I will show how the practice of giving and asking for reasons may also justify a different kind of semantics that can to explicate intuitionistic, classical (and linear) inference. In order to do that, however, we should replace the notion of incompatibility with something else.

4 Linear Logic and inferential practice

I briefly recall some features of linear logic which shows how linear logic could be a considered a good choice to state the relationship between different inferential practices. Linear logic has been introduced in (Girard 1986) as a resource conscious logic in which we can keep track of the *use* of hypotheses in a deduction. From a proof theoretical point of view, if we consider sequent calculus, a special attention should be devoted to its *structural rules* of weakening, contraction and exchange². The point of view introduced by linear logic in proof theory may be saying that structural side almost determine the logical side.

Consider a sequent of the form $\Gamma \vdash \Delta$, if we take contraction and weakening at a global level both on the right and on the left of the sequent symbol, when we define propositional connectives, they will behave classically.

As Gentzen somehow surprisingly proved, intuitionistic logic may be obtained from classical sequent calculus simply restricting sequents on the right to be one single formula. That is enough to reject, for instance, the provability of excluded middle. This may be interpreted as a quite extreme rejection of structural rules on the right.

Linear logic doesn't assume structural rules at a global level, rather one is allowed to perform weakening and contraction just in a controlled way.

The rejection of structural rules at a global level has strong consequences on the form of the logical connectives we can define. Briefly, it entails that we have two kinds of conjunction, and by duality, two types of disjunction: the reason is that we cannot identify anymore the additive presentation of the rule of conjunction (which identifies the contexts) with its multiplicative presentation (which make copies of the context):

$$\text{if } \Gamma \vdash A \text{ and } \Gamma \vdash B, \text{ then } \Gamma \vdash A \text{ and } B \quad (2)$$

$$\text{if } \Gamma \vdash A \text{ and } \Gamma \vdash B, \text{ then } \Gamma, \Gamma \vdash A \text{ and } B \quad (3)$$

The two formulations are equivalent if we assume contraction and weakening.

According to the rejection of structural rules at global level, in linear logic there are two distinct conjunctions: an additive conjunction denoted “&” (“with”) and a multiplicative conjunction “ \otimes ” (“times”).

The expressive power of classical (and intuitionistic) logic is retrieved by means of a controlled treatment of structural rules, which is achieved by means of *exponentials*,

² For reasons of space I cannot present here a detailed overview of linear logic, I refer to (Girard 2006).

denoted by $!A$ and $?A$. Those connectives, briefly speaking, allows structural rules on the left and on the right side of the sequent respectively.

By means of exponentials, one can translate classical and intuitionistic logic into linear logic, and this translation, besides preserving provability, allows to see the relationship between classical, intuitionistic and linear *proofs*. It is therefore apt to state the properties of *how* the inferential practice is performed.

As an example, we can consider the inference (1). In classical sequent calculus can be stated as follows³:

$$\text{if } A \vdash B \text{ and } A \vdash C, \text{ then } A \vdash B \wedge C \quad (4)$$

Its translation in linear logic can be defined as:

$$\text{if } !A \vdash B \text{ and } !A \vdash C, \text{ then } !A \vdash B \wedge C \quad (5)$$

That shows in which sense the classical inference (1) forgets any causal relation between premises and conclusions: “!” means that we can make any number of copies of A (by weakening). If we consider the toy causal relation we saw between the event of inserting an euro in a machine and the event of getting a coffee, that would amount saying that a single coin is the cause of any number of coffees or, symmetrically, that any number of coins is the cause of getting one single coffee.

As linear logic provides an analysis of proofs, it is interesting to view linear logic not as an alternative logic, but as a proof-theoretical analysis of logic itself, which shows how we can perform a sort of decomposition of classical and intuitionistic reasoning. I would like to stress here that the decomposition is performed in terms of *proofs*, in terms ways of *using* hypothesis in the inferential practice.

In the next paragraph I sketch an interpretation of linear logic semantics which aims to show how it may be related at least to some intuitions on the practice of giving and asking for reasons⁴.

4.1 Discursive practice for linear logic

In the original paper (Girard, 1986) there is an intuitive interpretation of *phase semantics*, which is the algebraic semantic providing a canonical model of linear logic, in terms of *phases of observations* and *facts*, the metaphor is inspired by physics and quantum mechanics.

I suggest a different interpretation based on *actions that may count as reasons* and *propositions*.

³ I use the classical symbol for conjunction since in this case the two conjunction of linear logic collapse in the classical meaning of *and*, since we are licensing structural rules.

⁴ The interpretation I propose would of course require a closer comparison with the notions of commitment and entitlement that Brandom analyzes investigating the practice of giving and asking for reasons. Here I just sketch how it works, to show that it can account for intuitionistic as well as classical consequence relation.

The intuition behind this interpretation is that there is interaction, between players involved in the practice of giving and asking for reasons, when there is a form of agreement between what counts as a reason for accepting A and what counts as a reason for rejecting A . It intuitively means that two opponents at least agree that they are challenging concerning a same issue, that they are playing the same game⁵. The idea is that we are going to interpret propositions as a sort of well behaved sets of actions that may count as reasons.

Start with a commutative monoid $(M, \bullet, 1)$, the elements of the monoid are interpreted as actions that may count as reasons. The multiplication of the monoid represents a concatenation of such actions, one may imagine in a discursive practice or game.

The unit 1 of the monoid represents a sort of actions with changes nothing: given any action p , performing 1 , doesn't matter: $p1 = p$.

One defines the following operation on subsets of the monoid, let $X, Y \subset M$, $X \multimap Y = \{m : \forall x \in X \ mx \in Y\}$.

A *phase space* is given by a commutative monoid together with the choice of a *pole* $\perp \subset M$.

Then one defines *negation* of a subset of X , $\sim X$ as $X \multimap \perp = \{y : \forall x \in X \ yx \in \perp\}$.

Negation allows to define directions: if A is a set of reason *for*, then $\sim A$ is a set of reason *against*. The pole represents what we may call a set of actions that count as reasons *for* and *against* at the same time.

Using negation, it is possible to define *facts*, as subsets of M such that $X = \sim \sim X$. The meaning of facts is sometimes explained intuitively saying that a fact is a set of elements that pass the same test.

In the interpretation I am proposing, facts are those sets of reasons on which the form of agreement I suggested holds. More precisely, facts are sets of reasons A such that a reason against a reason against A is a reason for A .

If we look at this property in terms of games with two opponent which are engaged in a dialogue, that simply means that the two players are playing the same game⁶:

⁵ It is important to remark that in this interpretation there is no content *before* interaction: the fact that a proposition may have a content depend on the fact that it shows this form of interaction. Of course, this is a very strong claim which is not justified here. It could be considered as strong form of pragmatism where actions are primitives and propositions are derived. The interesting point is that we can define logical vocabularies assuming just the form of agreement defined by the notion of fact, which I believe is not a demanding condition for a discursive practice.

⁶ There is an interpretation of linear logic in terms of game semantics which is compatible with the one I presented here. The reason why I didn't defined linear logic in terms of games is that it would require a longer exposition of proof-theoretical aspects of linear logic. In that interpretation formula are games, while a proof of formula A is a winning strategy for the game A . In this interpretation, there is also the possibility of considering strategies for a formula, they would correspond to *paraproofs* of that formula, or non logical proof of a given formula.

consider a proposition A , my move against my opponent's move is a move in the same game, we are still playing the game A .

Not all sets of actions have this property, when it holds, we can speak of *propositions*, so the intended interpretation of propositions is given by facts.

Among the properties that holds in this structure, one has that for any subset $X \subset M$, one can consider the smallest fact containing X given by $\sim \sim X$.

So one can prove for example that M and \perp are fact. Moreover, one can define the fact $\mathbf{1}$ as the negation of \perp , $\mathbf{1} = \sim \perp$, that is equivalent to say that $\mathbf{1} = \sim \sim \{1\}$, where 1 is the unit of the monoid. The usual semantic notions of truth in a model maybe restated in this setting. One defines a language which will be interpreted in this semantic structure, where the interpretation of a proposition A will be a fact in a phase space.

Then we can say that a proposition holds in a phase structure, when $\mathbf{1}$ belongs to the interpretation of A : that intuitively means there is no need for arguing on that fact anymore: being 1 in the sets of reasons for A , there is nothing more to do concerning the issue A .

Remark that if one takes the pole to be empty, then we have just two facts M and \emptyset , and we have the truth values semantics and classical logic.

The connectives of linear logic are then defined on facts. Without entering details, connectives will describes how to deal with reasons for complex statements.

For example, an action that counts as a reason for $A \otimes B$, should provide two reasons, one for A and one for B , while the meaning of $A \& B$ is that there are reasons that may count *both* for A and B .

The interpretation of exponentials is usually achieved considering actions that counts as reasons for a proposition $!A$ as actions which can be performed any number of times⁷. We can see, at least intuitively, what the toy example of causality shows. Classical inference (1) requires that any number of actions counts as a reasons A and that each of it count as a reasons for B (looking at the sequent symbol as inclusion).

While for the machine example, the reason for getting *this* cup of coffee is that I put *that* coin in the machine, not any number of coins.

The phase semantics is sound and complete respects linear logic sequent calculus. Therefore, if we state the practice which is sufficient in order to count as deploying a logical vocabulary in the terms I sketched, we can justify linear inferences; then, using the translation of classical and intuitionistic logic in linear logic, we can also justify classical and intuitionistic inferences. Looking at which kind of actions that count as reasons are required for the translation of classical and intuitionistic formula in linear logic, we can see how the inferential practice on reasons is articulated.

⁷ Technically, $!A$ is interpreted as the fact obtained from the intersection of the reasons for A with the set of the idempotents of the monoid (the element m in M , such that $mm = m$).

5 Conclusions

I argued that the incompatibility semantics is not apt to represent the variety of inferences which is interesting to consider at work in our inferential practice. We saw then how it is possible to justify in terms of inferential practice, in terms of giving and asking for reasons, a different practice from which we can elaborate linear logic. Then, I sketched how to state the relationship between classical, intuitionistic and linear reasoning, proposing therefore a more complex articulation of the notion of inferential practice itself. Of course, the approach I presented is just sketched and many more arguments should be provided.

I believe however that it is worthy to investigate in this direction, since, in particular it would provide philosophical foundation of the purely interactive account of logic given by some recent developments of linear logic. This foundation would be grounded in a notion of interaction which lays already in our discursive practice, as one can realize adopting the point of view of Brandom's analysis.

References

- Brandom R. B. (2008) *Between saying and Doing. Towards an Analytic Pragmatism*, Oxford, Oxford University Press.
- Dummett M. (1993) *The Logical Basis of Metaphysics*, Harvard University Press, Cambridge, Mass.
- Girard J. Y. (2006) *Le Point Aveugle I. Cours de Logique. Vers la perfection*, Hermann, Paris.
- Girard J. Y. (2006) *Le Point Aveugle II. Cours de Logique. Vers l'imperfection*, Hermann, Paris.

Can Negation be Defined in Terms of Incompatibility?

Nils Kurbis

1 Abstract

Every theory needs primitives. A primitive is a term that is not defined any further, but is used to define others. Thus primitives should be terms that can be expected to be understood by everyone.

Negation is a a very fundamental concept. Everyone understands it. No one has problems grasping it. It is a perfect choice for a primitive. Nonetheless, there have been attempts to define it in terms of allegedly more fundamental concepts.

The motivation behind such attempts is to provide a principled basis on which to settle the debate between rival logicians concerning the correct properties of negation. Most prominently, the debate between classicists and intuitionists is largely one about the laws governing negation. If negation is chosen as a primitive, no principled decision can be made.

I shall investigate how successful such attempts are. In particular, I shall argue that defining negation in terms of incompatibility fails, because the latter notion is conceptually rather more demanding notion than negation. Besides, the approach fails to decide between classicists and intuitionists. As a matter of fact different incompatibility-theorists come to different conclusions concerning which logic is the right one. Thus quite apart from the conceptual difficulties involved in choosing incompatibility as primitive, in the light of the debate between classicists and intuitionists, the approach does not fare any better than choosing negation as primitive.

2 On Primitives

Choice of primitives is an important issue in the philosophy of logic, and it is one that I think is not paid as much attention to as it deserves. What is a primitive? Every theory needs primitives. They are terms that are not defined any further, but in terms of which others are defined. Accordingly one should use terms as primitive that can be expected to be understood by everyone, of which we can expect everyone to have an intuitive, pre-theoretical understanding.

So why not *not*? If all this is the case, why would anyone ever want to define negation, rather than take it as a primitive? Everyone understands negation. No one has problems grasping it. It is a prime candidate and a perfect choice for a primitive.

Moreover, the most straightforward way that comes to mind to define negation, namely in terms of truth and falsity by ‘ $\neg A$ is true iff A is false’, doesn’t actually evade use of negation. Something needs to be said about the relation between truth and falsity, and this makes use of negation, e.g. ‘If A is true, then A is not false’. So it seems that we don’t get around negation.

Despite the fact that negation seems to be a most obvious choice for a primitive, philosophers have suggested that negation should be defined in terms of other concepts. Dummett suggests to define negation in terms of rules of inference. It has become fashionable to propose definitions of negation in terms of incompatibility.¹ Price and Rumfitt suggest to define negation in terms of assertion and denial, and they, too, make use of a primitive notion of incompatibility between speech acts.

3 Disputes over Logical Laws

The motivation behind attempts at defining negation in terms of some allegedly more fundamental notion is the aim to settle disputes between certain rival schools of logicians. Even though negation is such a simple notion, as a matter of fact intuitions have diverged concerning which logical laws hold

¹Cf., e.g., Christopher Peacocke (‘Understanding the Logical Constants. A Realist’s Account’, *Proceedings of the British Academy* 73 (1987), 153-200), Huw Price (‘Why ‘Not’?’, *Mind* 99 (1990), 221-238), Neil Tennant (‘Negation, Absurdity and Contrariety’, in *What is Negation?* edd. Dov Gabbay and Heinrich Wansing, (Dordrecht: Kluwer, 1999) 199-222) and to some extent, following them, Ian Rumfitt (‘“Yes” and “No”’, *Mind* 109 (2000), 781-823).

for it. It has been debated whether $A \vee \neg A$ is a logical law, whether from $A \& \neg A$ everything follows, and whether there is an understanding of negation on which the negation of a sentence containing ‘presupposition failures’ is defective in the same way the sentence itself. A more arcane dispute is the one whether contradictions can be true.

These disputes cannot be settled if negation is a primitive. A primitive is formalised on the basis of intuition, reflection and conceptual investigation. But as these diverge, each rival camp will start of with a different negation. There is then no basis for deciding the issue between them. What is missing is a common ground on which to debate which formalisation is ‘the correct one’.

Defining negation in terms of something else promises to provide a basis on which a principled decision can be made concerning which is the right way of formalising negation. From the perspective of the *justification of deduction*, finding such a basis for settling disputes over very fundamental logical laws is the Holy Grail of Logic. It is a criterion for the success of a theory of the justification of deduction if it can settle disputes over logical laws.

4 Why Dummett Fails

Dummett’s *proof-theoretic justification of deduction* was designed to settle this debate by formulating neutral requirements for definitions of the logical constants. Primitives of the theory are a ‘thin’ notion of truth and rules of inference.

I have shown that this project fails, because *negation* cannot be defined in this way. [For details see Section 4 (esp. 4.2.1) of ‘What is wrong with classical negation?'] Negation must enter the theory as an additional primitive. As a consequence, Dummett’s theory cannot settle the debate between classicists and intuitionists.

5 Incompatibility

The most prominent way of remedying Dummett’s theory is to define negation in terms of some notion of incompatibility—be it that its incompatibility between facts (Tennant), propositions (Brandom) or speech acts (Price,

Rumfitt).

A superficial look at the question whether negation can be defined in terms of incompatibility may elicit an obvious response. Interestingly, however, there are two contradictory such responses: ‘obviously yes’ and ‘obviously no’. The ‘obviously no’ camp would point out that ‘incompatibility’ is a negative notion; thus the definition is circular—this is in fact, I think, Russell’s reaction to the proposal. The ‘obviously yes’ camp would point out that there are several ways of defining $\neg A$ that use some notion of incompatibility, for instance the Sheffer Stroke ‘not both p and q ’.

At a more reflected level, what the ‘obviously yes’ camp needs to address is the question what the theoretical advantages of defining negation in terms of some notion of incompatibility are. The ‘obviously no’ camp needs to address the point that no circularity arises as a primitive notion of incompatibility, although undoubtedly a negative notion, is not analysed any further as ‘not compatible’.

The fundamental observation behind defining negation in terms of incompatibility is that there seems to be something incompatible about ‘ a is red’ and ‘ a is green’, without this relying on negation. It seems to be straightforward how to define negation on this basis, without any circularity: if p implies that a is red and that a is green, this should suffice for $\neg p$ to be true.

5.1 Some Unsuccessful Definitions

There are certain accounts of incompatibility that may be ruled out. For instance, if ‘ p is incompatible with q ’ amounts to ‘one of them is the negation of the other’, then the approach is either circular or a dispute over whether negation can be defined in terms of incompatibility is merely verbal. Such a notion of incompatibility is simply not different enough from the notion of negation to make the project worth while: negation is merely sold under a new heading.

For similar reasons, incompatibility should also not amount to something like ‘not both’, only expressed without the not. ‘Not both’, or rather ‘neither-nor’, may be viewed as a generalised negation, which applies to a number of sentences rather than only to one.

That ‘not both’ is not a suitable notion of incompatibility can also be seen by considering that on our intuitive understanding of ‘incompatibility’, no contingent or logically true sentence p is incompatible with itself. Quite to the contrary: ‘ p is incompatible with p ’ suggests itself as a definition of

‘ p is a contradiction’. If p is contingent or logically true, ‘ p is incompatible with p ’ should be a contradiction, but ‘not both p and p ’ is not.

What is probably the most obvious way of explaining the notion of incompatibility can also be ruled out, namely to explain ‘ p is incompatible with q ’ as ‘If p is true, then q is false, and if q is true, then p is false’. One reason has already been given, namely that an approach which appeals to truth and falsity is unlikely to succeed without an appeal to negation, as something has to be said about the relation between these two notions. Furthermore, if a definition of negation in terms of incompatibility helps itself to the notions of truth and falsity, one might as well define negation right away through the equivalence ‘ $\neg p$ is true if and only if p is false’. The notion of incompatibility would appear to be superfluous, as all the work could be done by the notions of truth and falsity.

To sum up, if a definition of negation in terms of incompatibility is proposed, then there should be a genuine difference between negation and incompatibility and the notion of incompatibility should do some real work. Otherwise the dispute is merely verbal or there are no theoretical benefits to be gained from employing the notion of incompatibility as a primitive.

5.2 Tennant’s Incompatibility

5.2.1 Outline

Huw Price has put the general idea behind defining negation in terms of incompatibility very neatly: ‘it is appropriate to deny a proposition p (or assert $\neg p$) when there is some proposition q such that one believes that q and takes p and q to be incompatible’.² Neil Tennant proposes a revision of Dummett’s theory in this direction. He suggests to view \perp as a ‘structural punctuation marker’³, which registers ‘metaphysico-semantic fact[s] of absurdity’⁴, such as ‘ a is red and a is green’ or ‘ a is here and a is over there simultaneously’. \perp is subject to the rule

$$\frac{A_1 \quad \dots \quad A_n}{\perp} \quad (1)$$

‘where by this we are to understand that A_1 to A_n are not jointly assertible,

²Huw Price, loc. cit., p.231.

³Tennant, loc. cit., p.199

⁴Ibid. p.202

that they are, that is, mutually inconsistent'⁵. According to Tennant, any speaker of a language grasps that certain atomic sentences are incompatible with each other. The notion of inconsistency 'arises by virtue of what the sentences mean and *various ways that we understand the world simply cannot be*.'⁶

Tennant goes on to give a proof-theoretic definition of negation in terms of introduction and elimination rules for it:

$$\frac{\overline{A}^i}{\Xi} \quad \frac{\neg A \quad A}{\perp} \quad \frac{\perp}{\neg A}^i \quad (2)$$

As \perp can only be arrived at if mutually incompatible sentences have been derived first, the introduction rule for \neg captures the thought that $\neg A$ is true just in case A entails mutually incompatible sentences. The elimination rule is chosen because it is harmonious with the introduction rule.⁷

Tennant's rules are of course to be understood as holding for an *interpreted* language, not a formal calculus. Theorems of the form $\vdash A_1 \supset \cdot A_2 \supset \cdot \dots A_{n-1} \supset \neg A_n$ can be deduced, which are not true on all interpretations of the formal language, but only on those which interpret $A_1 \dots A_{n-1}$ as 'mutually inconsistent'.

5.2.2 Problems

According to Tennant, \perp this is not a proposition at all: it is a 'punctuation mark'—one could as well use a blank space. Hence it is also not something which is always to be interpreted as being false. This has the strange consequence that interpreting $A_1 \dots A_n$ as sentences which may be true together cannot result in the rule becoming unsound. This, of course, is merely a rhetorical point, just as insisting on calling \perp a punctuation mark rather than a proposition is mere rhetoric. Certainly nothing in the rules Tennant has formulated dictates this interpretation. What is more serious is that the use of empty spaces may well be counterproductive in Tennant's framework,

⁵Ibid. p.217

⁶Ibid. p.217

⁷Tennant puts certain restrictions on these rules to fit his intuitionistic relevant logic, which need not concern us here. The three rules do not suffice to prove *ex falso quodlibet*. This could be remedied by adding a principle *ex adversis propositionibus quodlibet sequitur*.

as the validity of rules would then have to be explained with reference to notions of truth and falsity (cf. also ‘What is wrong with classical negation?’, section 4.3.3).

On a less *ad hominem* note, what turns out to be a substantial problem for Tennant’s approach is an attempt to express in the object language that sentences are incompatible. So far, it is not possible to express this in Tennant’s object language, as it only has interpreted sentence letters and the logical constants $\neg, \perp, \supset, \vee, \&, \forall, \exists$. Thus the language is incomplete, as obviously, we are able to say that ‘ a is red’ and ‘ a is green’ are incompatible. Let’s use I^n as an n -place predicate of propositions, where $I^n p_1 \dots p_n$ is to be interpreted as ‘ $p_1 \dots p_n$ are incompatible’.

Having extended the expressive power of the language in this way has, initially at least, the advantage of enabling us to give rules for negation that avoid the detour through \perp . Let’s restrict consideration to $n = 2$, and write Ipq . Modifying Tennant’s introduction rule for negation in the extended framework yields the following:

$$\frac{Iq_1q_2 \quad \frac{\overline{p}^i \quad \overline{p}^i}{\Xi_1} \quad \Xi_2}{\neg p} \quad q_2^i \quad (3)$$

$\neg p$ may be inferred p entails: $q_1 \dots q_n$ and $I^n q_1 \dots q_n$.

This rules capture the fundamental idea behind the definition of negation in terms of incompatibility. But using this rule alone to govern I results in too weak a logic of I . Given our intuitive understanding of incompatibility, we should have $Ip\neg p$, i.e. ‘ p and $\neg p$ are incompatible’, as a theorem. However, given only (3), this is not possible. Suppose you add the connectives I and the rule (3) to classical logic formalised in \neg and \supset . It is easily shown that $Ip\neg p$ is not derivable: interpret Ipq as being true if p and q are both false, and false otherwise. This interpretation, together with the standard interpretation of the connectives \neg and \supset , every assignment of truth-values to the atomic propositions satisfies all rules and axioms of the calculus, but no assignment satisfies $Ip\neg p$. Thus even the full force of classical logic does not suffice to derive $Ip\neg p$ as a theorem, hence it is not derivable in Tennant’s much weaker logic.

To capture the notion of incompatibility more adequately, further rules governing I must be added. But which rules? Obviously it would be counterproductive to add $Ip\neg p$ as an axiom, as that would mean to characterise

incompatibility with reference to negation. The rules we add must not use negation, if the approach of defining negation in terms of incompatibility is not to be thwarted.

Given Tennant's proof-theoretic outlook, the obvious first step towards more rules for I would be to try to formalise rules harmonious to (3). This meets with some difficulties, which are closely connected to the problem of formulation harmonious elimination rules for negation this is the introduction rule for negation:

$$\frac{\overline{A} \quad \overline{A} \quad \Pi \quad \Pi \quad \neg B \quad B}{\neg A}$$

The harmonious elimination rule would be *ex contradictione quodlibet*:

$$\frac{\neg A \quad A}{B}$$

But this rules leads to maximal formulas which cannot be removed from deductions in such a way that no negation rule is used in the transformation.

The remedy that can be used in the case of negation also works for the case of I . We need to use \perp . It has a straightforward introduction rule, which captures in one rule exactly the spirit, if not the letter, of Tennant's rule (1):

$$\frac{p \quad q \quad Ipq}{\perp} \tag{4}$$

\perp may be derived if two sentences have been derived which are incompatible. Negation can then be defined by Tennant's rules (2).

Applying the principle of harmony to (3) yields the following further rule governing I :

$$\frac{\overbrace{\overline{p}^i \quad \overline{q}^i} \quad \Pi}{\frac{\perp}{Ipq}^i} \tag{5}$$

This rules is an introduction rule for I .

Adding this rule does indeed yield $I_{p\neg p}$ as a theorem. However, it also yields something more, namely $\neg p \vdash Ipp$. This is quite unacceptable, at least for atomic p , given the intended interpretation of I , as noted before, as any contingent proposition is compatible with itself.

Hence rule is too strong for the intended interpretation of I as incompatibility. But on Tennant's proof-theoretic approach, he cannot easily evade the point that (5) is the additional rule governing I , as this is required by (4) and the principle of harmony. The connective governed by the rules (4) and (5) is of course the Sheffer function 'not both, p and q '. This is as close as we can get towards a notion of incompatibility in classical and intuitionistic logic. But it is not close enough. It does not capture many intuitions about incompatibility correctly. Hence following up Neil Tennant's notion of incompatibility thus does not lead to a convincing notion at all. In fact, given the difficulties surrounding formalising satisfactory rules for possibility in the proof-theoretic framework one can suspect that it is equally problematic in this framework to formalise the notion of incompatibility, which of course is also a modal notion.

There is thus a lack of fit between Tennant's proof-theoretic approach and his appeal to a primitive notion of incompatibility. There are no theoretical advantages to be had from this choice, rather than choosing negation. In fact, it seems positively harmful, as the notion of incompatibility is not one that can be adequately expressed in Tennant's own framework.

5.3 Brandom's Incompatibility

Robert Brandom attempts to give a semantics with the notion of incompatibility as the primitive which not only covers propositional logic, but also modal operators. According to Brandom, '*incompatibility* can be thought of as a sort of conceptual vector product of a *negative* and a *modal* component. It is *non-compossibility*.'⁸ It would of course be a blatant circularity to claim that incompatibility is defined as non-compossibility, and then to claim that negation can be defined in terms of this notion. So Brandom's remarks must be understood as merely heuristic, to get us on the right track of what notion of incompatibility he has in mind.

Brandom's heuristic procedure does, however, reveal that incompatibility is a more complicated notion than negation, and thus is not as good a choice

⁸Locke Lecture 5, p.16

for a primitive than negation. Brandom needs to appeal to the notions of conjunction, negation and possibility to get us on the right track of what he means by ‘incompatible’, because we have fairly good understanding of the former notions, but not really of the latter. In fact, Brandom himself characterises incompatibility in different ways which do not match up. In fact, two different ways of characterising incompatibility occur in one and the same passage: ‘to say that one way things could be is *incompatible* with another is to say that it is not *possible* that the second obtain if the first does—that if the first does, it is *necessary* that the second does not.’⁹ Thus p is incompatible with q is on the one hand said to be equivalent to $\neg\Diamond(p \supset q)$ and on the other hand to $\Box(p \supset \neg q)$, i.e. $\neg\Diamond(p\&q)$.

This may of course have just been a slip of the pen. But the equivocation might also have a deeper reason. If the first reading is adopted, it would indeed be a contradiction to say that p is incompatible with p , which is desirable given our intuitive understanding of this notion, as $\Diamond(p \supset p)$ is a logical truth (at least in **D** and hence in **S5**). On the second reading, incompatibility is non-compossibility. But this notion doesn’t quite match up with our intuitive understanding of incompatibility, at least not if the possibility used here is the one of **S5**, which, Brandom argues, is the modal logic that turns out to be validated by his incompatibility semantics. To see this, let’s have a look at compossibility and compatibility. We should expect them to be the same concepts, on Brandom’s account of incompatibility as non-compossibility. The problem is that every contingent or logically true sentence should be compatible with itself:¹⁰ if a sentence is not compatible with itself, that would suggest that it is a contradiction. So ‘ p is compatible with p ’ is logically true, if p is such a sentence. However, this shows that $\Diamond(p\&q)$ cannot correctly be interpreted as compatibility, for $\Diamond(p\&p)$ is not a logical truth, at least not in Brandom’s modal logic. Compatibility thus is not compossibility, at least not in the most obvious sense.

The notion of incompatibility is not one that is easily pinned down: it seems close to non-compossibility, but as compossibility doesn’t seem to be the same as compatibility, it isn’t clear how close it is. Incompatibility is thus not a good primitive: our intuitive, pre-theoretic understanding of it is not firm enough.

That our intuitions leave us behind when considering properties of Bran-

⁹Ibid p.10f

¹⁰Brandom agrees, I’ve asked him.

dom’s incompatibility is not surprising if one takes into account the object language connective expressing this notion. Brandom seeks to employ incompatibility as the sole primitive of the semantic theory. Thus what corresponds to it in the object language is a connective in terms of which all connectives of **S5** can be defined—it is the modal version of the Sheffer Stroke. The former is rather more complicated than the latter: $*$, to be interpreted as ‘is incompatible with’, suffices to define all other operators of **S5**, where $p * q$ is equivalent to $\neg\Diamond(p \& q) \vee (\Diamond(p \& q) \& \Diamond(p \& \neg q) \& \neg(p \& \neg q)) \vee (\Diamond(p \& q) \& \neg\Diamond(p \& \neg q) \& (p \& q))$. This connective is arguably not one of which we have an immediate, pre-theoretical understanding. In particular, it is not Brandom’s ‘non-compossibility’. Besides, it is worth noting that $*$ does not adequately capture an intuitive notion of incompatibility adequately: it is not logically true that $p * p$, which is equivalent to $\neg\Diamond p \vee (\Diamond p \& p)$, at least not in **S5**. In fact, $p * p$ can be used as the definition of $\neg p$, so another reason why $*$ does not express our intuitive notion of incompatibility.

In conclude that Brandom’s notion of incompatibility is not a suitable primitive. It is not clear what he has in mind when he speaks about incompatibility. Whenever he is explicit, it does not match up with other plausible requirements on a notion of incompatibility.

It is also worth noting the two incompatibility theorist Brandom and Tennant must have different notions of incompatibility in mind, despite the fact that their heuristic explanations of this notion are virtually identical: Tennant claims that a logic based on this notion is intuitionist (or, more precisely, the negation of his idiosyncratic intuitionist relevant negation), but Brandom argues that negation turns out to be classical. It is plausible to surmise that this is due to differing heuristic explanations of the notion of incompatibility. Tennant favours a ‘verificationist’ notion of truth, whereas Brandom favours a pragmatist one, which then means that ‘ p and q cannot be true together’ has different properties on each reading. It is thus questionable whether choosing the notion of incompatibility, rather than, say, negation, as a primitive succeeds in providing a neutral basis for settling the debate between classicists and intuitionists. The problem is that, because we haven’t got a strong enough pre-theoretic understanding of incompatibility, we need to resort to heuristic readings, which then smuggles illegitimate presuppositions into the theory.

As mentioned earlier, it is a criterion of success for a theory aiming at a justification of deduction that disputes over logical laws can be settled on its basis. However, as a matter of fact choosing incompatibility as a primitive

fails to solve the question whether negation is classical or intuitionist, as different incompatibility theories come to different conclusions about what kind of negation turns out to be definable in terms of incompatibility. Thus much of the motivation for choosing this primitive, rather than negation, has been lost.

6 Concluding Reflections on Incompatibility

There is something that the pairs ‘*a* is red’ and ‘*a* is green’, and ‘*a* is here’ and ‘*a* is over there’ have in common, and we can call this relation ‘incompatibility’. It is not difficult to give a general explanation of what incompatibility consists in: two sentences are incompatible, if they cannot be true together, or alternatively, if each entails the negation of the other. These are general characterisations of incompatibility, which make no reference to the specific content of the sentences which stand in this relation. Neither of them, however, is what theorists have in mind who propose to define negation in terms of incompatibility, as they are talking about a notion of incompatibility not explained any further in terms of truth, falsity and negation. Their notion of incompatibility is intimately tied to the specific *content* of sentences, rather than to general features of classes of sentences, such as truth, falsity or entailing negations of other sentences. In fact, the whole point seems to be that the notion is one tied intimately to the content of sentences, rather than being one that could be explained in a formal manner.

The last paragraph leads me to suspect that incoherent requirements need to be imposed on the notion of incompatibility. On the one hand, it is a notion tied to the particular content of sentences, on the other it needs to be a notion that applies across the board independently of the content of sentences, like a logical constant.

Now some pairs of sentences don’t exhibit this incompatibility, even though they may be said to exclude each other. One needs the *right kind of exclusiveness*: it would not suffice for logic that one can derive, say, ‘Beet-roots are revolting’ and ‘Scotch is disgusting’. There is a sense in which these two sentences exclude each other and cannot be true together – obviously the second is false and the first true – but that would merely result in a logic for my personal prejudice. That is to say, only *certain* atomic sentences which may be said to exclude each other could be used in a definition of negation in terms of incompatibility. ‘*a* is red’ and ‘*a* is green’ seem to exclude each other

in the right way, but ‘Scotch is disgusting’ and ‘Beetroot are revolting’ do not, because of their respective meanings. Hence the reasons why ‘ a is red’ and ‘ a is green’ constitute the right kind of exclusiveness is a matter of their *particular content*. If we characterise two atomic sentences as excluding each other this can only be because of their content. However, in order for the notion of exclusiveness to be of use in a definition of *negation*, rather than merely some indication that we find certain sentences unacceptable, there needs to be a *general* method of determining for *any* two atomic sentences whether or not they exclude each other in the desired way. We need to have a way of telling when we have arrived at two sentences which exclude each other in the right way. A general method is mandatory because the negation to be defined should cover any possible extension of the language by new atomic sentences: for any atomic sentences we may add to the language, it needs to be determined which pairs exclude each other. But this is precisely to say that the method needs to *abstract* from the content of atomic sentences. Hence the desired method for determining whether two atomic sentences exclude each other in the right way has to be general and independent of the content of the atomic sentences and at the same time cannot be general, but due to its nature must be particular and tied to the content of the atomic sentences. Hence there is no such method of characterising the right kind of exclusiveness of atomic sentences.

The only way I can see of reconciling these two opposing requirements is to say that, for instance, the reason why ‘ a is green and a is red’ constitute the kind of exclusiveness is that what is green cannot be red and conversely, if something is red, it is not green, hence if something is red as well as green, it is green as well as not green. But this makes use of negation.

A Reconstruction of the Hierarchical Relations Between Incompatibility-Entailment, Committive, and Permissive Consequences*

Edgar Andrade-Lotero and Catarina Dutilh-Novaes

E.J.AndradeLotero@uva.nl cdutilhnovaes@yahoo.com

ILLC/Department of Philosophy

Universiteit van Amsterdam

Nieuwe Doelenstraat 15, 1012 CP Amsterdam

Abstract

In this paper we propose to investigate the mutual relations among Brandom's three dimensions of semantic inferential articulation, namely, incompatibility entailment, committive, and permissive consequences. Brandom (Unpub.) argues (1) that incompatibility entailment implies committive consequence, and (2) that committive consequence in turn implies permissive consequence. However, under a closer scrutiny of this hierarchy, we see that a number of points are in need of further clarification. First, we question Brandom's claim that the hierarchy is strict. We show that, under reasonable assumptions, all three inferential relations are equivalent. Secondly, we discuss the soundness of Brandom's proofs of (1) and (2). We discuss both results against the background of the main pillars of Brandom's philosophy of language.

1 Introduction

One of the main argumentative lines in Brandom's philosophy of language is the reconstruction of "autonomous discursive practices" in terms of propositions, assertions, and declarative sentences (cf. BSAD, p. 117).¹ The latter is a syntactic notion, and we will not deal with it here. The other two notions, namely propositions and assertions, are treated by semantics and pragmatics, respectively.

On the semantic side, the notion of semantic content of a proposition is defined in terms of its specific inferential articulation (cf., for instance, MIE, pp. 133ff).² Such inferential articulation is understood in a normative way: "Content is understood in terms of proprieties of inference, and those are understood in terms of the norm-instituting attitudes of taking or treating moves as appropriate or inappropriate in practice" (MIE, p. 134). Proprieties of inference are normative statuses, and as such they are grounded on the practice of giving and asking for reasons.

Thus, according to Brandom, the content of an expression is conferred by the intentional states that ground its use in language. "For it is the practical inferential proprieties acknowledged by such attitudes [of treating an inference as correct or incorrect] that make noises and marks mean what they mean" (MIE, p. 174). Nevertheless, Brandom's explanatory strategy consists in explaining simultaneously both the content of

*We would like to thank Martin Stokhof for his insightful comments on an early version of this paper.

¹In this paper, BSAD is an abbreviation of Brandom (2008); MIE is an abbreviation of Brandom (1994).

²This inferential articulation can be conceptualized in terms of the notion of incompatibility (cf. BSAD, ch. 5).

expressions and the content of intentional states. It doesn't come as surprise, therefore, that semantics and pragmatics are explained by the same account of the practice of giving and asking for reasons.³ This leads us to consider this practice in more detail.

In the practice of giving and asking for reasons, the fundamental move is that of making an assertion. To make an assertion is to make a particular move:

According to the model [of the game of giving and asking for reasons], to treat a performance as an assertion is to treat it as the undertaking or acknowledging of a certain kind of *commitment* . . . Doxastic commitments are essentially a kind of deontic status for which the question of *entitlement* can arise. (MIE, p. 142, emphasis in the original)

Being committed and being entitled are normative statuses, which are grounded on the participants' mutual attributions of these statuses to each other. In Brandom's framework, the ability to attribute commitments and entitlements are conditions of possibility for the practice of assertion. Besides these fundamental abilities of attributing commitment and entitlement, a story must be told as to how they combine in order to give rise to the inferential practices which eventually constitute the contents of assertions. Indeed, "the inferential articulation [of assertions], in virtue of which they deserve to be understood as propositionally contentful, consists in consequential relations among the particular doxastic commitments and entitlements" (MIE, p. 142). Thus, not only are the abilities to attribute commitment and entitlement fundamental, but also the abilities to *relate* commitments (entitlements) to other commitments (entitlements). How the ability to attribute commitments and entitlements, and the ability to interrelate these attributions, combine in order to underwrite the required inferences is spelled out in terms of the dimensions of authority and responsibility of assertions. Brandom claims that:

In producing assertions, performers are doing two sorts of things. They are first *authorizing* further assertions (and the commitments they express), both concomitant commitments on their part (inferential consequences) and claims on the part of their audience (communicational consequences). In doing so, they become responsible in the sense of answerable for their claims. That is, they are also *undertaking* a specific task *responsibility*, namely the responsibility to show that they are *entitled* to the commitment expressed by their assertions, should that entitlement be brought into question. (MIE, p. 173, emphasis in the original)

There is both an *intrapersonal* and an *interpersonal* dimension to the inheritance of commitments (entitlements). Both of them are essential to the reconstruction of the practice of giving and asking for reasons (as it is clear from the quote above). But it is only the intrapersonal dimension of the inheritance that we will focus on, the "inferential consequences," as Brandom calls them. This doesn't mean, as we will see, that we don't need to take the other dimension into account.

There are three different notions of inferential relations, based on the notions of commitment and entitlement. They are defined as follows (cf. BSAD, pp. 120f; MIE, pp. 168–9):⁴

Commissive consequence: We say that p commissive entails q , denoted as $p \square \rightarrow q$, if and only if whenever a speaker S is committed to p , S is committed to q .

³A further question that would be interesting to ask is whether this account allows for a difference between literal meaning and speaker meaning, and if so, how they could be reconstructed.

⁴The symbolic convention is ours. It is also worth noticing that although it is acknowledged that inferences can have more than one premiss, we will restrict ourselves in this paper to the single-premiss case.

This kind of inference corresponds to examples like “A is to the West of B, so B is to the East of A; This monochromatic patch is green, so it is not red” (MIE, p. 168).

Permissive consequence: We say that p permissive entails q , denoted as $p \diamondrightarrow q$, if and only if whenever S is entitled to p , S is entitled to q .

This kind of inference corresponds to examples like “the claim that this is a dry, well-made match can serve as a justification entitling someone to the claim that it will light if struck” (MIE, p. 169).

The third inferential relation is based on the notion of *incompatibility*. We say that p is incompatible with q if and only if commitment to p precludes entitlement to q , and commitment to q precludes entitlement to p . The notion of incompatibility defines the semantic content of propositions and performances—the content of p is the class of all q such that p and q are incompatible (cf. BSAD, ch. 5)—, and gives rise to the notion of incompatibility entailment:

Incompatibility entailment: We say that p incompatibility entails q , denoted as $p \circrightarrow q$, if and only if for all r , if r is incompatible with q , r is incompatible with p .

This kind of inference corresponds to examples like “If my first pet (in fact, let us suppose, a fish) *had been* a donkey, it *would have been* a mammal” (BSAD, p. 122).

Brandom (Unpub., p. 23–25) argues that one can use the dimensions of authority and responsibility of assertions (cf. above) in order to determine the following connections between the three inferential relations:

If p incompatibility entails q , then p committive entails q . (1)

If p committive entails q , then p permissive entails q . (2)

In the next section, we will reconstruct Brandom’s hierarchy of semantic inferential relations, as discussed in Brandom (Unpub.). In section 3, we will show that the hierarchy is actually different, as both the reciprocal of (1) actually follows from (2), and the reciprocal of (2) follows from (1). After an assessment of the consequences of this result, we will go on in section 4 to challenge Brandom’s proofs of (1) and (2). This will trigger a number of considerations, summarized in section 5.

2 The Hierarchy of Inferential Relations

Now we take up the discussion about the mutual connections between Brandom’s three semantic inferential relations. It is on the basis of the analysis of the notion of commitment in terms of its dimensions of authority and responsibility as described in §1 that claims (1) and (2) can be proved. To this effect, we can reconstruct the notion of assertion in the following way. Suppose that S is a speaker and p a sentence. What S is doing when he asserts p is undertaking a commitment, which can be spelled out along the two dimensions of responsibility and authority as follows: We say that S is committed to p if and only if:

Responsibility: S has to respond to any challenge to p .

Authority: S authorizes others to be committed to p . This means that S has to be prepared for making his own any challenge to p addressed to other people. Or, conversely, that other people can discharge responsibility to show entitlement to p by deferring it to S .

In the model of giving and asking for reasons, a challenge is a fundamental notion. In line with Brandom’s contention that assertion is the fundamental speech-act (cf. MIE, p. 172), a challenge needs to be reconstructed in terms of assertion, otherwise assertion would not be the basic speech-act Brandom wants it to be. Therefore, one “can challenge an assertion only by making an assertion incompatible with it” (MIE, p. 178).⁵

With this reconstruction at hand, we will first prove that $p \circ \rightarrow q$ entails $p \square \rightarrow q$, that is, claim (1), and next we will proceed to prove claim (2).

Proof of (1): Suppose that $p \circ \rightarrow q$ and that S is committed to p . We will show that S is committed to q by showing (i) that S has to respond to any challenge to q ; and (ii) that S authorizes others to assert q :

(i) Suppose q is appropriately challenged with r , which is incompatible with q . Then p is appropriately challenged with r as well, because $p \circ \rightarrow q$ and therefore r is incompatible with p . In other words, if $p \circ \rightarrow q$ and r is a challenge to q , then r is a challenge to p . Furthermore, since S is committed to p , he has to respond to r . We have just shown that S has to respond to any challenge to q .

(ii) By assumption, S authorizes others to be committed to p . This means that he will make his own any challenge to p addressed to other people. But since any challenge to q is a challenge to p , he also has to make his own any challenge to q . \square

This proves that $p \circ \rightarrow q$ entails $p \square \rightarrow q$. Now, we will prove claim (2), that is, that $p \square \rightarrow q$ entails $p \diamond \rightarrow q$.

Proof of (2): Suppose that $p \square \rightarrow q$ and that S is not entitled to q . This means that if S were to assert q , he could not fulfill his responsibility to show entitlement to q . This would not be the case if S were entitled to p , since he could discharge his responsibility with respect to q by asserting p . The reason is that we have assumed that $p \square \rightarrow q$.⁶ Therefore, S is not entitled to p either. By counterpositive, it follows that if S is entitled to p , he is entitled to q , and we are done. \square

Brandom claims that the order is “strict”—i.e., that the ‘if... then’ in (1) and (2) are not reversible. Although he doesn’t give any explicit reason for this claim, we can give the following charitable explanation. The reason why this hierarchy is not reversible is the following. These three inferential relations are intended to be recognizable in terms of more familiar inferential relations, such as deduction and induction. Commissive and permissive consequences relate to these familiar notions in the following way. The former corresponds to a generalization to the material case of *deductive* inferences, and the latter to a generalization to the material case of *inductive* inferences (cf. BSAD, p. 121; MIE, pp. 168–9). The third inferential relation, namely incompatibility entailment, is characterized by Brandom as a “counterfactual-supporting, *modally* robust inferential relation” (ibid.; cf. §3). For ease of presentation, we will call this latter relation, the *robust inference*. These corresponding inferential relations, namely robust, deductive, and inductive inferences, are, purportedly, *strictly* ordered by strength. This means that the inferential relations to which they each correspond must also be ordered by such a strict order. We will challenge this picture in the next section.

⁵The definition of challenging p by asserting an incompatible q assumes that p and q are recognized as incompatible by both speaker and challenger, and not only that they are so at the level of normative, social statuses.

⁶It comes as no surprise that inheriting entitlement depends on inheriting commitment. For there are only four ways in which a speaker can discharge his responsibility of showing entitlement to a claim, and only one of them is inferential, namely, that a commitment follows from another.

A few words on the relevance of this hierarchy, and a closer examination of it, are in order. As we showed in §1, the semantic content of an assertion is understood in terms of its inferential articulation. In ch. 5 of BSAD, Brandom gives an analysis of the inferential articulation along the dimension of incompatibility entailment. But, as he acknowledges (cf. BSAD, footnote 5, p. 123), this is just *one* dimension. The other two dimensions are to correspond to generalizations to the *material* case of deductive and inductive inferences. That we are dealing with generalization to the material case means, on the one hand, that those inferences are grounded on the meaning of the (logical and non-logical) expressions in the sentences involved—at least for the case of deduction—and, on the other hand, that the present analysis of inferential articulation is dealing with inferences that occur in our everyday-use of sentences and expressions, which presumably include to a large extent inductive inferences. The hierarchy can shed light on how these different inferential practices are organized with respect to each other, and therefore it shows how all the dimensions of the semantic content of expressions are organized, thus giving a more fine-grained analysis of the notion of semantic content.

3 A Closer Scrutiny

Brandom claims (cf. Brandom, Unpub., p. 25) that the ‘if... then’ in (1) and (2) are not reversible. However, under a closer scrutiny, we can see that the claim is not sound. It can be proved that (2) implies that the converse of (1) holds. In other words, from the fact that $p \sqsupset \rightarrow q$ implies $p \diamond \rightarrow q$, we can prove that (3) holds:

$$\text{If } p \text{ committive entails } q, \text{ then } p \text{ incompatibility entails } q. \quad (3)$$

This means that if (1) and (2) hold, incompatibility entailment is *equivalent* to committive consequence. The proof of (3), given (2), is the following:

Proof of (3): Suppose that $p \sqsupset \rightarrow q$. We need to show that $p \diamond \rightarrow q$. Let r be such that it is incompatible with q . We shall prove that r is incompatible with p . Consider the notion of p being incompatible with r paraphrased as follows:

- (i) If S is committed to p , S is not entitled to r .
- (ii) If S is committed to r , S is not entitled to p .

(i) If, on the one hand, we assume that S is committed to p , by the assumption that $p \sqsupset \rightarrow q$ it follows that S is committed to q . Since r is incompatible with q , S cannot be entitled to r .

(ii) On the other hand, if we assume that S is committed to r , since r is incompatible with q , then S cannot be entitled to q . Moreover, since $p \sqsupset \rightarrow q$ implies that $p \diamond \rightarrow q$ (this is the assumption that (2) holds), then S cannot be entitled to p . \square

We have just shown that r is incompatible with p . Therefore, if $p \sqsupset \rightarrow q$, then $p \diamond \rightarrow q$. In view of (1), we have that $p \sqsupset \rightarrow q$ if and only if $p \diamond \rightarrow q$.

As we said above, this hierarchy has consequences for the way in which we can recognize these three inferential relations in terms of more familiar inferential relations. We noted above that committive consequence is intended to correspond to deductive inferences, and that permissive consequence is intended to correspond to ‘robust’ inference. The third inferential relation is characterized as follows:

The fact that the *properties* of being a donkey and being a mammal stand in the relation of incompatibility entailment means that every *property* incompatible with being a mammal is incompatible with being a donkey. . . . We could say: “Necessarily anything that is a donkey is a mammal.” (BSAD, pp. 122, emphasis in the original)

The notion of incompatibility entailment, therefore, has an intrinsic modal ingredient. The question now is whether, under the light of the equivalence between committive consequence and incompatibility entailment, we can keep the promise that committive consequence corresponds to deductive inference. The problem at stake is whether the notion of deductive consequence is the same as, or equivalent to, the notion of robust inference. There are a number of caveats triggered by this consequence, but for reasons of space we cannot develop them here.⁷

What does come as surprise is that, under the assumption of symmetry of the incompatibility relation,⁸ we can prove that (1) implies the reciprocal of (2), that is (4):

If p permissive entails q , then p committive entails q . (4)

Proof of (4): We will prove that if $p \diamondrightarrow q$ then $p \circrightarrow q$. The desired result will follow by (1). Suppose that $p \diamondrightarrow q$ and that r is incompatible with q . The latter assumption implies that if S is committed to r , S cannot be entitled to q . Since entitlement to p implies entitlement to q , it follows that S cannot be entitled to p either. By symmetry (see footnote 8), it follows that r is incompatible with p and we are done. \square

The consequences of (4) are obvious. It implies that all three inferential relations are equivalent. This destroys all chances to maintain the characterization of these inferences in terms of deduction and induction. We cannot hold that committive consequence corresponds to deduction and that permissive consequence corresponds to induction.⁹

4 The Analysis of Commitment

Brandom’s derivation of claim (1) seems to conflate two different notions of committive consequence. One of these notions occurs at the attitude-level, whereas the other occurs at the status-level. On the one hand, there is a notion of committive consequence in terms of the individual ability to attribute a commitment as a consequence of attributing another commitment —i.e., an intrapersonal inheritance of commitment. On the other hand, there is a notion of committive consequence brought about by the social normative status conferred to the individual by the attributions of the members of a linguistic community. Therefore, the notion of commitment to a sentence, as obtained as the result of a committive consequence at the attitude-level, is in principle different from the notion of a commitment to a sentence, as obtained as the result of

⁷There is an issue concerning the concrete definition of robust inference, and the modal vocabulary in terms of it. Moreover, the failure of the principle of extensionality in modal logic requires a more fine-grained analysis of the equivalence between (classical) deduction and robust inference.

⁸The relation of incompatibility is symmetric by definition. What is actually meant is something stronger. It says that if one denies that p and q are incompatible, one is denying both that (i) If S is committed to p , S is not entitled to q , and (ii) If S is committed to q , S is not entitled to p . In other words, symmetry means that (i) implies (ii) and viceversa. We take that this is what Brandom has in mind when he imposes the condition of symmetry over the incompatibility relations that are “suitable as semantic primitives” (cf. BSAD, p. 123).

⁹This claim was a bit odd from the beginning, under the light of (2). The reason is that (2) implies that any deductive inference corresponds to an inductive inference. In other words, since $p \squarerightarrow q$ means that there is a sound deduction from p to q , (2) implies that there is a plausible induction from p to q . This doesn’t make much sense. Deduction and induction are categorically different concepts, not concepts that differ in ‘strength’.

a committive consequence at the status-level. Brandom calls the former an *acknowledged* commitment,¹⁰ and the latter a *consequential* commitment (cf. MIE, pp. 194ff). Both of these notions are at play in Brandom’s reconstruction of the propositional conceptual content of a sentence (cf. MIE, p. 174). But we shall show that the gist of the proof of claim (1) relies on a conflation of these notions of commitment.

It is important to disentangle the individual and the social aspects in the game of giving and asking for reasons. The game consists of a number of participants, each of whom has the ability to keep score of each other’s attributions of commitments and entitlements, as well as treating certain inferences as good (cf. MIE, pp. 182ff). These are individual abilities, as is manifest in Brandom’s insistence that this game is “doubly perspectival”, as opposed to, say Baseball, where there is one single official score (cf., particularly, MIE, p. 185). On the one hand, *S*’s assertion that *p* consists in undertaking a commitment, which, among other things, implies a self-attribution of commitment — an acknowledged commitment. On the other hand, what commitments follow from *S*’s assertion that *p* are evaluated from the perspective of each of the participants, including *S* himself. But these commitments need not be the same for everyone. This creates a difference between consequential and acknowledged commitments.

To make the point clearer, consider the case where *S* is committed to incompatible claims. To refrain from considering *S* as irrational, one needs to make a difference between *S*’s own commitments and what *S* is committed to according to the community —i.e., there are commitments that are not incompatible according to him, but that are so according to the community. For instance, suppose that, according to *S*’, $r \Box \rightarrow q$, but this is not so for *S*. Furthermore, suppose that *p* and *q* are incompatible (according to both *S* and *S*’). If *S* commits himself to *p* and *r*, he need not be attributing himself incompatible commitments. But according to *S*’, *S* is also committed to *q*. This results from *S*’s committive consequence from *r* to *q*. Therefore, according to *S*’, *S* is making incompatible commitments, since *p* and *q* are incompatible and *S*’ is attributing *S* commitment both to *p* and *q*. In this example, *S* is not committed to *q* in this sense that it is not one of his *acknowledged* commitments. On the other hand, *S* is committed to *q* from the point of view of *S*’ —the *consequential* commitment.¹¹

That this difference is important can be seen from Brandom’s insistence that “this is one of the benefits of this sort of approach over causal-functional accounts of intentional states” (MIE, p. 196). The reason is that one can attribute oneself incompatible commitments —i.e., commitments that are not incompatible for oneself, but that are so according to the community—, and thus one can be mistaken, that is, one can be ‘illogical’, and yet rational.

The difference between the status- and the attitude-level is also important to the extent that Brandom claims that:

The way in which the collaboration of attitudes adopted from two socially distinct perspectives —attributions of commitments to oneself and by others— is required to institute discursive commitments is the central theme of this work. It is in terms of the social-perspectival character of discursive deontic statuses that the notion of *objectivity* is to be made intelligible (MIE, pp. 166f).

The two levels of acknowledged and consequential commitments are conflated in the proof of (1). Recall that Brandom is discussing *intrapersonal* inheritance of commitments. This means that the discussion is centered around the practical ability of

¹⁰Actually, an acknowledged commitment is a commitment self-attributed, regardless of whether it was self-attributed as part of making an assertion, or as a consequence of an intrapersonal inheritance of commitment.

¹¹There is also the issue as to whether *S*’ can be mistaken in his assessment that $r \Box \rightarrow q$. Moreover, we are not told a story as to how to compute consequential commitments out of the individual scores kept by the participants (cf. below).

each participant to be disposed to treat himself as committed to q as a consequence of treating himself as committed to p , and to project this inference on the rest of the participants. Therefore, the discussion is conducted at the level of *acknowledged* commitments. *Consequential* commitments, since they are essentially social, cannot be the focus of the discussion. Yet, the proof of (1) introduces the social dimension of authority and responsibility. Along the latter dimension, for instance, Brandom shows that if $p \circ \rightarrow q$ (for S), and if S undertakes the commitment that p , then S has to respond to any challenge to q . The missing step is to go from the social status of ‘having to respond to a challenge to q ’ back to S ’s individual self-attribution of commitment to q . But this step is not allowed in as much as we stick to a difference between the social- and the attitude-level of commitments.

A similar analysis can be conducted to the proof of (2). In other words, the proof of (2) also relies in a conflation of the two levels of *entitlement*.

5 Conclusion

First, we should note that the entailment from $p \diamond \rightarrow q$ to $p \circ \rightarrow q$ relies on the assumption of the symmetry of the incompatibility relation. This assumption seems to have consequences for the metamathematical results on the incompatibility relation (cf. BSAD, ch. 5). In as much as one feels compelled to keep these results, one must stick to this assumption and, therefore, to the entailment in question.

Secondly, the idea that incompatibility entailment, committive consequence, and permissive consequence correspond to robust inference, deductive inference, and inductive inference respectively, must be abandoned —at least if robust inference is meant to be something different from deduction. There are no three different notions of inference, but just one. This result might have consequences with respect to the socrekeeping dynamics described in ch. 3 of MIE.

Thirdly, in view of the equivalences obtained in §3 and an analysis of the notion of commitment, it will be interesting to measure the extent to which different participants of the game of giving and asking for reasons can have different inferential attitudes and different attributions of incompatibilities. *Prima facie*, this discussion seems to be at odds with the possibility of maintaining a difference between the status- and the attribution-level of commitments, as discussed in §4.

Fourth, it follows from our discussion in §4, although we cannot argue for it in detail for reasons of space, that the appropriate level for the discussion of the hierarchical relations between the three semantic inferential relations, and therefore for the discussion of semantic content, is the *status*-level. But unlike the inferential relations at the attribution-level, Brandom doesn’t spell out when someone is committed to a sentence at the status-level. Is S committed to p if all the participants, including himself, attribute him commitment to p ? But since there are differences between these two levels, this requirement is too strict. How many participants then should attribute S commitment to p in such a way that S is committed to p at the status level? Half of them? Most of them? Only the relevant ones?

In this paper, we have explored the internal logical relations between the definitions offered by Brandom of these three kinds of entailment relations. From a different, external perspective, we suspect that Brandom may be misguided in focusing on incompatibility-entailment rather than on committive consequence as the most interesting notion of entailment/consequence from the point of view of his pragmatist inferentialism. Indeed, (under a different formulation) committive consequence may do better justice than incompatibility-entailment to the pragmatist and in particular *constructivist* intuitions and assumptions that are at the heart of his inferentialist philosophy (cf. his references to Dummett in MIE). Under a different formulation, committive consequence may define a constructivist logic, perhaps similar to intuitionistic logic, in

which case the hierarchy might end up being: committive consequence (roughly similar to intuitionistic logic) \Rightarrow incompatibility-entailment (classical logic) \Rightarrow permissive consequence (inductive logic). But for now, this shall remain as a suggestion for future work.

References

- Brandom, R. B., 1994: *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press.
- 2008: *Between saying and doing. Towards an Analytic Pragmatism*. Oxford University Press.
- Unpub.: Conceptual content and discursive practice, Unpublished Manuscript.

Intentionality and the Philosophy of Mind

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 it is 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.

Mind the consequences of inferentialism and normativism: conceptual mental episodes ain't in the head (at all)

Pierre Steiner

Pierre.Steiner@utc.fr

Webpage: <http://php.mental-works.com/~costech/v2/pages/infos-chercheur.php?id=82>

Université de Technologie de Compiègne – COSTECH/CRED – BP 60319, 60203 Compiègne Cedex, France

Abstract. In this short paper, we briefly expose a position dubbed *Vehicleless Externalism for Conceptual Mental Episodes (VECME)*. According to it, the constitutive relations there are between the production of conceptual mental episodes by an individual and the inclusion of this individual in social discursive practices make it impossible to equate, even partially, conceptual mental episodes with the occurrence of intracranial events. Conceptual mental episodes do not have subpersonal vehicles; they have owners: persons in interpretational practices. In Section 1, the context and the scope of this position, notably in Brandom's philosophical system, is clarified. In Section 2, we situate the roots of *VECME* in a marriage between inferentialism and normativism. Section 3 presents some aspects of *VECME*.

1 Externalism(s) and the dependences between mental activities and linguistic environments

Let us begin our argument with a quote from Wilfrid Sellars, a quote itself often used by Brandom: "Clearly human beings could dispense with all discourse, though only at the expense of having nothing to say" (Sellars 1980, p.152). In the Brandomian framework, the point here is not merely to hold that the framing of new ideas, intentions or beliefs by an individual depends on the fact this individual enjoys discursive abilities. Indeed, discursive or linguistic abilities only exist as special kinds of practices, nested in social practices. It is rather the fact some individual is a member of these social practices that enables him to acquire linguistic capacities, and thus a conceptually-articulated form of mental life.

Whereas numerous contemporary philosophers also remind us that language is a constitutive tool for the development of mental life¹ (Carruthers, Clark, Dennett), Brandom reminds us that the use of that tool is not free. If language is a tool, it is a tool that *enables* us to do and think about many things because its use is based on

¹ For instance, for abilities such as memorizing, categorizing, introspecting, reasoning, metarepresenting, formalizing, acquisition and formation of new concepts.

constraints we freely abide by. These normative constraints (inheritance of normative statuses for instance) define or constitute the meaningfulness of the new abilities and performances we gain by using these tools. These normative constraints are related to social practices. These constraints we bind ourselves by define the space in which meaning and expressive freedom may appear. This is a basic difference between normative and natural constraints (Brandom 1979).

Besides these socially-mediated causal relations between linguistic tools and mental activity, one generally considers that there is another essential relation between our conceptual mental life and the linguistic social practices we are embedded in: classical semantic externalism (Putnam, Burge) considers that there are important taxonomic relations between the referential dimension of the contents of one's beliefs (at least the ones involving some kinds of concepts) and the linguistic environment "in" which these thoughts occur. Still, this externalist framework leaves intact the possibility that mind and mental episodes are still in the head. Even if the referential dimension or even all the semantic dimensions of content can only be individuated if one considers the relations an agent entertains with its environment, the bearers or the vehicles of this content have an intracranial localization (they are physically specifiable entities). We here have what McDowell and McGinn called a *Duplex Conception of the Mental*: mental episodes are in the head (that is: their vehicles are in the head); but they should be conceived, attributed or individuated in terms of their relations to what is outside the head.

We want to argue here, following amongst others Brandom, that the fact individuals are members of social practices is not only a cause or an enabling condition of their having a meaningful cognitive life, and a descriptive constraint one should respect in the individuation of their mental contents: it entails the fact that the conceptual mental lives of these individuals are not in their head. Or, at least: conceptual mental episodes (*CME*) such as judging, thinking, considering or intending that p^2 do not have vehicles, at least in the authoritative sense in philosophy of mind: internal material entities or events that, by intrinsically providing or encoding some kind of content would, modulo some functional and causal roles, be the realizers of these *CME*. To put it otherwise: *CME* ain't in the head at all – even in a layered model of what happens in the intracranial world. They are not realized in, supervenient on or constituted by the occurrence of inner events. This thesis we dub Vehicleless Externalism on Conceptual Mental Episodes (*VECME*).

In an abstract way, one might figure a vehicle as what remains in one event of thinking that p if one descriptively strips this event of all its intentional and phenomenal properties: a nonintentionally specifiable phenomenon is supposed to remain; this meaningless bundle of subpersonal material properties can be described and individuated with reference to its formal, syntactical, or neurophysiological properties. Still, it is a content-*providing* material entity (Clark 2006); it is an entity that comes to *bear* some content (Rowlands 2003, p.156). We here understand the 'providing relation' in a twofold sense: first it might mean the vehicle does *encode* the

² Beliefs are not here included in *CME*: beliefs are mental *states* (whose occurrences are judgments). Still, if one wants to consider beliefs as mental episodes, then one should see them, in a Brandomian framework, as *commitments to the truth of a proposition*.

very same content as the content of the *CME*, as it has the same formal structure as this content. Second it might mean the vehicle encodes some content (at *t*) that causally explains the simultaneous presence of the personal-level content without being a formally-conform translation of it. The first case applies to a symbolic conception of the vehicles of *CME* (the vehicles of *CME* are Mentalese sentences). The second case is proper to subsymbolic conceptions of the vehicles of *CME*.

According to *VECME*, if *CME* are not encoded or realized in inner vehicles, they are not encoded or realized in external vehicles as well, that is, in external material entities or substances (books, diaries, sentences, pictures,...). *CME* rather have *owners*: persons (and not subpersonal parts of them, or impersonal items they would manipulate), actively standing in interpretational and institutional practices.

The basic idea behind *VECME* is not new³. Brandom devoted a paper to it in 2004 anyway. It can even be seen a consequence of an idea he already defended in 1994: neither linguistic intentionality nor mental intentionality have foundational priority in the explanation of mental and linguistic activities (1994, p.151). Both of them are inherited from the intentionality of social practices. Still, I believe the originality of *VECME* has not been sufficiently appreciated, and especially its importance in Brandom's system. Actually, it might even somehow be strange of trying to understand Brandom's originality in recent (analytic) philosophy of language by still (possible) endorsing a basic assumption of (analytic) contemporary philosophy of mind⁴ and philosophy of language that Brandom's work also attempts to undermine. This main assumption is vehicle-internalism: the idea that the conceptual mental episodes agents produce in the game of giving and asking for reasons ultimately have an intracranial localization (although they would be individuated and attributed from social practices). On this vehicle-internalist picture, as public productions, assertions are *expressions* of mental phenomena; they *manifest* the presence of intracranially-located phenomena (beliefs, intentions, meaning,..) that, at least partially (for the asserter), endow these assertions with meaning (before inferential, environmental, indexical and contextual components are considered).

VECME can be seen as a logical consequence of both Brandom's sophisticated inferentialism and normativism (the thesis that meaning is constitutively normative). If it is correct, *VECME* makes it very difficult to endorse both an inferentialist and normativist model of linguistic activities and a traditional vehicle-internalist model of mental episodes, as Mark Lance for instance does by holding that, in his inferentialist story, "to attribute a belief is, roughly, to attribute an internal state which disposes one to make a public move of asserting a propositional content within the social game of giving and asking for reasons." (Lance 1998, p.220).

³ Philosophers such as Wittgenstein, Arthur Collins, Lynne Ruder-Baker or Vincent Descombes endorsed or endorse it. Note that some of Brandom's main philosophical influences (Sellars, Rorty) do not. Ditto for Davidson.

⁴ Anomalous monism, functionalism, reductionism, supervenience, or dual aspect theories all endorse vehicle-internalism.

2 The normative and inferentialist grounds of *VECME*

A controversial version of the “meaning is normative” idea is the following (and we will adopt it here): there is no meaning prior to the norms that define this meaning. That is: meaning is constitutively normative. It is not only the case that means implies some oughts: more fundamentally, “for every means, there is an ought that implies it” (Gibbard 1994, p.101). Still, we will see, these *oughts* are more constraints than obligations. For this version of the “meaning is normative” slogan, it is not only the case that applying a concept can sometimes oblige one to apply another concept, often preclude one from applying a different concept, or permit one to apply other concepts: these normative constraints constitute the content of the initial concept (and the correct conditions of its mastery as well).

This essential normative character of meaning is here best understood in an inferentialist conception of meaning (but note that inferentialism does not entail normativism, and conversely).

As it is well known, according to inferentialism, conceptual content should not be (primarily) analysed in terms of reference or representation, but in terms of inference, more precisely in terms of inferential relations between conceptual content and other contents, but also perception and action (which are non-inferential entries and exits of a system). For the inferentialism we consider here, the meaning of a sentential expression (and not only of a logical concept) is its use, role, or function within its language, including perceptual entries and behavioural exits. Inferential relations constitute the inferential role of the content, which is to be equated with its meaning. But meaning is not just inferential role. This inferential role is indeed the product of inferential rules. The inferential uses, roles, or functions are thus framed and constrained by inferential rules. The meaning of a linguistic concept or expression therefore consists in its *normatively framed* inferential role in some language. This normative foundation of inferences makes inferential rules, and not inferences or inferential relations, the building stone of semantics.

From this point of view, the contentfulness of *CME* consists in their inferential role, being governed by inferential rules that are socially instituted and used (and not being governed by mere causal-functional relations within one’s cognitive economy, as some functional role semantics for mental content hold). These inferential rules - typically expressed by conditionals, material or formal - do not primarily consist in obligations for speakers or believers; they rather constraint our linguistic practices by delimitating what, on an inferential point of view, we may and may not do by entertaining conceptual contents. They should better be seen as normative uniformities characterizing the pattern-governed behaviours of speakers. Negative uniformities play the most important role; the rules that govern them are therefore to be construed as constraints, rather than obligations or incentives (Sellars 1974; Peregrin 2006). Inferential rules do not prescribe what ought to be done, but what is allowed and what is not allowed to do when one asserts that *p*.

For Brandom’s inferentialism, the inferences that articulate conceptual contents are mainly material inferences (and not primarily and necessarily formal inferences); they are cashed out in terms of normative statuses (entitlements and commitments), objects of normative attitudes (undertakings (or acknowledgement), attributions (or ascriptions)). The content of what we say and judge is inferentially articulated by

being pragmatically determined in normative practices of scorekeeping. These practices confer contents to performances; they institute and exercise the linguistic norms concerning how it is correct to use expressions, under what circumstances it is appropriate to produce some conceptual performances and what the appropriate consequences of these performances are. Inferential relations are thought of in terms of preservation, inheritance and exclusion relations among commitments and entitlements. The crucial point here is that these statuses do not supervene on physical inner properties of their owners: even when they are *undertaken* by an individual, their significance is always related to their *attribution* by other scorekeepers and to the *attribution* (by other scorekeepers as well) of other statuses that follow from them and that define their content. This point is not only crucial when we consider the content of *CME*: it is central when it comes to define what *judging* some content is.

Judging that *p* (*in fore interio*) is endorsing *p*; it is undertaking a propositionally contentful commitment to *p* (Brandom 2004, p.244). The conceptual content (inferential role) of our judgments is mainly determined by the commitments and entitlements that are (or would be) attributed to us from the commitments we undertake in this judging. Brandom's inferentialism comes with a fine-grained interpretationism or phenomenalism. For interpretationism, any performance can be contentful only in virtue of the fact that it is or can be treated as such in content-conferring practices. What one is committed or entitled to, depends on what one is *taken to be* committed or entitled to. It is misleading to say that for this interpretationism, *S* is judging that *p* if and only if someone interprets (or would be ready to interpret) him as judging that *p*. It is better to say that *S* is judging that *p* if and only if he is able to acknowledge the various deontic statuses that some member(s) of the linguistic community would be ready to attribute to him and that inferentially define the content of its judging. Attributions and ascriptions need not be actual; they might just be potential. Even actual, they are more implicit than explicit: implicit attribution can equate with the tracking of *S*'s inferential moves and with dispositions to correct them. The members of the community do not primarily attribute mental episodes: they rather attribute the deontic statuses that define the inferential scope of the contents of these episodes (but also what these episodes *are*: intendings, judgments,...), and are ready to correct deviant uses of concepts. For this interpretationism, the nature of the propositional content of a mental episode (thinking, fearing, intending, judging that *p*) is defined by its pragmatically and socially-mediated inferential articulation.

In this interpretationism, nobody can judge that *p* if he is not recognizable as mastering the inferential articulation of *p* in attributed commitments and entitlements. Understanding content is knowing the circumstances in which one becomes committed or entitled towards this content, but also the consequential commitments and entitlements one inherits by endorsing this content. Judging that *p* is not necessarily judging that one ought to judge various things entailed by *p*; it is being disposed to move to and to accept (or to make accept) these entailed contents. It is not actual interpretation that matters here, but interpretability. Only interpretability is required for someone genuinely to be binding herself by conceptual norms. This interpretability requires the mastery of the pragmatically-determined inferential proprieties that articulate the contents of *CME*. A good criterion enabling us to

establish the presence of this practical conceptual knowledge by an agent is the fact he is disposed to correct deviant inferential uses of the content made by other agents. Let us now turn to the consequences of inferentialism and normativism for *VECME*.

3 Conceptual mental episodes, discursive practices, and intracranial processes

Judging something is undertaking a commitment to some content that is defined in terms of other commitments and entitlements. Judging constitutively includes undertaking a commitment to the truth of a proposition. This normative *stand* can only exist in normative practices (Brandom 2004, p.244), where *one* is able to justify this commitment and to exhibit inferential mastery of this content (especially its relations with actions). The normative statuses (commitments, entitlements) that make up the contents of judgments and that define what it is for some agent to judge something are the products of social practical attitudes (Brandom 2004, p.249). Their gain or loss is therefore not reflected at all in some intrinsic properties of their owner. Undertakings of commitments, endorsing claims are social proprieties, not natural or material properties that could be instantiated in subpersonal inner events. No subpersonal item can engage in these practices; these practices have as objects the abilities of subjects, not what happens inside of their crane. If the brain can exhibit *responses* to various states of affairs, these responses cannot constitute *our responsibilities* or *undertakings of commitments* to things being thus-and-so.

The status of being committed to the truth of *p* or to take a propositionally contentful *stand* is not vehicled by inner processes, for this status can only be realized by *persons* in social-interpretational (scorekeeping) practices. From this point of view, judging that *p* is like borrowing a book at the library or being a mayor: these social statuses do not supervene on the intrinsic properties of their owner. Two individuals might be internally physically indiscernible, one and not the other might yet be a believer, a mayor or a book borrower only *in virtue of* the environmental social practices he is *actively* embedded in.

It is not only content that is determined by inferential norms. More basically, *judging* some content is itself determined by practices that confer the *having* of contents *by* attributing contentful commitments and entitlements. It is not only impossible to refer to some things without the appropriate environment: it is impossible to produce meaningful thoughts without actively standing in these normative practices. It is not only the case that *p* is inferentially and socially determined: *judging* it is a social status, an attributed or attributable commitment to the truth of *p* (whose undertaking comes with other commitments and entitlements, practically mastered by the person). This status belongs to a *person*, nothing inside of this person helps to realize it or constitutes it. *All* depends on the nature of the social practices in which the person is, and on her abilities to inferentially use the content of what she judges and so of committing herself to other judgments. The social practices define the content of the judgment, but also the behavioural and inferential factors in virtue of which the person is interpretable as *judging* that *p* (these behavioural factors include the fact the person can herself be an interpreting person). Normative practices do not attribute

judgements; yet it is only by actively being a member of them (notably by being able to inferentially use conceptual contents) that a person is interpretable as entertaining contents *in fore interio* – that is, commitments that are judgments (and not assertings).

No inner event can give or amount to the endorsement and commitment to some conceptual content. Intracranial processes are not only non-sufficient for that, they are non-necessary as well. From this claim should follow the claim that *CME*, as social status (commitments), are not even *partially* realized in or by some physical substance. There is no *first* an inner event that would *then* become an intentional episode, or an event that could be both described in nonintentional and intentional terms. The non-intentional specification of what happens in my head when I judge that *p* is unable to shed any explanatory lights on the happening of judging that *p* (and not only on the reference or on the content of this judging).

VECME can accept that *cognitive* operations and events could exist independently of their potential attribution by mental agents – but this does not apply to *CME*, which are necessarily attributable or attributed social status (commitments to the truth of a proposition). If it turned out that *S* was unable to justify or to inferentially articulate the content *p* of the *CME* he claims to have (notably by showing brain-scans proving that “there is something in there”), we might refuse to attribute him the endorsement and thus the judging of *p* (and thus the *CME*). Conversely, more than often, we attribute *CME* and conceptual commitments to agents without looking at what happens inside of their head. All that matters are the behavioural capacities of the agents, as long as we can understand them as being inferential capacities (that is, as abilities to engage in pattern-governed semantic behaviour).

The point is not only to say that the subpersonal internalist specifications of the so-called vehicles of *CME* cannot make intelligible the interpersonal normative properties in virtue of which personal mental episodes *are CME* (that is, endorsements of conceptual contents). It is to say that these interpersonal normative properties are essential in the definition of *CME*, and that they do not have necessary relations *at all* with subpersonal physical operations (which are thus *useless* for the definition and identification of *CME*). What is in the head is neither sufficient nor necessary for the occurrence of the essential features of *CME*: their normative properties. Nothing inside the head can *encode* the content persons *express* in their judgments. For the latter one is necessarily undertaken by the person, in virtue of its active insertion in linguistic practices. Nothing subpersonal can *do* what persons as *persons do*.

One might argue that the fact some agent does not have inferential mastery of the content of some *CME* is a matter of intracranial parameters. But inferential skills are *inferential* skills only in virtue of inferential norms instituted by the community. Their exercise depends on intracranial processes, but the latter ones do not constitute what are correct or incorrect inferential skills and knowledge. True, sometimes neuropsychological damages may *cause behavioural deficiencies* that can be *taken as* inferential behavioural deficiencies. But what *causes* the deficiency does not *constitute* the rightness or wrongness of the deficient behaviour. Behavioural deficiencies might put into question the fact the subject really entertains some judgments or assertings only depending on our interpretations and charity principles – and not from what the brain would show or tell us. Behavioral inferential capacities

are pattern governed behaviour abilities only from the outside, from actual or potential interpretations. There are only in the head the material conditions and operations that are empirically required for the possession and exercise of the behavioural skills that serve as *criteria* for the attribution of social statuses and thus mental attitudes and contents. Behavioural skills and performances are criteria for the phenomenal presence of the *CME* not because they are causally produced by intracranial operations, but because of our conceptions of what must contextually be the behavioural criteria of the presence of a *CME*.

When we tend to attribute mental episodes to other agents, we attribute them social statuses, as abilities to accomplish the pattern-governed behaviours defining the meaning of the contents they judge or think. Brain factors do not influence our attributions. Sure, they influence the behavioural capacities of agents. But not in a way that is constitutive of the normative character of these capacities (making them pattern governed behaviour similar to the undertaking of commitments and entitlements). Intracranial processes rather causally influence the production of the pattern governed behaviour on the basis of which persons are interpretable as the owners of these *CME*. Put otherwise, inner events do not cause the occurrence of *CME*; they rather contribute to cause the occurrence of behavioural skills (which are *inferential*, *correct* or *appropriate* outside – they are skills for pattern governed behaviour) on the basis of which commitments and entitlements are attributable to the agent.

4 Conclusion

The point of *VECME* is definitely not to hold that the conditions for identifying and having meaningful thoughts are not in the head; it is to hold that thoughts themselves ain't there *at all*. *VECME* follows the lead of what Bjørn Ramberg (2000) has called *post-ontological* philosophy of mind, for which the difference between minded and non-minded creatures is not defined by some mind-independent ontological facts of the matter, but is rather related to our descriptive policies (vocabularies) and especially here to the *deontological* facts that support intentional ascriptions (“attributing an intentional state is attributing a normative status”).

So far, we have not talked about the relations between *CME* and intracranial events yet. The last paragraphs were about the relations between inferential behaviour and intracranial events: we hope we have shown why the latter ones were not *causes* of the former ones, so that intracranial events cannot not be seen as the realizers of *CME* that would be *correctly* attributed to an individual from his/her behaviour *because* that behaviour would be caused by the realizers of *CME*.

In order to tackle the issue of the relations between *CME* and intracranial events – and to better understand how the latter ones are not (even partial) vehicles of the former ones –, it is enlightening to systematically consider the differences between the two basic meaning-use relations⁵ of the intentional and physicalist vocabularies respectively: what one must do in order to (sufficiently) deploy the specified

⁵ In Brandom's (2008) sense.

vocabulary, and what the specified vocabulary allows us to specify when a person engages in conceptual mental life. The respective *causal* properties of each vocabulary might then usefully be clarified.

References

- Brandom, R. (1979), "Freedom and Constraint by Norms", *American Philosophical Quarterly*, 16 (3), pp.187-196.
- Brandom, R. (1994), *Making It Explicit. Reasoning, Representing, and Discursive Commitment*, Cambridge (MA)/London, Harvard University Press.
- Brandom, R. (2004), "From a Critique of Cognitive Internalism to a Conception of Objective Spirit: Reflections on Descombes' Anthropological Holism", *Inquiry*, 47, pp.236-253.
- Brandom, R. (2008), *Between saying and doing. Towards an analytic pragmatism*, Oxford, Oxford University Press.
- Clark, A. (2006), "Material Symbols", *Philosophical Psychology*, Vol.19, No.3, June 2006, pp.1-17.
- Gibbard, A. (1994), "Meaning and Normativity", *Philosophical Issues*, Vol.5, Truth and Rationality, pp.95-115.
- Lance, M. N. (1998), "Some reflections on the sport of language", *Philosophical Perspectives*, 12, pp.219–240.
- Peregrin, J. (2006), "Meaning as an Inferential Role", *Erkenntnis*, 64, pp.1-36.
- Sellars, W. (1974), "Meaning as Functional Classification", *Synthese*, 27, pp.417-437.
- Sellars, W. (1980), *Pure pragmatics and possible worlds. The early essays of Wilfrid Sellars*, J.F. Sicha (Ed.), Atascadero (CA), Ridgeview.
- Ramberg, B. (2000), « Post-Ontological Philosophy of Mind: Rorty versus Davidson », in Brandom, R. (ed.), *Rorty and his critics*, Oxford/Malden, Blackwell pp.351-370.
- Rowlands, M. (2003), *Externalism. Putting Mind and World Back Together Again*, Montreal & Kingston/Ithaca: McGill-Queen's University Press.

How Analytic Philosophy Has Failed Cognitive Science

Robert Brandom

I. Introduction

We analytic philosophers have signally failed our colleagues in cognitive science. We have done that by not sharing central lessons about the nature of concepts, concept-use, and conceptual content that have been entrusted to our care and feeding for more than a century.

I take it that analytic philosophy began with the birth of the new logic that Gottlob Frege introduced in his seminal 1879 *Begriffsschrift*. The idea, taken up and championed to begin with by Bertrand Russell, was that the fundamental insights and tools Frege made available there, and developed and deployed through the 1890s, could be applied throughout philosophy to advance our understanding of understanding and of thought in general, by advancing our understanding of concepts—including the particular concepts with which the philosophical tradition had wrestled since its inception. For Frege brought about a revolution not just in *logic*, but in *semantics*. He made possible for the first time a *mathematical* characterization of meaning and conceptual content, and so of the structure of sapience itself. Henceforth it was to be the business of the new movement of analytic philosophy to explore and amplify those ideas, to exploit and apply them wherever they could do the most good. Those ideas are the cultural birthright, heritage, and responsibility of analytic philosophers. But we have not done right by them. For we have failed to communicate some of the most basic of those ideas, failed to explain their significance, failed to make them available in forms usable by those working in allied disciplines who are also professionally concerned to understand the nature of thought, minds, and reason.

Contemporary cognitive science is a house with many mansions. The provinces I mean particularly to be addressing are cognitive psychology, developmental psychology, animal psychology (especially primatology), and artificial intelligence. (To be sure, this is not all of cognitive science. But the points I will be making in this paper are not of similarly immediate significance for such other subfields as neurophysiology, linguistics, perceptual psychology, learning theory, and the study of the mechanisms of memory.) Cognitive psychology aims at reverse-engineering the human mind: figuring out how we do what we do, what more basic abilities are recruited and deployed (and how) so as to result in the higher cognitive abilities we actually display. Developmental psychology investigates the sequence of stages by which those abilities emerge from more primitive versions as individual humans mature. Animal psychology, as I am construing it, is a sort of combination of cognitive psychology of non-human intelligences and a phylogenetic version of ontogenetic human developmental psychology. By contrast to all these empirical inquiries into actual cognition, artificial intelligence swings free of questions about how any actual organisms do what they do, and asks instead what constellation of abilities of the sort we know how to implement in artifacts might in principle yield sapience.

Each of these disciplines is in its own way concerned with the empirical question of how the trick of cognition is or might be done. Philosophers are concerned with the normative question of what counts as doing it—with what understanding, particularly discursive, conceptual understanding consists in, rather than how creatures with a particular contingent constitution, history, and armamentarium of basic abilities come to exhibit it. I think Frege taught us three fundamental lessons about the structure of concepts, and hence about all possible abilities that deserve to count as concept-using abilities.¹ The conclusion we should draw from his discoveries is that concept-use is intrinsically stratified. It exhibits at least four basic layers, with each capacity to deploy concepts in a more sophisticated sense of ‘concept’ structurally presupposing the capacities to use concepts in all of the more primitive senses. The three lessons that generate the structural hierarchy oblige us to distinguish between:

- concepts that only *label* and concepts that *describe*,

¹ It ought to be uncontroversial that the last two of the three lessons are due to Frege. Whether he is responsible also for the first is more contentious. Further, I think both it and a version of the second can be found already in Kant. (As I argue in my 2006 Woodbridge Lectures, “Animating Ideas of Idealism: A Semantic Sonata in Kant and Hegel,” forthcoming in the *Journal of Philosophy*.) But my aims here are not principally hermeneutical or exegetical—those issues don’t affect the question of what we philosophers ought to be teaching cognitive scientists—so I will not be concerned to justify these attributions.

- the *content* of concepts and the *force* of applying them, and
- concepts expressible already by *simple* predicates and concepts expressible only by *complex* predicates.

AI researchers and cognitive, developmental, and animal psychologists need to take account of the different grades of conceptual content made visible by these distinctions, both in order to be clear about the topic they are investigating (if they are to tell us how the trick is done, they must be clear about exactly which trick it is) and because the empirical and in-principle possibilities are constrained by the way the abilities to deploy concepts in these various senses structurally presuppose the others that appear earlier in the sequence. This is a point they have long appreciated on the side of basic *syntactic* complexity. But the at least equally important—and I would argue more conceptually fundamental—hierarchy of *semantic* complexity has been largely ignored.

II. First Distinction: From Labeling to Describing

The Early Modern philosophical tradition was built around a *classificatory* theory of consciousness and (hence) of concepts, in part the result of what its scholastic predecessors had made of their central notion of Aristotelian forms. The paradigmatic cognitive act is understood as classifying: taking something particular as being of some general kind. Concepts are identified with those general kinds.

This conception was enshrined in the order of logical explanation (originating in Aristotle's *Prior Analytics*) that was common to *everyone* thinking about concepts and consciousness in the period leading up to Kant. At its base is a doctrine of *terms* or *concepts*, particular and general. The next layer, erected on that base, is a doctrine of *judgments*, describing the kinds of classificatory relations that are possible among such terms. For instance, besides classifying Socrates as human, humans can be classified as mortal. Finally, in terms of those meta-classifications grouping judgments into kinds according to the sorts of terms they relate, a doctrine of *consequences* or *sylogisms* is propounded, classifying valid inferences into kinds, depending on which classes of classificatory judgments their premises and conclusions fall under.

It is the master-idea of classification that gives this traditional order of explanation its distinctive shape. That idea defines its base, the relation between its layers, and the theoretical aspiration that animates the whole line of thought: finding suitable ways of classifying terms and judgments (classifiers and classifications) so as to be able to classify inferences as good or bad solely in virtue of the kinds of classifications they involve. The fundamental meta-conceptual role it plays in structuring philosophical thought about thought evidently made understanding the concept of classifying itself a particularly urgent philosophical task. Besides asking what differentiates various kinds of classifying, we can ask what they have in common. What is it one must *do* in order thereby to count as *classifying* something as being of some kind?

In the most general sense, one classifies something simply by responding to it differentially. Stimuli are grouped into kinds by the response-kinds they tend to elicit. In this sense, a chunk of iron classifies its environments into kinds by rusting in some of them and not others, increasing or decreasing its temperature, shattering or remaining intact. As is evident from this example, if classifying is just exercising a reliable differential responsive disposition, it is a ubiquitous feature of the inanimate world. For that very reason, classifying in this generic sense is not an attractive candidate for identification with conceptual, cognitive, or conscious activity. It doesn't draw the right line between thinking and all sorts of thoughtless activities. Pan-psychism is too high a price to pay for cognitive naturalism.

That need not mean that taking *differential responsiveness* as the genus of which *conceptual classification* is a species is a bad idea, however. A favorite idea of the classical British empiricists was to require that the classifying response be entering a *sentient* state. The intrinsic characters of these sentient states are supposed to sort them immediately into repeatable kinds. These are called on to function as the *particular* terms in the base level of the neo-Aristotelian logical hierarchy. *General* terms or concepts are then thought of as sentient state-kinds derived from the particular sentient state-kinds by a process of *abstraction*: grouping the base-level sentient state-repeatables into higher-level sentient state-repeatables by some sort of perceived *similarity*. This abstractive grouping by similarity is itself a kind of classification. The result is a path from one sort of consciousness, sentience, to a conception of another sort of consciousness, sapience, or conceptual consciousness.

A standing felt difficulty with this empiricist strategy is the problem of giving a suitably naturalistic account of the notion of sentient awareness on which it relies. Recent information-theoretic accounts of

representation (under which heading I include not just Fred Dretske's theory, which actually goes by that name, but others such as Jerry Fodor's asymmetric counterfactual dependence and nomological locking models²) develop the same basic differential responsiveness version of the classic classificatory idea in wholly naturalistic modal terms. They focus on the information conveyed about stimuli—the way they are grouped into repeatables—by their reliably eliciting a response of one rather than another repeatable response-kind from some system. In this setting, unpalatable pan-psychism can be avoided not, as with traditional empiricism, by insisting that the responses be sentient states, but for instance by restricting attention to flexible systems, capable in principle of coming to encode many different groupings of stimuli, with a process of *learning* determining what classificatory dispositions each one actually acquires. (The classical American pragmatists' program for a naturalistic empiricism had at its core the idea that the structure common to evolutionary development and individual learning is a Test-Operate-Test-Exit negative feedback process of acquiring practical habits, including discriminative ones.³)

Classification as the exercise of reliable differential responsive dispositions (however acquired) is not by itself yet a good candidate for *conceptual* classification, in the basic sense in which applying a concept to something is *describing* it. Why not? Suppose one were given a wand, and told that the light on the handle would go on if and only if what the wand was pointed at had the property of being *grivey*. One might then determine empirically that speakers are grivey, but microphones not, doorknobs are but windowshades are not, cats are and dogs are not, and so on. One is then in a position reliably, perhaps even infallibly, to apply the *label* 'grivey'. Is one also in a position to *describe* things *as* grivey? Ought what one is doing to qualify as applying the *concept grivey* to things? Intuitively, the trouble is that one does not know what one has found out when one has found out that something is grivey, does not know what one is taking it to be when one takes it to be grivey, does not know what one is describing it *as*. The label is, we want to say, uninformative.

What more is required? Wilfrid Sellars gives this succinct, and I believe correct, answer:

It is only because the expressions in terms of which we describe objects, even such basic expressions as words for the perceptible characteristics of molar objects, locate these objects in a space of implications, that they describe at all, rather than merely label.⁴

The reason 'grivey' is merely a *label*, that it classifies without informing, is that nothing *follows* from so classifying an object. If I discover that all the boxes in the attic I am charged with cleaning out have been labeled with red, yellow, or green stickers, all I learn is that those labeled with the same color share *some* property. To learn what they *mean* is to learn, for instance, that the owner put a red label on boxes to be discarded, green on those to be retained, and yellow on those that needed further sorting and decision. Once I know what *follows* from affixing one rather than another label, I can understand them not as *mere* labels, but as *descriptions* of the boxes to which they are applied. Description is classification with *consequences*, either immediately practical ("to be discarded/examined/kept") or for further classifications.

Michael Dummett argues generally that to be understood as conceptually contentful, expressions must have not only *circumstances* of appropriate application, but also appropriate *consequences* of application.⁵ That is, one must look not only *upstream*, to the circumstances (inferential and non-inferential) in which it is appropriate to apply the expression, but also *downstream* to the consequences (inferential and non-inferential) of doing so, in order to grasp the content it expresses. One-sided theories of meaning, which seize on one aspect to the exclusion of the other, are bound to be defective, for they omit aspects of the use that are essential to meaning. For instance, expressions can have the same circumstances of application, and different consequences of application. When they do, they will have different descriptive content.

² Dretske, Fred: *Knowledge and the Flow of Information* (MIT Press—Bradford, 1981), Fodor, Jerry: *A Theory of Content* (MIT Press—Bradford, 1990).

³ I sketch this program in the opening section of "The Pragmatist Enlightenment (and its Problematic Semantics)" *European Journal of Philosophy*, Vol 12 No 1, April 2004, pp. 1-16.

⁴ Pp. 306-307 (§107) in: Wilfrid Sellars: "Counterfactuals, Dispositions, and Causal Modalities" In *Minnesota Studies in the Philosophy of Science, Volume II: Concepts, Theories, and the Mind-Body Problem*, ed. Herbert Feigl, Michael Scriven, and Grover Maxwell (Minneapolis: University of Minnesota Press, 1958), p.225-308.

⁵ I discuss this view of Dummett's (from his *Frege: Philosophy of Language* second edition [Harvard University Press 1993], originally published in 1974), at greater length in Chapter Two of *Making It Explicit* [Harvard University Press, 1994], and Chapter One of *Articulating Reasons* [Harvard University Press, 2000].

1] I will write a book about Hegel,

and

2] I foresee that I will write a book about Hegel,

say different things about the world, describe it as being different ways. The first describes my future activity and accomplishment, the second my present aspiration. Yet the circumstances under which it is appropriate or warranted to assert them—the situations to which I ought reliably to respond by endorsing them—are the same (or at least, can be made so by light regimentation of a prediction-expressing use of ‘foresee’). Here, to say that they have different descriptive content can be put by saying that they have different truth conditions. (That they have the same assertibility conditions just shows how assertibility theories of meaning, as one-sided in Dummett’s sense, go wrong.) But that same fact shows up in the different positions they occupy in the “space of implications.” For from the former it follows that I will not be immediately struck by lightning, that I will write some book, and, indeed, that I will write a book about Hegel. None of these is in the same sense a consequence of the second claim.

We might train a parrot reliably to respond differentially to the visible presence of red things by squawking “That’s red.” It would not yet be *describing* things as red, would not be applying the concept red to them, because the noise it makes has no significance for it. It does not know that it follows from something’s being red that it is colored, that it cannot be wholly green, and so on. Ignorant as it is of those inferential consequences, the parrot does not grasp the concept (any more than we express a concept by ‘grivey’). The lesson is that even observational concepts, whose principal circumstances of appropriate application are non-inferential (a matter of reliable dispositions to respond differentially to non-linguistic stimuli) must have inferential consequences in order to make possible description, as opposed to the sort of classification effected by non-conceptual labels.

The rationalist idea that the inferential significance of a state or expression is essential to its *conceptual* contentfulness is one of the central insights of Frege’s 1879 *Begriffsschrift* (“concept writing”)—the founding document of modern logic and semantics—and is appealed to by him in the opening paragraphs to define his topic:

...there are two ways in which the content of two judgments may differ; it may, or it may not, be the case that all inferences that can be drawn from the first judgment when combined with certain other ones can always also be drawn from the second when combined with the same other judgments...I call that part of the content that is the same in both the conceptual content [begriffliche Inhalt].⁶

Here, then, is the first lesson that analytic philosophy ought to have taught cognitive science: there is a fundamental conceptual distinction between classification in the sense of *labeling* and classification in the sense of *describing*, and it consists in the *inferential* consequences of the classification: its capacity to serve as a premise in inferences (practical or theoretical) to further conclusions. (Indeed, there are descriptive concepts that are purely *theoretical*—such as gene and quark—in the sense that in addition to their inferential consequences of application, they have *only* inferential *circumstances* of application.) There is probably no point in fighting over the minimal circumstances of application of the concepts concept and conceptual. Those who wish to lower the bar sufficiently are welcome to consider purely classificatory labels as a kind of concept (perhaps so as not to be beastly to the beasts, or disqualify human infants, bits of our brains, or even some relatively complex computer programs wholly from engaging in conceptually articulated activities). But *if* they do so, they must *not* combine those circumstances of application with the consequences of application appropriate to genuinely *descriptive* concepts—those that *do* come with inferential significances downstream from their application.

Notice that this distinction between labeling and describing is untouched by two sorts of elaborations of the notion of labeling that have often been taken to be of great significance in thinking about concepts from the classical classificatory point of view. One does not cross the boundary from labeling to describing just because the reliable capacity to respond differentially is *learned*, and in that sense flexible, rather than

⁶ Frege, *Begriffsschrift* (hereafter *BGS*), section 3. The passage continues: “In my formalized language [Begriffsschrift]...only that part of judgments which affects the possible inferences is taken into consideration. Whatever is needed for a correct inference is fully expressed; what is not needed is...not.”

innate, and in that sense rigid. And one is likewise developing the classical model in an orthogonal direction insofar as one focuses on the metacapacity to learn to distinguish arbitrary Boolean combinations of microfeatures one can already reliably discriminate. From the point of view of the distinction between labeling and describing, that is not yet the capacity to form *concepts*, but only the mastery of *compound* labels. That sort of structural articulation upstream has no *semantic* import at the level of description until and unless it is accorded a corresponding inferential significance downstream.

III. Ingredient vs. Free-Standing Content: Semantically Separating Content from Force

Once our attention has been directed at the significance of applying a classifying concept—downstream, at the consequences of applying it, rather than just upstream, at the repeatable it discriminates, the grouping it institutes—so that *mere* classification is properly distinguished from *descriptive* classification, the necessity of distinguishing different *kinds* of consequence becomes apparent. One distinction in the vicinity, which has already been mentioned in passing, is that between *practical* and *theoretical* (or, better, *cognitive*) consequences of application of a concept. The significance of classifying an object by responding to it one way rather than another may be to make it appropriate to *do* something else with or to it—to keep it, examine it, or throw it away, to flee or pursue or consume it, for example. This is still a matter of inference; in this case, it is *practical* inferences that are at issue. But an initial classification may also contribute to further classifications: that what is in my hand falls under both the classifications raspberry and red makes it appropriate to classify it also as ripe—which in turn has practical consequences of application (such as, under the right circumstances “falling to without further ado and eating it up,” as Hegel says in another connection) that neither of the other classifications has individually. Important as the distinction between practical and cognitive inferential consequences is, in the present context there is reason to emphasize a different one.

Discursive intentional phenomena (and their associated concepts), such as assertion, inference, judgment, experience, representation, perception, action, endorsement, and imagination typically involve what Sellars calls “the notorious ‘ing’/‘ed’ ambiguity.” For under these headings we may be talking about the *act* of asserting, inferring, judging, experiencing, representing, perceiving, doing, endorsing, and imagining, or we may be talking about the *content* that is asserted, inferred, judged, experienced, represented, perceived, done, endorsed, or imagined. ‘Description’ is one of these ambiguous terms (as is ‘classification’). We ought to be aware of the distinction between the act of describing (or classifying), applying a concept, on the one hand, and the content of the description (classification, concept)—*how* things are described (classified, conceived)—on the other. And the distinction is not merely of theoretical importance for those of us thinking systematically about concept use. A distinctive level of conceptual sophistication is achieved by concept users that themselves distinguish between the contents of their concepts and their activity of applying them. So one thing we might want to know about a system being studied, a non-human animal, a prelinguistic human, an artifact we are building, is whether *it* distinguishes between the *concept* it applies and what it *does* by applying it.

We can see a basic version of the distinction between semantic content and pragmatic force as in play wherever *different* kinds of practical significance can be invested in the *same* descriptive content (different sorts of speech act or mental act performed using that content). Thus if a creature can not only say or think that the door is shut, but also ask or wonder whether the door is shut, or order or request that it be shut, we can see it as distinguishing in practice between the content being expressed and the pragmatic force being attached to it. In effect, it can use descriptive contents to do more than merely describe. But this sort of practical distinguishing of pragmatic from semantic components matters for the semantic hierarchy I am describing only when it is incorporated or reflected in the *concepts* (that is, the *contents*) a creature can deploy. The capacity to attach different sorts of pragmatic force to the same semantic content is not sufficient for *this* advance in structural semantic complexity. (Whether it is a necessary condition is a question I will not address—though I am inclined to think that in principle the answer is ‘No’.)

For the inferential consequences of applying a classificatory concept, when doing that is describing and not merely labeling, can be either *semantic* consequences, which turn on the *content* of the concept being applied, or *pragmatic* consequences, which turn on the *act* one is performing in applying it. Suppose John issues an observation report: “The traffic light is red.” You may infer that it is operating and illuminated,

and that traffic ought to stop in the direction it governs. You may also infer that John has a visually unobstructed line of sight to the light, notices what color it is, and believes that it is red. Unlike the former inferences, these are not inferences from what John *said*, from the *content* of his utterance, from the concepts he has applied. They are inferences from his *saying* it, from the pragmatic force or significance of his *uttering* it, from the fact of his *applying* those concepts. For what he has *said*, that the traffic light is red, could be true even if John had not been in a position to notice it or form any beliefs about it. Nothing about John follows just from the color of the traffic light.⁷

It can be controversial whether a particular consequence follows from how something is described or from describing it that way, that is, whether that consequence is part of the descriptive content of an expression, the concept applied, or stems rather from the force of using the expression, from applying the concept. A famous example is expressivist theories of evaluative terms such as ‘good’. In their most extreme form, they claim that these terms have no descriptive content. *All* their consequences stem from what one is doing in using them: commending, endorsing, or approving. In his lapidary article “Ascriptivism,”⁸ Peter Geach asks what the rules governing this move are. He offers the archaic term ‘macarize’, meaning to characterize someone as happy. Should we say that in apparently describing someone as happy we are not really describing anyone, but rather performing the distinctive speech act of macarizing? But why not then discern distinctive speech acts for *any* apparently descriptive term?

What is wanted is a criterion for distinguishing semantic from pragmatic consequences, those that stem from the content of the concept being applied from those that stem from what we are doing in applying that concept (using an expression to perform a speech act). Geach finds one in Frege, who in turn was developing a point made already by Kant.⁹ The logical tradition Kant inherited was built around the classificatory theory of consciousness we began by considering. Judgment was understood as classification or predication: paradigmatically, *of* something particular *as* something general. But we have put ourselves in a position to ask: is this intended as a model of judgeable contents are constructed, or of what one is doing in judging? Kant saw, as Frege would see after him, that the phenomenon of *compound* judgments shows that it *cannot* play both roles. For consider the hypothetical or conditional judgment

3] If Frege is correct, then conceptual content depends on inferential consequences.

In asserting this sentence (endorsing its content), have I predicated correctness of Frege (classified him as correct)? Have I described him as correct? Have I applied the concept of correctness? If so, then predicating or classifying (or describing) is not judging. For in asserting the conditional I have *not* judged or asserted that Frege is correct. I have at most built up a judgeable content, the antecedent of the conditional, by predication. For embedding a declarative descriptive sentence as an unasserted component in a compound asserted sentence strips off the pragmatic force its free-standing, unembedded occurrence would otherwise have had. It now contributes *only* its *content* to the *content* of the compound sentence, to which alone the pragmatic force of a speech act is attached.

This means that embedding simpler sentences as components of compound sentences—paradigmatically, embedding them as antecedents of conditionals—is the way to discriminate consequences that derive from the *content* of a sentence from consequences that derive from the *act* of asserting or endorsing it. We can tell that ‘happy’ *does* express descriptive content, and is *not* simply an indicator that some utterance has the pragmatic force or significance of macarizing, because we *can* say things like:

4] If she is happy, then John should be glad.

⁷ One might think that a similar distinction could be made concerning a parrot that merely reliably responsively discriminated red things by squawking “That’s red.” For when he does that, one might infer that there was something red there (since he is reliable), and one might also infer that the light was good and his line of sight unobstructed. So both sorts of inference seem possible in this case. But it would be a mistake to describe the situation in these terms. The squawk is a label, not a description. We infer from the parrot’s producing it that there is something red, because the two sorts of events are reliably correlated, just as we would from the activation of a photocell tuned to detect the right electromagnetic frequencies. By contrast, John offers *testimony*. What he says is usable as a premise in our own inferences, not just the fact that his saying it is reliably correlated with the situation he (but not the parrot) reports (though they both respond to it).

⁸ *The Philosophical Review*, Vol. 69, No. 2, 221-225. Apr., 1960.

⁹ I discuss this point further in the first lecture of “Animating Ideas of Idealism” [op.cit.].

For in asserting that, one does *not* macarize anyone. So the consequence, that John should be glad, must be due to the descriptive content of the antecedent, not to its force.

Similarly, Geach argues that the fact that we can say things like:

5] If being trustworthy is good, then you have reason to be trustworthy,

shows that ‘good’ *does* have descriptive content.¹⁰ Notice that this same test appropriately discriminates the different descriptive contents of the claims:

6] Labeling is not describing,

and

7] I believe that labeling is not describing.

For the two do not behave the same way as antecedents of conditionals. The stuttering inference

8] If labeling is not describing, then labeling is not describing,

is as solid an inference as one could ask for. The corresponding conditional

9] If I believe that labeling is not describing, then labeling is not describing,

requires a good deal more faith to endorse. And in the same way, the embedding test distinguishes [1] and [2] above. In each case it tells us, properly, that different descriptive contents are involved.

What all this means is that any user of descriptive concepts who can also form compound sentences, paradigmatically conditionals, is in a position to distinguish what pertains to the semantic *content* of those descriptive concepts from what pertains to the *act* or pragmatic *force* of describing by applying those concepts. This capacity is a new, higher, more sophisticated level of concept use. It can be achieved *only* by looking at compound sentences in which other descriptive sentences can occur as unasserted components. For instance, it is only in such a context that one can distinguish *denial* (a kind of speech act or attitude) from *negation* (a kind of content). One who asserts [6] has *both* denied that labeling is describing, *and* negated a description. But one who asserts conditionals such as [8] and [9] has negated descriptions, but has *not* denied anything.

The modern philosophical tradition up to Frege took it for granted that there was an special attitude one could adopt towards a descriptive conceptual content, a kind of minimal force one could invest it with, that must be possible independently of and antecedent to being able to endorse that content in a judgment. This is the attitude of merely *entertaining* the description. The picture (for instance, in Descartes) was that *first* one entertained descriptive thoughts (judgeables), and *then*, by an in-principle subsequent act of will, accepted or rejected it. Frege rejects this picture. The principal—and in principle fundamental—pragmatic attitude (and hence speech act) is judging or endorsing.¹¹ The capacity merely to entertain a proposition (judgeable content, description) is a late-coming capacity—one that is parasitic on the capacity to endorse such contents. In fact, for Frege, the capacity to entertain (without endorsement) the proposition that *p* is just the capacity to endorse *conditionals* in which that proposition occurs as antecedent or consequent. For that is to explore its descriptive content, its inferential circumstances and consequences of application, what it follows from and what follows from it, what would make it true and what would be true if it were true, without endorsing it. This is a new kind of distanced attitude toward one’s concepts and their contents—one that becomes possible only in virtue of the capacity to form compound sentences of the kind of which

¹⁰ Of course, contemporary expressivists such as Gibbard and Blackburn (who are distinguished from emotivist predecessors such as C.L. Stevenson precisely by their appreciation of the force of the Frege-Geach argument) argue that it need not follow that the right way to understand that descriptive content is not by tracing it back to the attitudes of endorsement or approval that *are* expressed by the use of the expression in free-standing, unembedded assertions.

¹¹ In the first essay of “Animating Ideas of Idealism” [op.cit.] I discuss the line of thought that led Kant to give pride of place to judgment and judging.

conditionals are the paradigm. It is a new level of cognitive achievement—not in the sense of a new kind of empirical knowledge (though conditionals can indeed codify new empirical discoveries), but of a new kind of semantic self-consciousness.

Conditionals make possible a new sort of hypothetical thought. (Supposing that postulating a distinct attitude of supposing would enable one to do *this* work, the work of conditionals, would be making the same mistake as thinking that denial can do the work of negation.) Descriptive concepts bring empirical properties into view. Embedding those concepts in conditionals brings the contents of those concepts into view. Creatures that can do that are functioning at a higher cognitive and conceptual level than those who can only apply descriptive concepts, just as those who can do that are functioning at a higher cognitive and conceptual level than those who can only classify things by reliable responsive discrimination (that is, labeling). That fact sets a question for the different branches of cognitive science I mentioned in my introduction. Can chimps, or African grey parrots, or other non-human animals not just use concepts to describe things, but also semantically discriminate the contents of those concepts from the force of applying them, by using them not just in describing, but in conditionals, in which their contents are merely entertained and explored? At what age, and along with what other capacities, do human children learn to do so? What is required for a computer to demonstrate this level of cognitive functioning?

Conditionals are special, because they make *inferences* explicit—that is, put them into endorsable, judgeable, assertible, which is to say propositional form. And it is their role in inferences, we saw, that distinguishes descriptive concepts from mere classifying labels. But conditionals are an instance of a more general phenomenon. For we can think of them as operators, which apply to sentences to yield further sentences. As such, they bring into view a new notion of conceptual content: a new principle of assimilation, hence classification, of such contents. For we begin with the idea of sameness of content that derives from sameness of pragmatic force, attitude, or speech act. But the Frege-Geach argument shows that we can also individuate conceptual contents more finely, not just in terms of their role in free-standing utterances, but also accordingly as substituting one for another as arguments of operators (paradigmatically the conditional) does or does not yield compound sentences with the same free-standing pragmatic significance or force. Dummett calls these notions “free-standing” and “ingredient” content (or sense), respectively. Thus we might think that

10] It is nice here,

and

11] It is nice where I am,

express the same attitude, perform the same speech act, have the same pragmatic force or significance. They not only have the same circumstances of application, but the same consequences of application (and hence role as antecedents of conditionals). But we can see that they have different *ingredient* contents by seeing that they behave differently as arguments when we apply another operator to them. To use an example of Dummett’s,

12] It is *always* nice here,

and

13] It is *always* nice where I am,

have very different circumstances and consequences of application, different pragmatic significances, and *do* behave differently as the antecedents of conditionals. But this difference in content, this sense of “different content” in which they patently do have different contents, is one that shows up *only* in the context of compounding operators, which apply to sentences and yield further sentences. The capacity to deploy such operators to form new conceptual (descriptive) contents from old ones accordingly ushers in a new level of cognitive and conceptual functioning.

Creatures that can not merely label, but describe are *rational*, in the minimal sense that they are able to treat one classification as providing a *reason* for or against another. If they can use conditionals, they can distinguish inferences that depend on the *content* of the concept they are applying from those that depend

on what they are *doing* in classifying something as falling under that concept. But the capacity to use conditionals gives them more than just that ability. For conditionals let them *say* what is a reason for what, *say that* an inference is a good one. And for anyone who can do that, the capacity not just to *deny* that a classification is appropriate, but to use a *negation* operator to form new classificatory contents means brings with it the capacity to say that two classifications (classifiers, concepts) are incompatible: that one provides a reason to withhold the other. Creatures that can use this sort of sentential compounding operator are not just *rational*, but *logical* creatures. They are capable of a distinctive kind of *conceptual self-consciousness*. For they can describe the rational relations that make their classifications into descriptions in the first place, hence be conscious or aware of them in the sense in which descriptive concepts allow them to be aware of empirical features of their world.

IV. Simple versus Complex Predicates

There is still a higher level of structural complexity of concepts and concept use. I have claimed that Frege should be credited with appreciating both of the points I have made so far: that descriptive conceptual classification beyond mere discriminative labeling depends on the inferential significance of the concepts, and that semantically distinguishing the inferential significance of the contents of concepts from that of the force of applying them depends on forming sentential compounds (paradigmatically conditionals) in which other sentences appear as components. In each of these insights Frege had predecessors. Leibniz (in his *New Essay on the Human Understanding*) had already argued the first point, against Locke. (The move from thinking of concepts exclusively as reliably differentially elicited labels to thinking of them as having to stand in the sort of inferential relations to one another necessary for them to have genuine descriptive content is characteristic of the advance from empiricism to rationalism.) And Kant, we have seen, appreciated how attention to compound sentences (including “hypotheticals”) requires substantially amending the traditional classificatory theory of conceptual consciousness. The final distinction I will discuss, that between *simple* and *complex* predicates, and the corresponding kinds of concepts they express, is Frege’s alone. No-one before him (and embarrassingly few even of his admirers after him) grasped this idea.

Frege’s most famous achievement is transforming traditional logic by giving us a systematic way to express and control the inferential roles of *quantificationally complex* sentences. Frege could, as the whole logical tradition from Aristotle down to his time (fixated as it was on syllogisms) could not, handle iterated quantifiers. So he could, for instance, explain why

14] If someone is loved by everyone, then everyone loves someone,

is true (a conditional that codifies a correct inference), but

15] If everyone loves someone, then someone is loved by everyone,

is not. What is less appreciated is that in order to specify the inferences involving arbitrarily nested quantifiers (‘some’ and ‘every’), he needed to introduce a new kind of predicate, and hence discern a structurally new kind of *concept*.

Our first grip on the notion of a predicate is as a *component* of sentences. In artificial languages we combine, for instance, a two-place predicate ‘P’ with two individual constants ‘a’ and ‘b’ to form the sentence ‘Pab’. Logically minded philosophers of language use this model to think about the corresponding sentences of natural languages, understanding

16] Kant admired Rousseau,

as formed by applying the two-place predicate ‘admired’ to the singular terms ‘Kant’ and ‘Rousseau’. The kind of inferences that are made explicit by *quantified conditionals*—inferences that essentially depend on the contents of the predicates involved—though, require us also to distinguish a one-place predicate, related to but distinct from this two-place one, that is exhibited by

17] Rousseau admired Rousseau,

and

18] Kant admired Kant,

but *not* by [16].

19] Someone admired himself,

that is, something of the form $\exists x[Pxx]$, follows from [17] and [18], but not from [16]. The property of being a self-admirer differs from that of being an admirer and from that of being admired (even though it entails both).

But there is no *part* of the sentences [17] and [18] that they share with each other that they don't share also with [16]. Looking just at the sub-sentential expressions out of which the sentences are built does not reveal the respect of similarity that distinguishes self-admiration from admiration in general—a respect of similarity that is crucial to understanding why the conditional

20] If someone admires himself then someone admires someone,

$(\exists x[Pxx] \rightarrow \exists x\exists y[Pxy])$ expresses a good inference, while

21] If someone admires someone then someone admires himself,

$(\exists x\exists y[Pxy] \rightarrow \exists x[Pxx])$ does not. For what [17] and [18] share that distinguishes them from [16] is not a *component*, but a *pattern*. More specifically, it is a pattern of cross-identification of the singular terms that two-place predicate applies to.

The repeatable expression-kind 'admires' is a *simple* predicate. It occurs as a component in sentences built up by concatenating it appropriately with a pair of singular terms. 'x admires x' is a *complex* predicate.¹² A number of different complex predicates are associated with any multi-place simple predicate. So the three-place simple predicate used to form the sentence

22] John enjoys music recorded by Mark and books recommended by Bob,

generates not only a three-place complex predicate of the form $Rxyz$, but also two-place complex predicates of the form $Rxxy$, $Rxyx$, and $Rxyy$, as well as the one-place complex predicate $Rxxx$. The complex predicates can be thought of as patterns that can be exhibited by sentences formed using the simple predicate, or as equivalence classes of such sentences. Thus the complex self-admiration predicate can be thought of either as the *pattern*, rather than the *part*, that is common to all the sentences {"Rousseau admired Rousseau," "Kant admired Kant," "Caesar admired Caesar," "Brutus admired Brutus," "Napoleon admired Napoleon,"...}, or just as that set itself. Any member of such an equivalence class of sentences sharing a complex predicate can be turned into any other by a sequence of *substitutions* of all occurrences of one singular term by occurrences of another.

Substitution is a kind of *decomposition* of sentences (including compound ones formed using sentential operators such as conditionals). After sentences have been built up using simple components (singular terms, simple predicates, sentential operators), they can be assembled into equivalence classes (patterns can be discerned among them) by regarding some of the elements as systematically replaceable by others. This is the same procedure of noting invariance under substitution that we saw applies to the notion of free-standing content to give rise to that of ingredient content, when the operators apply only to whole sentences. Frege called what is invariant under substitution of some sentential components for others a '*function*'. A function can be applied to some arguments to yield a value, but it is not a *part* of the value it yields. (One can apply the function *capital of* to Sweden to yield the value Stockholm, but neither Sweden nor *capital of* is part of Stockholm.) He tied himself in some metaphysical knots trying to find a clear way

¹² This point, and the terminology of 'simple' and 'complex' predicates, is due to Dummett, in the second chapter of his monumental *Frege's Philosophy of Language* [op.cit.].

of contrasting functions with *things* (objects). But two points emerge clearly. First, discerning the substitutional relations among different sentences sharing the same simple predicate is crucial for characterizing a wide range of inferential patterns. Second, those inferential patterns articulate the contents of a whole new class of concepts.

Sentential compounding already provided the means to build new concepts out of old ones. The Boolean connectives—conjunction, disjunction, negation, and the conditional definable in terms of them ($A \rightarrow B$ if and only if $\sim(A \& \sim B)$)—permit the combination of predicates in all the ways representable by Venn diagrams, corresponding to the intersection, union, complementation, and inclusion of sets (concept extensions, represented by regions), and so the expression of new concepts formed from old ones by these operations. But there is a crucial class of new concepts formable from the old ones that are *not* generable by such procedures. One cannot, for instance, form the concept of a C such that for every A there is a B that stands to that C in the relation R. This is the complex one-place predicate logicians would represent as having the form $\{x: Cx \& \forall y \in A \exists z \in B [Rxz]\}$. As Frege says, such a concept cannot, as the Boolean ones can, be formed simply by putting together pieces of the boundaries of the concepts A, B, and C. The correlations of elements of these sets that concepts like these, those expressed by complex predicates, depend on, and so the inferences they are involved in, cannot be represented in Venn diagrams.

Frege showed further that it is just concepts like these that even the simplest mathematics works with. The concept of a natural number is the concept of a set every element of which has a successor. That is, for every number, there is another related to it as a successor ($\forall x \exists y [\text{Successor}(x, y)]$). The decisive advance that Frege's new quantificational logic made over traditional logic is a *semantic, expressive* advance. His logical notation can, as the traditional logic could not, form *complex* predicates, and so both express a vitally important kind of concept, and logically codify the inferences that articulate its descriptive content.

Complex concepts can be thought of as formed by a four-stage process.

- First, put together simple predicates and singular terms, to form a set of sentences, say $\{Rab, Sbc, Tacd\}$.
- Then apply sentential compounding operators to form more complex sentences, say $\{Rab \rightarrow Sbc, Sbc \& Tacd\}$.
- Then substitute variables for some of the singular terms (individual constants), to form complex predicates, say $\{Rax \rightarrow Sxy, Sxy \& Tayz\}$.
- Finally, apply quantifiers to bind some of these variables, to form new complex predicates, for instance the one-place predicates (in y and z) $\{\exists x [Rax \rightarrow Sxy], \forall x \exists y [Sxy \& Tayz]\}$.

If one likes, this process can now be repeated, with the complex predicates just formed playing the role that simple predicates originally played at the first stage, yielding the new sentences $\{\exists x [Rax \rightarrow Sxd], \forall x \exists y [Sxy \& Taya]\}$. They can then be conjoined, and the individual constant *a* substituted for to yield the further one-place complex predicate (in z) $\exists x [Rzx \rightarrow Sxd] \& \forall x \exists y [Sxy \& Tzyz]$. We can use these procedures to build to the sky, repeating these stages of concept construction as often as we like. Frege's rules tell us how to compute the inferential roles of the concepts formed at each stage, on the basis of the inferential roles of the raw materials, and the operations applied at that stage. This is the heaven of concept formation he opened up for us.

V. Conclusion

The result of all these considerations, which have been in play since the dawn of analytic philosophy, well over a century ago, is a four-stage *semantic* hierarchy of ever more demanding senses of “concept” and “concept use.” At the bottom are concepts as reliably differentially applied, possibly learned, *labels* or classifications. Crudely behaviorist psychological theories (such as B. F. Skinner's) attempted to do all their explanatory work with responsive discriminations of this sort. At the next level, concepts as *descriptions* emerge when merely classifying concepts come to stand in *inferential, evidential, justificatory* relations to one another—when the propriety of one sort of classification has the practical significance of making others appropriate or inappropriate, in the sense of serving as *reasons* for them. Concepts of this sort may still all have observational uses, even though they are distinguished from labels by also having

inferential ones.¹³ Already at this level, the possibility exists of empirical descriptive concepts that can *only* be properly applied as the result of inferences from the applicability of others. These are *theoretical* concepts: a particularly sophisticated species of the genus of descriptive concepts.

At this second level, conceptual content first takes a distinctive *propositional* form; applications of this sort of concept are accordingly appropriately expressed using *declarative sentences*. For the propositional contents such sentences express just are whatever can play the role of premise and conclusion in *inferences*. And it is precisely being able to play those roles that distinguishes applications of descriptive concepts from applications of merely classificatory ones. Building on the capacity to use inferentially articulated descriptive concepts to make propositionally contentful judgments or claims, the capacity to form sentential *compounds*—paradigmatically *conditionals*, which make endorsements of material inferences relating descriptive concept applications propositionally explicit, and *negations*, which make endorsements of material incompatibilities relating descriptive concept applications propositionally explicit—brings with it the capacity to deploy a further, more sophisticated, kind of conceptual content: *ingredient* (as opposed to free-standing) content. Conceptual content of this sort is to be understood in terms of the contribution it makes to the content of *compound* judgments in which it occurs, and only thereby, indirectly, to the force or pragmatic significance of endorsing that content.

Ingredient conceptual content, then, is what can be *negated*, or *conditionalized*. The distinctive sort of definiteness and determinateness characteristic of this sort of conceptual content becomes vivid when it is contrasted with contents that cannot appear in such sentential compounds. My young son once complained about a park sign consisting of the silhouette of what looked like a Scottish terrier, surrounded by a red circle, with a slash through it. Familiar with the force of prohibition associated with signs of this general form, he wanted to know: “Does this mean ‘No Scotties allowed’? Or ‘No dogs allowed’? Or ‘No animals allowed’? Or ‘No pets allowed’?” Indeed. A creature that can understand a claim like “If the red light is on, then there is a biscuit in the drawer,” without disagreeing when the light is not on, or immediately looking for the biscuit regardless of how it is with the light, has learned to distinguish between the content of descriptive concepts and the force of applying them, and as a result can entertain and explore those concepts and their connections with each other without necessarily applying them in the sense of endorsing their applicability to anything present. The capacity in this way to free oneself from the bonds of the here-and-now is a distinctive kind of conceptual achievement.

The first step was from merely *discriminating* classification to *rational* classification (‘rational’ because inferentially articulated, according to which classifications provide reasons for others). The second step is to *synthetic logical* concept formation, in which concepts are formed by logical compounding operators, paradigmatically conditionals and negation. The final step is to *analytical* concept formation, in which the sentential compounds formed at the third stage are *decomposed* by noting invariants under substitution. This is actually the same method that gave us the notion of ingredient content at the third stage of concept formation. For that metaconcept arises when we realize that two sentences that have the same pragmatic potential as free-standing, force-bearing rational classifications can nonetheless make different contributions to the content (and hence the force) of compound sentences in which they occur as unendorsed components—that is, when we notice that substituting one for the other may change the free-standing significance of asserting the compound sentence containing them. To form *complex* concepts, we must apply the same methodology to sub-sentential expressions, paradigmatically singular terms, that have multiple occurrences in those same logically *compound* sentences. Systematically assimilating sentences into various equivalence classes accordingly as they can be regarded as substitutional variants of one another is a distinctive kind of *analysis* of those compound sentences, as involving the application of concepts that were not *components* out of which they were originally constructed. Concepts formed by this sort of analysis are substantially and in principle more expressively powerful than those available at earlier stages in the hierarchy of conceptual complexity. (They are, for instance, indispensable for even the simplest mathematics.)

This hierarchy is not a *psychological* one, but a *logical* and *semantic* one. Concepts at the higher levels of complexity presuppose those at lower levels not because creatures of a certain kind cannot in practice, as a matter of fact, deploy the more complex kinds unless they can deploy the simpler ones, but because in principle it is impossible to do so. Nothing could count as grasping or deploying the kinds of concepts that

¹³ A key part of the higher *inferential* grade of conceptuality (which includes the former, but transforms it) is that it is *multipremise material* inferences that one learns to draw as conclusions (=responses) now to Boolean combinations of the relatively enduring states that result from one’s own responses.

populate the upper reaches of the hierarchy without also grasping or deploying those drawn from its lower levels. The dependencies involved are not empirical, but (meta)conceptual and normative. The Fregean considerations that enforce the distinctions between and sequential arrangement of concept-kinds do not arise from studying how concept-users actually work, but from investigation of what concept use fundamentally is. They concern not how the trick (of concept use) is done, but what counts as doing it—a normative, rather than an empirical issue. That is why it is philosophers who first came across this semantic hierarchical metaconceptual structure of concept-kinds.

But cognitive scientists need to know about it. For it is part of the job of the disciplines that cognitive science comprises to examine—each from its own distinctive point of view—all four grades of conceptual activity: the use of more complex and sophisticated kinds of concepts, no less than that of the simpler and less articulated sorts. The move from merely classificatory to genuinely descriptive concepts, for instance, marks a giant step forward in the phylogenetic development of sapience. I do not think we yet know what non-human creatures are capable of taking that step. Human children clearly do cross that boundary, but when, and by what means? Can non-human primates learn to use conditionals? Has anyone ever tried to teach them? The only reason to focus on that capacity, out of all the many linguistic constructions one might investigate empirically in this regard, is an appreciation of the kind of semantic self-consciousness about the rational relations among classifications (which marks the move from classification to rational description) that they make possible. Computer scientists have, to be sure, expended some significant effort in thinking about varieties of possible implementation of sentential compounding—for instance in exploring what connectionist or parallel distributed processing systems can do. But they have not in the same way appreciated the significance of the question of whether, to what extent, and how such “vehicleless” representational architectures can capture the full range of concepts expressed by complex predicates. (Their lack of syntactically compositional explicit symbolic representations prohibits the standard way of expressing these concepts, for that way proceeds precisely by substitutional *decomposition* of such explicit symbolic representations.) These are merely examples of potentially important questions raised by the hierarchy of conceptual complexity that cognitive scientists have by and large not been moved so much as to ask.

Why not? I think it is pretty clear that the answer is *ignorance*. Specifically, it is ignorance of the considerations, put forward already by Frege, that draw the bright metaconceptual lines between different grades of concepts, and arrange them in a strict presuppositional semantic hierarchy. Any adequately trained cognitive scientist—even those working in disciplines far removed from computational linguistics—can be presumed to have at least passing familiarity with the similarly four-membered Chomsky hierarchy that lines up kinds of grammar, automaton, and *syntactic* complexity of languages in an array from most basic (finite state automata computing regular languages specifiable by the simplest sort of grammatical rules) to most sophisticated (two-stack pushdown automata computing recursively enumerable language specifiable by unrestricted grammatical rules). But the at least equally significant *semantic* distinctions I have been retailing have not similarly become a part of the common wisdom and theoretical toolbox of cognitive science—even though they have been available for a half-century longer.

The cost of that ignorance, in questions not asked, theoretical constraints not appreciated, promising avenues of empirical research not pursued, is great. Failure to appreciate the distinctions and relations among fundamentally different kinds of concepts has led, I think, to a standing tendency systematically to overestimate the extent to which one has constructed (in AI) or discerned in development (whether by human children or non-human primates) or reverse-engineered (in psychology) what *we* users of the fanciest sorts of concepts do. That underlying ignorance is culpable. But it is not the cognitive scientists themselves who are culpable for their ignorance. The ideas in question are those that originally launched the whole enterprise of analytic philosophy. I think it is fair to say that as we philosophers have explored these ideas, we have gotten clearer about them in many respects. For one reason or another, though, we have not shared the insights we have achieved. We are culpable for having kept this treasure trove to ourselves. It is high time to be more generous in sharing these ideas.

Two Aspects of Content: Semantic Inferentialist's Reconciliation of Authoritative Self-Knowledge and Content Externalism

Shuhei, Shimamura

s0025@hotmail.com

Doctoral Course, Department of Philosophy,
Graduate School of Humanities and Sociology,
The University of Tokyo, Japan

Abstract: This paper aims to dissolve the problem of the compatibility between authoritative self-knowledge of propositional attitudes and content externalism. The basic strategy to do this is to appeal to pair concepts of semantic inferentialism, i.e., the propositional content and the representational content. Inferentialists can explain how the former content can be authoritatively known from the first-person perspective, and why the latter content is partly determined by external factors. In addition to this, it is also shown how this inferentialistic conception of content succeeds in dodging the attack on authoritative self-knowledge derived from the physical externalism about the mental content. What makes it possible is that the inferentialistic conception of content rejects the traditional but problematic presumption that the intension is the determinant of the extension. In this way, this paper reconciles the authority of self-knowledge of propositional attitudes with the thesis of (physical) externalism.

1 Introduction: The Problem of Compatibility

It is, on the one hand, often stressed that our language has a public character. Content externalism is one of the clear manifestations of the character. On the other hand, it is also difficult to deny that our self-knowledge of mental states (including propositional attitudes) has a certain kind of authority. Moreover, since the self-knowledge is normally achieved relying on no inference, the authority is based on no empirical evidence. Although both of these claims (the public character of our language and the non-inferential authority of our self-knowledge) seems quite plausible respectively, there is, at least apparently, a tension in simultaneously accepting both of them. This is what I call 'the problem of compatibility', which is the subject of this paper.

But what exactly is the problem? Although the controversy is complicated, the gist of the problem is easy to grasp.¹ Content externalism claims that the contents of linguistic expressions are determined partly by certain (environmental or social) factors external to an individual speaker's inner state. Once we endorse this thesis concerning the contents of linguistic expressions, then it becomes almost unavoidable to endorse the same thesis concerning the contents of propositional attitudes, since those two sorts of contents are arguably dependent on each other. But the external factors that

¹ The following formulation of the compatibility problem is first proposed by Boghossian (1989).

the thesis claims partly determine the contents of our statements and propositional attitudes can be completely unknown to ourselves. Consequently, the thesis of content externalism entails that the contents of linguistic expressions and propositional attitudes are in a sense beyond our own grasps. Now this gives rise to the problem of compatibility. For authoritative self-knowledge of propositional attitudes seems to demand the opposite. It seems that we can authoritatively know our own propositional attitudes without relying on any evidence or inference based on it. But if the contents of propositional attitudes are determined partly by certain external factors, of which we can be completely unaware, then the authoritative self-knowledge is threatened.

That seems to be a contradiction! But is that really contradictory? The crucial point on which we have to reflect is what the word ‘contents’ means here. To begin with, (1) it is not at all obvious that the contents of which we can have authoritative self-knowledge and the contents that are beyond our grasps are the same. Then, if they are different, (2) in what do the former and the latter contents consist respectively? And finally (3) how are the former contents authoritatively known to us through no evidence and no inference, and why are the latter contents beyond our own grasps?

2 Key Ideas from Semantic Inferentialism

In my opinion, the answers to all these questions can be given adequately from the semantic inferentialist’s standpoint. Therefore, briefly turning aside from the problem of compatibility, I introduce the main points of semantic inferentialism.²

Semantic inferentialism is concerned with the question of what it is for statements uttered or propositional attitudes held by us to be meaningful or to have contents. And it answers the question basically in the spirit of the use theory of meaning. Traditionally, philosophers have divided meaning into two distinctive categories—the intension / the extension, or in Brandom’s terminology the concomitant aspect / the representational aspect of the propositional content.³ Roughly speaking, many of the philosophers try to explain the former based on the latter. Semantic inferentialists, however, base on our normative practice of inference and invert the order of explanation.⁴

To begin with, according to inferentialists’ explanation, a statement one utters has some propositional content when it is *inferentially articulated*, i.e., when it counts as a move in the game of giving and asking for reasons. Then, the propositional content of a claim has two different aspects, which are the inferentialist’s alternatives of the traditional distinction of intension and extension. First, the *concomitant aspect* of the propositional content that the uttered statement has consists in the *inferential commitments*⁵ that the original statement, along with the background beliefs of *the utterer*,

² This view is rigorously and exhaustively developed in Brandom (1994), and we can see the outline of it in Brandom (2001).

³ Strictly speaking, the distinction of extension / intension, slightly but importantly, differs from the inferentialist’s version of their alternatives, the distinction of representational aspect / concomitant aspect. We will see the difference at length in section four below.

⁴ See Brandom (1994, pp. 93-4, pp. 135-6). This inversion is crucial for our discussion. See section four.

⁵ The adjective ‘inferential’ is contrasted with ‘substitutional.’ For details, see the following paragraphs and note seven.

rationally makes her prepared to undertake, and the *entitlements* that *the utterer* rationally regards as justifying the original statement.⁶ Thus, we understand the concomitant aspect of the propositional content of a statement (for short, I simply call this the ‘propositional content’ of a statement) in terms of *its distinctive roles in intrapersonal inference*.

Second, in order to understand the *representational aspect* of the propositional content of a statement (in short, ‘representational content’ of a statement), we should attend not merely to the inferentially articulated dimension, but also to the socially articulated dimension of our game of giving and asking for reasons. To begin with, the representational content of a claim, contrasted with the propositional content of the claim (i.e., what is *said* by the claim), is the object that the claim talks *about*. And what the object is like is determined by the identity condition of the object. Then the representational content of the claim consists in *the inferential roles of various true identity statements that describe the identity condition of the object*. But what inferential roles do these identity statements play? Their roles are enabling us to make new *substitutional commitments*⁷ through substitutional inferences, i.e., the inferences that draw a consequence by simultaneously replacing a certain term occurring in a premise with another term based on an identity statement. Therefore, grasping the representational content of a statement consists in the ability to derive various substitutional commitments from the original statement together with those true identity statements through making substitutional inferences. But for what purpose does such an ability serve? We employ this kind of ability in our social and communicative practice. Grasping the representational content of a claim is indispensable for assessing the claim made by another person from *de re* standpoint and for extracting useful information from the claim. In this sense, the representational content of a statement is not only inferentially but also *socially* articulated in our inferential practice.⁸

Finally, all the above explanations can be, *mutatis mutandis*, applied to the explanations of the contents of beliefs. Additionally, for the sake of argument, I presuppose that the contents of the other propositional attitudes than beliefs (e.g., desires and intentions) can be also explained in similar ways.

3 The Basic Line

⁶ See Brandom (1994, pp. 168-70, pp. 186-90). Strictly speaking, a third factor, i.e., the incompatibility relation between commitments and entitlements, is also needed in full explanation of inferential articulation of the propositional content. But, for reasons of space, we have to leave it out.

⁷ The distinction between the inferential commitment and the substitutional commitment consists in the difference between *de dicto* viewpoint and *de re* viewpoint. An inferential commitment of a claim is an inferential consequence of the claim together with the *speaker's* auxiliary background beliefs. So, the speaker herself is prepared to acknowledge the commitment from her own standpoint. Oppositely, a substitutional commitment of a claim is the commitment undertaken by *a interlocutor* from *de re* viewpoint. The substitutional commitment is undertaken through the substitutional inference based on an identity statement to which *the interlocutor* is committed, regardless that the speaker may acknowledge the commitment or not.

⁸ See Brandom (1994, pp. 136-40, pp. 370-6, pp. 495- 520).

Now we can come back to the series of questions concerning the compatibility problem. The basic line of my position is very simple. Let me answer briefly the questions (1) and (2) at a stroke. In my opinion, the two sorts of contents, i.e., the contents that one can authoritatively know without relying on any evidence and the contents that are in a sense beyond one's own grasp, are different. And my suggestion is that, from the semantic inferentialist's viewpoint, we should regard the former as 'propositional contents' and the latter as 'representational contents.'⁹

But how can we authoritatively know the propositional contents of our own attitudes based on no evidence? As we saw in section two, grasping the propositional content of one's own belief is grasping the inferential commitments and the entitlements connected to the belief. And the inferential commitments of a belief, on the one hand, are the statements that the belief, together with other auxiliary background beliefs of the believer, makes her prepared to acknowledge.¹⁰ On the other, the entitlements of the belief are the statements that the believer is prepared to acknowledge as justification of the belief. For our purpose, what is important in these characterizations of commitments and entitlements is that both of them contain the normative phrase of '(the believer is) prepared to acknowledge....' In virtue of this feature, we

⁹ But I have to note that this direction has been already suggested in Brandom (1994, p. 507). So what I try to do below is to flesh out his idea and thereby to apply it to an attempt of dissolving the compatibility problem.

¹⁰ Strictly speaking, there can be an inferential commitment that a subject is not prepared to acknowledge. Brandom pointed out that there are two ways of undertaking an inferential commitment. (Brandom 1994, pp. 193-8) One way is to directly acknowledge the commitment, and the other is to undertake the commitment as a consequence of the other commitments that she explicitly acknowledged. Let us call the former 'commitments as acknowledgement' and the latter 'commitments as consequence.' The latter (the commitment as consequence) poses an apparent problem to our claim that we have authoritative self-knowledge about the propositional contents of our own beliefs. For a commitment as consequence can be unknown to its subject since her capacity of inference may not be complete. Indeed, this kind of incompleteness of inferential capacity is a familiar and widely observed phenomenon. For example, consider a person who knows all the axiom of Euclidean geometry but who does not know the Pythagorean theorem. Making complex inferences often requires a special talent. Now, does this kind of unconscious commitments form the counterexamples to our claim?

Probably, this is true. But I think that it is no flaw of inferentialistic conception of mental content. For our ability of self-knowledge is *actually* limited to some extent. For instance, consider a person who explicitly acknowledges both that free wills exist and that Newtonian mechanics is true. Then, an incompatibilist philosopher appears and starts to persistently persuade her step by step. The persuasion succeeded, and she accepts and acknowledges all these steps, e.g., that the existence of free will entails the ability to do otherwise, that the ability requires the rejection of determinism, but that Newtonian mechanics do imply determinism etc.. Now she convinces that the original two beliefs are incompatible and may say, 'I did not completely grasp what my beliefs mean!'

In this sense, our authority of self-knowledge about propositional contents that we are going to explain is in fact not perfect. But by contrast with normal knowledge, there still remains something special in the self-knowledge about propositional contents. This is what we are going to explain. In the following argument, for simplicity, I will focus exclusively on the commitment as acknowledgement. But our argument can be also applied to the commitment as consequence, to the extent that we normally admit the authoritative self-knowledge about them.

can explain the self-knowledge at issue, appealing to so-called ‘Transparency Thesis.’ In the following, I will elucidate this point.

To begin with, let us turn to the explanation of our grasping the inferential commitments of our own beliefs. Suppose that B(P) is one of my beliefs, and Cs are the inferential commitments associated with B(P). As we saw in the above paragraph, a statement S is C if and only if B(P), together with other auxiliary background beliefs of mine (call them ‘B(Qs)’), makes me prepared to acknowledge S. However, in fact, *being prepared to acknowledge S* is equivalent to *believing* that S. So, the above biconditional can be put in the following way; a statement S is C if and only if B(P), together with B(Qs), makes me *believe* that S. In other words, A statement S is C given that B(P) and B(Qs), I believe that S. This means that in order to judge whether a given statement S is C or not, I only have to judge whether I believe that S or not given that B(P) and B(Qs).

Here, it is helpful to remember the ‘Transparency Thesis’ about self-ascription of beliefs, which is first proposed and named by Roy Edgley, and is afterward developed by Gareth Evans and Richard Moran.¹¹ According to the thesis, in order to judge whether I believe that S or not, all I have to do is judging whether it is the case that S or not. (The relevant question is directed not to myself, but to the world.) Why? For asserting the following sentence (what is called ‘Moore’s sentence’) sounds awkwardly irrational; ‘It is the case that S, but I do not believe that S.’ Note that the Transparency Thesis holds only in the case of *self*-ascription. (It is not at all irrational to assert that it is case that S, but *she* does not believe that S.) Now, let us apply this thesis to the question at issue. In order to judge whether I believe that S or not given that B(P) and B(Qs), all I have to do is judge whether it is the case that S or not, given that P and Qs. As long as I am equipped with minimal rationality and with the normal ability to make inference concerning the world, I can judge whether a given statement S is C or not, based on this kind of transparency procedure. Therefore, we can authoritatively know what inferential commitments our own beliefs have.

This type of explanation, *mutatis mutandis*, can be applied to the grasp of the entitlements of one’s own beliefs. In addition to the previous notation, suppose that Es are the entitlements associated with B(P). Then, a statement S is E if and only if I am prepared to acknowledge that E is a justification of B(P). But again, for the synonymy of ‘being prepared to acknowledge’ and ‘believe’, we can translate the above biconditional into the following form; a statement S is E iff I believe that S is a justification of B(P). Through the transparency procedure, the right side of this biconditional can be judged by judging the question of whether it is the case that S justifies P or not. Consequently, we, who have minimal rationality and the ability to assess evidence, can authoritatively know what entitlements our own beliefs have. Now, we obtain the explanation of the authoritative knowledge about the propositional contents (the inferential commitments and the entitlements) that our own beliefs have.

Next, why are the representational contents of one’s attitudes beyond one’s own grasp? According to the inferentialist’s view, grasping the representational content of a claim —grasping the object that the claim talks about—demands of us to grasp the substitutional commitments that the original claim plus the relevant *true* identity

¹¹ See Edgley (1969, p. 90), Evans (1982, p. 225-6), and Moran (1988, pp. 142-8; 2001, pp. 61-5).

statements make us undertake. Here, the adjective ‘true’ is essential. For instance, my belief that the Morning Star is very bright is, needless to say, directed to the Morning Star. Of course, I know that. But this belief can be also said to be directed to Venus since it is true that the Morning Star = Venus, whether or not I know it. In other words, my belief that the Morning Star is very bright conveys some information about Venus too. I, however, may not be able to utilize the information since the truth of the relevant identity statement is objective, and I have to make certain empirical inquiry to know whether it is true or not. In a word, *a posteriority* of true identity statements makes a lot of room for my not completely grasping the representational contents of my own beliefs.

These are the basic line of my argument. And they seem to suggest a way to reconcile authoritative self-knowledge with content externalism. For our picture can admit both that we can, in the propositional sense, authoritatively know the contents of our own propositional attitudes, and that the contents of our propositional attitudes are, in the representational sense, determined by certain external factors. In the following section, we will turn to investigate at length exactly what implications this line of thought has to content externalism.

4 Implications to the Physical Externalism¹²

4-1 Putnam’s Argument

Now, let us consider Putnam’s physical externalism (see Putnam (1975)). In order to establish the thesis, Putnam devised a science-fictional thought experiment of Twin Earth, which is a ‘Doppelganger’ of Earth except that the liquid that people call ‘water’ has the different micro-structure, XYZ, although as to the macro-level properties XYZ is indistinguishable from H₂O. Then, imagine a person who lived in 1750 on Twin Earth. Since the technique of chemical analysis had not been yet invented in 1750, she did not have any knowledge of the micro-structure of XYZ and cannot distinguish it from H₂O. Hence, in 1750 she had the exactly same set of beliefs about the substance that she called ‘water’ as her counter-part on Earth. This, however, entails that the meaning of the word ‘water’ is not determined by her mental states. For despite that she and her counter-part share the exactly same types of mental state, the reference of ‘water’ differs between Twin Earth and Earth. Consequently, Putnam says, “‘meanings’ just ain’t in the head!” (ibid., p. 227) and claims that it is partly determined by the physical environment where the speaker has been living.

In order to be more rigorous, the following two points are worth noting. The first is that when Putnam uses the word ‘mental state’ above, he uses it in the ‘narrow’ sense, i.e., that the mental states that he supposes are individualistic in the sense that they supervene just on subject’s inner states, not on the physical / social environment of

¹² Although the social externalism proposed by Burge (1979) is also important, for reasons of space, we have to omit the consideration of it. My prospect is that though Burge himself asserts that his externalism encroaches on the territory of propositional contents of our beliefs, inferentialists can reconcile Burge’s argument for social externalism with authoritative self-knowledge again. I will elaborate on this reconciliation some other time.

the subject. He calls this premise ‘methodological solipscism’ about the mental. And the second note is that when he says, “‘meanings’ just ain’t in the head!”, the word ‘meanings’ refers to *intensions* in the traditional sense. Traditionally, the intension of a word is supposed to be something that determines the extension (reference) of the word, although the explanations of exactly what intensions are diverge among philosophers. According to this terminology, the core of the above thought experiment can be reformulated as follows: the extension (reference) of the word ‘water’ is determined partly by the physical environment of a speaker; but every (narrow) mental state that the speaker holds is determined by her inner states, regardless of her physical environment; therefore, the word’s intension (the determinant of its extension) cannot be fit inside her (narrow) mental states.

From this argument, Putnam concludes that we have to reject either of the two familiar claims, i.e., (1) the claim that the speaker’s mental states determine the intension of a word, or (2) the claim that the intension of a word determines its extension. As is widely known, Putnam selected the first option of rejecting the claim (1). This means that the understanding of word’s meaning— connecting an intension with a word—is not a matter of the mental since the connection obtains partly by the causal link between a speaker and her environment. Thus, the physical externalism follows.

Although Putnam himself applies his argument only to the meanings of words, his argument, as McGinn (1977) correctly pointed out, can be also applied to their mental counter-parts, the contents of propositional attitudes. And this expansion will lead us into the compatibility problem of self-knowledge and physical externalism. Again, imagine that Emily, who lived on Twin Earth in 1750, held a belief that water is transparent. To what was her belief directed, or what did her belief represent? For the same reason as the above, the answer will be XYZ, not H₂O. But from the assumptions of the argument, she had the exactly same types of bodily state and narrow mental state as her counter-part on Earth, i.e., she had the brain states and the phenomenal states¹³ that were qualitatively identical to her counter-part’s. So what her belief represents is not fully determined by these inner states. It is partly determined by her physical environment. And if we suppose that, as parallel to the case of the intension, the content of a propositional attitude is the determinant of what the attitude represents, then the content also bulges out of these inner states, i.e., the content does not supervene on them. Such a content is generally called ‘wide content.’ Then, the problem of compatibility between self-knowledge of one’s own propositional attitudes and physical externalism occurs. If the contents of our propositional attitudes are wide, i.e., if they are partly determined by speaker’s physical environment, of which we can be completely ignorant, how can we have non-inferential authoritative self-knowledge of our own propositional attitudes?

4-2 Reconciling Putnam’s Argument with Self-Knowledge

Roughly speaking, our reply to this problem is rejecting claim (2) and sustaining claim (1), which is just the opposite of Putnam’s position. In his argument, the rejection of (1), which is equivalent to endorsing content externalism, depends on the

¹³ For the argument, here I presuppose that the phenomenal states are generally narrow. But this presupposition is not indispensable for my discussion.

rightness of (2). However, is claim (2) really essential for our concept of content (or intension)? Isn't there any other conception of content (or intension) that does not imply claim (2)? Here, the standpoint of semantic inferentialism is crucial. As I pointed out in section two, inferentialism reverses the traditional order of explanation between the extension and the intension. Traditionally, philosophers tend to explain what the intension is based on its relation to the extension, presupposing an independent explanation of what the extension is. And the view that the philosophers widely shared was that the intension is something giving the sufficient condition of the extension (although it is open exactly what the intension consists in). However, the inferentialist radically departs from this way of thinking. Inferentialist's alternatives of the intension and the extension, i.e., the propositional content and the representational content, are both explained in terms of our inferential practice. What is crucial for our discussion is that inferentialist's explanation of the two concepts does *not* imply the traditional claim (2). I want to elaborate on this point below.

According to inferentialism, on the one hand, grasping the propositional content of a statement (or a belief) is grasping the inferential role that the statement (or the belief) plays, namely the inferential commitments that the subject who asserts the statement (or holds the belief) undertakes and the entitlements that she regards as support for the statement (or the belief). On the other hand, grasping the representational content of a statement (or a belief) consists in grasping the substitutional commitments derived from the statement along with the relevant true identity statements. It follows from this explanation that grasping the propositional content of a statement (or a belief) does not necessarily contain fully grasping its representational content. For, as I pointed out in section three, the relevant true identity statements may be a posteriori.

For example, consider the statement,

The Morning Star is observed at peep of day. (a)

And imagine a rational person, Nancy, who asserts this statement with a good grasp of its *propositional content* (its entitlements and its inferential commitments): e.g., she grasps what perceptual statements she can use to support the statement, and she grasps that by asserting the statement she is committed to the statement that the Morning Star is a certain large substance floating in the space etc.. But, in this case, it is possible that she refrains from undertaking the commitment to the following statement,

The Evening Star is observed at peep of day. (b)

Since the following identity statement is undoubtedly a posteriori, she may be ignorant of it.

The Morning Star = the Evening Star. (i)

Nevertheless, we, who has already known that (i) is empirically true, know that her statement (a) is directed not only to the Morning Star, but also to the Evening Star. So we can derive the commitment to (b) from the commitment to (a) together with the

commitment to (i). In other words, the inferential role of (i), in this context, is enabling us to make substitutional inference from (a) to (b). Nancy, however, cannot make such a substitutional inference. Consequently, she does not grasp the substitutional commitment derived from the proper substitutional inference, and therefore, she does not fully grasp the representational content of her own statement (a).

To summarize, the inferentialist defines the propositional content of a statement in terms of the entitlements and the inferential commitments that a rational utterer connects to the statement. So, in uttering the statement, the rational utterer, as I argued in section three, always grasps the propositional content of her statement. On the contrary, grasping the representational content of the statement consists in grasping the proper substitutional commitments (like (b)) that can be obtained through proper substitutional inferences based on *true* identity statements (like (i)). But since the truth value of the identity statements is objective, and knowing the value is an empirical matter, it is possible that the rational utterer does not know which identity statement is true. Therefore, the representational content of the statement can be unknown to the utterer. In a word, grasping the representational content of a statement demands the utterer of empirical inquiry although grasping its propositional content does not.

If grasping the propositional content of a statement does not necessarily contain grasping its representational content, then it follows that its propositional content is not something providing the necessary and sufficient condition of its representational content. This is just the same as rejecting the above claim (2). Consequently, the inferentialist can dissolve the compatibility problem, in full agreement with Putnam on his externalist insight that Emily's belief that water is transparent is directed to XYZ, not to H₂O. Since the inferentialist rejects (2), accepting the authoritative self-knowledge about propositional contents of one's own belief is no bar to the belief's representational content being partly determined by her physical environment, of which she may be completely unaware. Indeed, the inferentialist, as we saw in the above paragraph, positively asserts that the representational content of a propositional attitude varies depending on her physical environment.¹⁴ For the inferentialist too, the representational content is an empirical matter.

Concluding Remarks

In my diagnosis, the problem of compatibility between self-knowledge and externalism arises, due to a bias traditionally shared in several conceptions of intension or mental content, i.e., the bias that the intension (or the mental content) somehow determines the extension. Inferentialism, however, rejects the bias, and thereby can reconcile the distinctive self-knowledge with the externalist's insight. This shows an important advantage of inferentialistic conception of meaning and mental content. For each of the distinctiveness of our self-knowledge about propositional attitudes and the public character of our language is not easy to throw away. The large quantity of papers concerning this topic shows the seriousness of this dilemma. But inferentialism dissolves the dilemma by offering two new conception of meaning (or mental con-

¹⁴ Brandom is much self-conscious of this externalistic character of his view. He focuses on this topic in Brandom (1994, pp. 631-3, pp. 645-649).

tent), i.e., the propositional content and the representational content. In other words, by offering them, inferentialism neatly accommodates both the subjective, authoritatively known aspect and the public, empirically known aspect of meaning (or mental content). In this sense, inferentialist's conception of meaning (or mental content) brings us a significant insight.

References

- Boghossian, P. (1989), 'Content and Self-Knowledge,' *Philosophical Topics*, 17, pp. 5-26.
- Brandom, R. B. (1994), *Making It Explicit*, Harvard University Press.
- Brandom, R. B. (2000), *Articulating Reasons*, Harvard University Press.
- Burge, T. (1979), 'Individualism and the Mental,' in *Studies in Metaphysics: Midwest Studies in Philosophy*, vol. 4, reprinted in his *Foundations of Mind*, Clarendon Press, 2007, pp. 100-50.
- Edgley, R. (1969), *Reason in Theory and Practice*, Hutchinson.
- Evans, G. (1982), *The Varieties of Reference*, ed. J. McDowell, Oxford University Press.
- McGinn, C. (1977), 'Charity, Interpretation and Belief', *Journal of Philosophy*, 74, pp. 521-35.
- Moran, R. (1988), 'Making up your mind: Self-Interpretation and Self-Constitution' *Ratio* 1, pp. 135-51.
- Moran, R. (2001), *Authority and Estrangement: An Essay on Self-Knowledge*, Princeton University Press.
- Putnam, H. (1975), 'The Meaning of "Meaning,"' in his *Mind, Language and Reality: Philosophical Papers*, Cambridge University Press, pp. 215-71.

Explicating as Distancing

Jeremy Wanderer

Logical locutions are characterised by Brandom as having two interlinked features. The first is that the abilities required to deploy such locutions are algorithmically elaborated from the abilities required to participate in an autonomous discursive practice. The second is that such locutions function to make explicit what is implicit in the performances and claims that feature in such an autonomous discursive practice.

Brandom endorses the following extension of this conception: it is possible for there to be an autonomous discursive practice that lacks logical locutions. This has been dubbed the 'layer-cake picture' of sapience, in the sense that, as in a layer-cake where each lower layer is able to stand independently of the one above it (but not the other way round), the layer of rationality in Brandom's model of sapience is able to stand independently of the logical layer (but not the other way round).

Many critics of Brandom have found this layer-cake picture unintelligible. According to these critics, the ability to make reasons explicit as reasons is required to be able to criticize them. Any social practice in which performers lack critical abilities, such as one lacking the logical locutions required to make reasons explicit, is not really a rational (discursive) practice at all.

This paper explores Brandom's uneasy relation to a (Kant-inspired) tradition that provides an obvious home for such a critique, a tradition for which the notion of rational freedom requires the ability to distance oneself from brute impulses. Brandom sometimes conveys the impression that his model of discursive practice should be treated as a development of the way of thinking about freedom captured by this tradition. The discussion here suggests instead that his model constitutes a significant departure from this tradition.

More specifically, it is argued that:

- (a) the notion of explication receives an alternative treatment when viewed within the trajectory of this tradition,
- (b) this treatment mandates the abandonment of the layer cake picture of sapience to which Brandom is committed, and that
- (c) abandoning this picture need not undermine the insight afforded by the conception of logical locutions as elaborated from, and explicative of, an autonomous discursive practice.

What emerges is an alternative understanding of explication as distancing, in which the abilities required to deploy rational and logical abilities are independent though interdependent aspects of our sapience. According to this, not every exercise of rationality need involve the deployment of distinctively logical locutions, but unless one has some logical capabilities that could be deployed, thereby allowing for critical self-reflection upon such an exercise, these exercises would not be instances of rationality.

Bibliography:

Brandom, R. (1979) Freedom and Constraint by Norms *American Philosophical Quarterly* 16: 187-196

- Brandom, R. (1994) Making It Explicit Harvard University Press.
- Brandom, R. (2005) Responses *Pragmatics and Cognition 13*: 227-249.
- Brandom, R. (2008) Between Saying and Doing Oxford University Press.
- Brandom, R. (2009) Reason In Philosophy: Animating Ideas Harvard University Press.
- Hampshire, S. (1965) Freedom of the Individual Princeton University Press.
- Korsgaard, C. (1996) The Sources of Normativity Cambridge University Press.
- McDowell, J. (1994) Mind and World Harvard University Press.
- McDowell, J. (2005) Motivating inferentialism: Comments on Making It Explicit (Ch. 2) *Pragmatics and Cognition 13*: 121–140
- McDowell, J. (ms.) *Five Lectures on Action* delivered at the University of Chicago, 2006.
- Macbeth D. (forthcoming) Inference, Meaning, and Truth in Brandom, Sellars, and Frege in: Wanderer and Weiss (eds.) Reading Brandom (Routledge).
- Moran, R. (2001) Authority and Estrangement Princeton University Press.
- Laurier, D. (2005) Pragmatics, Pittsburgh style *Pragmatics and Cognition 13*: 141–160.
- Wanderer, J. (2008) Robert Brandom (Acumen)

The Mind-Body Problem and Brandom's Analytic Pragmatism

François-Igor Pris

francois-igor.pris@uni-erfurt.de

Erfurt University (Nordhäuserstraße 63, 99089 Erfurt, Germany)

Abstract. I propose to solve the hard problem in the philosophy of mind by means of Brandom's notion of the pragmatically mediated semantic relation. The explanatory gap between a phenomenal concept and the corresponding theoretical concept is a gap in the pragmatically mediated semantic relation between them. It is closed if we do not neglect the pragmatics.

1 Introduction

In the second section, I will formulate the hard problem. In the third section, I will describe a pragmatic approach to the problem and propose to replace the classical non-normative physicalism/naturalism with a normative physicalism/naturalism of Wittgensteinian language games. In subsection 3.1, I will give a definition of a normative naturalism. In subsection 3.2, I will make some suggestions concerning an analytic interpretation of the second philosophy of Wittgenstein. In the fourth section, I will propose a solution to the hard problem within Brandom's analytic pragmatism by using the notion of the pragmatically mediated semantic relation. In the fifth section, I will make some suggestions about possible combinatorics related to pragmatically mediated semantic relations. In the sixth section, I will consider pragmatic and discursive versions of the mind-body identity $M=B$. In the last section, I will conclude that the explanatory gap is a gap in a pragmatically mediated semantic relation between B and M. It is closed if we do not neglect pragmatics.

2 The Hard Problem

The *hard problem* in the philosophy of mind can be formulated as follows. Let us suppose that type-type physicalism is true, and in particular, every phenomenal property M *is* (identical to) a physical property B, $M=B$. Let us suppose, for instance, that the property of being in pain is identical to the property of being excited C-fibers. Then it seems that M could have been identical to some other physical property than B. The problem arises from trying to explain why M is B and not some other physical property B' or not even a physical property at all. In other words, there is the so-called *explanatory gap* between M and B.

Of course, from the dualist's point of view, the explanatory gap between M and B is ontological, and it cannot be closed. The most popular physicalist approach to the hard problem says that the explanatory gap between M and B is purely epistemic, and is due to the specific character of the phenomenal concept M which picks out the same physical property as the theoretical concept B. The phenomenal concepts are generalizations of ordinary language concepts like "pain" and "red". The former can also be used non-discursively or introspectively. The conceptual expression of identity, $M=B$, where M is a phenomenal concept, B is a theoretical concept of the corresponding property, would produce the illusion of an ontological gap, because of the phenomenal character of M.

Nevertheless, it is known that in this case the problem re-appears on the conceptual level: how do we close the gap between phenomenal concepts and theoretical concepts or, if this gap cannot be closed, how do we explain the specificity of phenomenal concepts from the physicalist's point of view?

I propose to solve the hard problem by means of Brandom's (2008) notion of *pragmatically mediated semantic relation* between a phenomenal vocabulary (concept) and the corresponding theoretical vocabulary (concept); the explanatory gap can be closed pragmatically.

3 A Pragmatic Approach to the Hard Problem

There is a quietist pragmatic view on the hard problem: the hard problem is a pseudo-problem, and the explanatory gap is a philosophical illusion due to unnatural or confused use of language. In ordinary language, the explanatory gap does not appear. This is a pure dissolution of the problem. I will speak about a *solution* to the hard problem, because I will make an appeal to the meta-conceptual apparatus of meaning-use analysis developed by Brandom, which makes explicit the implicit pragmatic relations between vocabularies. This will allow us to understand the nature of the explanatory gap and give a constructive response to the question posed by Wittgenstein: « The feeling of unbridgeable gulf between consciousness and brain-process: How does it come about that this does not come into the considerations of our ordinary life? » (Wittgenstein, PI, 412).

I will suppose that pragmatism is normative: practices/language games are governed, in general, by implicit conceptual norms. I will consider these norms to be rules in the Wittgensteinian sense. Not only the practical abilities, or pragmatic concepts, but also the theoretical concepts can be understood as such norms/rules. Their exercise must be correct in practice.

One can see the notion of practice as a generalization of the notion of experience, and hence, phenomenal concepts as pragmatic concepts of phenomenal experiences. Then the explanatory gap is an apparent gap between a phenomenal concept and a theoretical concept referring to one and the same phenomenal experience or a gap between the concepts and their referent. Within the experience itself, where both concepts are implicit, the gap is absent. In Wittgensteinian terms, the explanatory gap is a gap between a rule and its application – a language game governed by this rule. It is closed pragmatically within the language game.

From such a pragmatic perspective, the identity $M=B$ is a language game governed by the concepts/rules B, M . The formal identity $M=B$ is also a rule.¹ In principle, the concepts B and M can be used differently: $M=B'$ and $M'=B$ are language games which are different from $M=B$. This means that the identity as a language game is not unambiguous. The type-type physicalism is true, although not in its classical non-normative version, but as a normative physicalism/naturalism – physicalism/naturalism of language games, which are both natural and normative at the same time.

3.1 A Normative Naturalism

A normative physicalism/naturalism can be understood in a few equivalent ways. It can be understood in the sense that there are non-reducible normative facts, that is, facts that certain things ought to be the case, or facts about how certain things ought to be, which are compatible with physicalist/naturalist ontology, or it can be understood as saying that at least some physical/natural facts are irreducibly normative or have a normative dimension in the sense that they must, in principle, be justifiable. In other words, according to normative physicalism/naturalism, norms are implicit in nature (in physical/natural facts). An assertional Wittgensteinian language game can be viewed as a normative fact or a judgment about a normative fact.

3.2 An Analytic Interpretation of the Second Philosophy of Wittgenstein

I think there is an analytic interpretation of the second philosophy of Wittgenstein (making it analytically explicit) as a specific normative naturalism of language games, which in its spirit is very close to Brandom's analytic pragmatism. The second philosophy of Wittgenstein is not a radical pragmatism, even though Wittgenstein stresses the investigation of the phenomenon of coming into being of new language games. It implicitly contains a reconciliation of the analytic philosophy and pragmatism. The idea of therapy of philosophy with help of natural ordinary language plays in it a secondary role. Wittgenstein himself is not a Wittgensteinian quietist. The real concern of the second philosophy of Wittgenstein is ontological – about the nature of “language games”.

The Wittgensteinian slogan *meaning is use* is not a reduction of the meaning to use, but a synthesis of both, an identity of meaning and use, or a pragmatic meaning-use relation.

Of course, the identity of meaning and use doesn't mean that there is no distinction between the “meaning” and the “use”. Wittgenstein also says that the meaning *is in the use*, that is, associated with the use. The slogan should be understood in the sense that there is no sharp border between the meaning and the use, between the

¹ This rule is a pragmatic metarule for the language game “ M is necessarily B ” elaborated from the game “ M is B ”. The game “ M is necessarily B ” can be formally represented as the Kripkean identity *necessary a posteriori*, that is, as the identity between two modally rigid designators.

conceptual and physical/natural. The natural use has a normative dimension – in a broad sense, it is a meaning. Conversely, every meaning can be naturalized: it is not only a theoretical entity associated with a use, but it is itself a certain use. This is a pragmatic naturalistic interpretation of the slogan. In Brandom’s notation this interpretation takes the following form: PV-suff (the discursive practice P deploys the vocabulary V) means “V is P” (the meaning V is the use P).

Here is a semantic interpretation of the slogan: the meaning of a word is the rule for its use. “Meaning is (in the use)” means: the application of the rule/meaning generates the use - a language game governed by this rule. In Brandom’s notation this means: VP-suff, that is, “the meaning V specifies the use P”.

If both PV-suff* \equiv PP’-suff & P’V-suff (V is elaborated from P) and VP-suff (V specifies P) at the same time, that is in Brandom’s notation, V is a LX-vocabulary relative to P, then one can say that the meaning V= P’ pragmatically supervenes on the use P. In philosophy of mind, the claim “consciousness is use” is an analogue of the claim “meaning is use”. Hence one can also define the notion of pragmatic supervenience of consciousness. It is known that the notion of modal supervenience is purely logical; it cannot, for example, explain the mental causation. Maybe a pragmatic unpacking of the notion of modal supervenience, if it exists, could allow us to resolve the problem of mental causation.

4 The Identity M=B as a Pragmatically Mediated Semantic Relation

I’d like to remind the reader that my proposal is to interpret the identity B=M analytically as a pragmatically mediated semantic relation between B and M. What is a pragmatically mediated semantic relation?

Let us look, for example, at Brandom’s interpretation of Sellar’s claim that the vocabulary of appearance, how things appear, V_{look} , presupposes the corresponding objective vocabulary of existence, how things actually are, V_{is} , as a $V_{is}V_{look}$ – *necessary* pragmatically mediated relation,

$V_{is}V_{look} - nec = P_{is}V_{is} - suff \ \& \ P_{is}P_{look} - nec \ \& \ P_{look}V_{look} - suff,$

where I use the symbol of conjunction & to connect the different components of the resultant pragmatically mediated semantic relation.

The meaning of notation can be described as follows. $V_{is}V_{look} - nec$ means: the capacity for deploying the vocabulary V_{is} is necessary to deploy the vocabulary V_{look} . $P_{is}V_{is}-suff$ means: the practice P_{is} is sufficient to deploy the vocabulary V_{is} . $P_{is}P_{look} - nec$ means: the practice P_{is} is necessary to elaborate the practice P_{look} . $P_{look}V_{look} - suff$ means: the practice P_{look} is sufficient to deploy the vocabulary V_{look} (Brandom 2008).

V_{is} and V_{look} are deployed in different practices, P_{look} and P_{is} , – the phenomenological vocabulary of appearance pragmatically emerges from the objective vocabulary, – but these practices are about one and the same object.

Let us take another example.

According to Brandom, non-indexical vocabulary can serve as an adequate *pragmatic metavocabulary* for indexical vocabulary, $V_{non-index}V_{index} - suff$. This pragmatically mediated semantic relation is an expressive bootstrapping.

For example, though the indexical ‘now’ is not semantically equivalent to the description “the time of utterance”, there is a pragmatically mediated semantic relation between them. One could write this pragmatically mediated semantic relation as the identity “*now is the time of utterance*”.

The practical rules for this language game are: “If at time t a speaker wants to assert that some property P holds at time t, it is correct to say ‘P holds now’ ”, “If a speaker at time t asserts ‘P holds now’, the speaker is committed to the property holding at time t.” (Brandom 2008, pp. 25-26)

Now, it is known that in some important respects phenomenal concepts resemble indexical concepts.

One can suppose that there is a pragmatically mediated semantic relation between M and B which is analogous to the pragmatically mediated semantic relation between indexicals and their third-person objective descriptions.

Let us suppose that, speaking in Brandom’s terms, there are two vocabularies (or only two concepts): a theoretical one V_{th} (B) and a phenomenal one V_{ph} (M), such that V_{th} (B) is $V_{th}P$ (BP) - suff, where the new

notation VP-suff means “the vocabulary V is sufficient to specify the practice P” (in this case V makes explicit or explains P), and the practice P, in turn, is PV_{ph} (PM) - suff to deploy the vocabulary V_{ph} (the concept M). I suggest that a pragmatically mediated semantic relation of such a kind might make explicit one of the possible pragmatic meanings of the identity M=B.²

In terms of the example above, the theoretical concept of being excited C-fibers specifies a practice which is sufficient to deploy the phenomenal concept of being in pain: *excited C-fibers = pain*. The specification itself is a certain pragmatic act, which can be activated or not. The referent, the property of being in pain, is a certain “practice”. It cannot be determined by only one concept - the theoretical concept of being excited C-fibers or the phenomenological concept of being in pain. This determination requires both these concepts: it is the whole pragmatic identity that determines the referent.³

More generally, a pragmatically mediated semantic relation includes what Brandom calls PP’ – suff relation: P is sufficient for the elaboration of the practice P’, and/or PP’- nec relation (see the pragmatically mediated semantic relation between V_{look} and V_{is} above). One can suggest that a *correlation* between B and M might be interpreted as a pragmatically mediated semantic relation of the following kind:

$$VBVM - \text{suff} = VB P - \text{suff} \text{ (or } PVB - \text{suff) \& PP}' - \text{suff} \text{ \& P'VM} - \text{suff} \text{ (or VMP}' - \text{suff)} \quad (1)$$

The pragmatic relation (1) means: “the concept B is sufficient to characterize the concept M”.

The PP’ – suff relation and VBVM - suff - the resultant pragmatically mediated semantic relation between vocabularies - can be understood as the Wittgensteinian extension of the initial practice P (vocabulary VB) by means of a pragmatic projection. Every normative practice can be extended in accordance with its norms/rules. The norms/rules themselves can be modified in the course of the extension. Such extension must satisfy the condition of family resemblance. I interpret the Wittgensteinian notion of family resemblance as follows: There is such-and-such family resemblance between two practices/language games if and only if these practices/language games are governed by the corresponding common, in general implicit, rule. From this point of view, language does have a “downtown”. It makes sense to speak of a family resemblance relation only between possible uses of a rule, which are not arbitrary and determined by a natural justifiable extension of the domain of its established uses.⁴

Since the pragmatically mediated semantic relation is not univocal, the expression of the identity “B is M” or the correlation “B correlates with M” is not univocal either. This expression is a certain practice that might involve the following elements:

VP-suff, VP-nec, PV-suff, PV-nec, PP’ – suff, PP’ – nec, VV-suff, VV-nec.

² Functional reductionism says that the mind-body identity M=B means that the physical property B is an empirical realizer of the *a priori* functional role F of the mental property M. Since there is no restriction on the possible physical realizers, such position is, metaphysically, functionalism: the realizers are not supposed to have anything in common, or resemble each other (see Block (Forthcoming) for a critique of functional reductionism as physicalism ontologically, but as functionalism metaphysically). Let us suppose now that, in addition, there is some kind of pragmatic relation between the functional role F and the property M and hence between F and B. This condition would play the role of a restriction on possible realizers of the functional role F. Then the functionalism is transformed into normative physicalism: M=F=B - identity as a pragmatically mediated relation, or a Wittgensteinian language game governed by the rules M, F and B. Let M=F=B’ be a different language game with a different pragmatic relation and a different realizer, B’. Then the family resemblance relation of some kind between B and B’ due to the common implicit rule M would play the role of a common first-order physical property in classical physicalism.

³ The *explanatory gap* between a phenomenal state and its neurological description cannot be closed merely by a more sophisticated description. An appeal to phenomenal concepts turns out to be necessary. Conversely, using neurological concepts is necessary to justify using phenomenal concepts; otherwise, the corresponding phenomenal knowledge may turn out to be “private knowledge”. The mind-body identity M=B must be viewed as a pragmatic connection between a phenomenal concept M and a neurological concept B. Neither the concept M nor the concept B alone fixes the referent of the identity. So-called *neuro-phenomenology* by Varela takes into account this conceptual dualism by proposing to extend the science by introducing into it phenomenal concepts and the corresponding practices of their use. (Varela 1995)

⁴ For Wittgenstein, as I understand him, the capacity to make assertions or claims, or equivalently, the capacity to use discursive rules/concepts is PV-nec for the deployment of any discursive vocabulary. In other words, this is the core of capacities common to all discursive practices. Furthermore, the capacity to use rules/concepts and the capacity to justify their use – the capacity of inferring - are mutually PP-nec. Hence, in these respects, I see Brandom’s position as exactly Wittgensteinian.

5 Combinatorics of Pragmatically Mediated Semantic Relations

It might be possible to introduce a system of “combinatorics” related to these elements.

If, for example PV-suff, where $V=V_{\max}$ is the maximal vocabulary deployed in the practice P, then also VP-suff. Moreover, $PV_{\max}\text{-suff} = V_{\max}P\text{-suff} = V_{\max}P\text{-nec} = PV_{\max}\text{-nec}$, or $V_{\max} = P$: meaning is use (see subsection 3.2).

More generally, one might suppose that $PV\text{-suff}^* = VP\text{-suff}$, where $PV\text{-suff}^* \equiv PP' - \text{suff} \ \& \ P'V - \text{suff}$ is the process of pragmatic elaboration/explicitation, VP-suff itself is a practice.⁵

Conversely, in general VP-suff does not entail PV-suff: a pragmatic metavocabulary V is not explicitly deployed in specified practice; it is implicit in it.

Note that the practice $P' = VP\text{-suff}$ – the practice of specifying P using V – is a practice in which the vocabulary V is deployed: $(VP\text{-suff})V\text{-suff}$. It seems that the latter practice can be pragmatically elaborated from the practice P: $P(VP\text{-suff}) - \text{suff}$. Then the vocabulary V both specifies and is elaborated from P.

Hence my hypothesis is that the elaboration/explicitation $PV\text{-suff}^* = PP'\text{-suff} \ \& \ P'V\text{-suff}$ and the specification VP-suff is one and the same process: $(VP\text{-suff}) = (P' = V, \text{ i.e., } (P'V\text{-suff})) = PV\text{-suff}^*$, where V is a LX-vocabulary relative to P. The only difference between the elaboration and the specification is in the order of determination: the former makes explicit an implicit rule/concept contained in P, the latter applies the explicit rule/concept to P. In subsection 3.2, I suggested that the LX - relation might be used to define “pragmatic supervenience”.

Further, one and the same practice P might be sufficient to deploy two different vocabularies V_M and V_B , for example, in the following sense: $PV_M - \text{suff}^* = PP' - \text{suff} \ \& \ P'V_M - \text{suff}$, $PV_B - \text{suff}^* = PP'' - \text{suff} \ \& \ P''V_B - \text{suff}$. Then $PV_M - \text{suff}^* \ \& \ PV_B - \text{suff}^* = VMP - \text{suff} \ \& \ VBP - \text{suff}$, that is, both V_M and V_B are LX-vocabularies relative to P.

Brandom shows that the practice of rational rectification Prs is $PrsV_N - \text{suff}$ and $PrsV_M - \text{suff}$, where V_N is the vocabulary of normative incompatibility and V_M is the vocabulary of modal incompatibility. In detail, $P_1P_N - \text{suff} \ \& \ P_NV_N - \text{suff} \ \& \ P_2P_M - \text{suff} \ \& \ P_MV_M - \text{suff}$, where P_1, P_2 are two sub-practices of Prs. The two senses of ‘incompatibility’ – normative and modal – are not homonyms. This reflects the existence of an intimate pragmatic relation between both vocabularies.⁶

It seems to me that the relations between conceptual normativity and modality can be understood in Wittgensteinian terms as follows. A factual judgment is a Wittgensteinian assertional language game. The corresponding rule/concept is implicit in the fact viewed as a purely pragmatic language game from which the judgment is pragmatically elaborated and which it makes explicit or explains.⁷ The modal question is the question about how the fact might be or how the fact could have been. The possibilities here are not the Kripkean metaphysical possibilities but the “real possibilities” of the Wittgensteinian rule use, that is, the modal language games “might be” (“could have been”) associated with the factual language game.

In Brandom’s terms, the intentional explanatory gap – a gap between words and world – is a gap between the subjective/normative and the objective/modal poles of intentional relations, or, in my terms, it is a gap between the Wittgensteinian rule and its applications. Brandom understands intentionality to be a pragmatically mediated semantic relation which closes the intentional explanatory gap. The explanatory gap between the phenomenal and physical can be closed analogously.

⁵ $PV\text{-suff}^* = VP\text{-suff}$ does not entail $PV\text{-nec}$, since V could also specify a different practice P' , $VP'\text{-suff}$. However, there must be a family resemblance between P and P' .

⁶ Note that for Wittgenstein, the modal incompatibility, such as “something cannot be both red and green at the same time”, is a conceptual truth: a sentence of the “philosophical grammar”. In this sense, modal incompatibility is conceptual incompatibility.

⁷ I interpret the Wittgensteinian notion of a rule as determined by its explicit general formulation and a set of its established applications between which there is a relation of family resemblance due to the common rule. It is the ordinary language notion of a rule. The general formulation of a rule can be viewed as a formulation of a necessary law, the applications – as the possibilities. In such an interpretation, the conceptual normativity is the modality.

6 Pragmatic Relations between Pragmatic and Discursive Identities

According to Brandom, there is an intimate, pragmatically mediated semantic relation between purely pragmatic intentionality, partly pragmatic - partly discursive intentionality and purely discursive intentionality. Respectively, the identity $M=B$ can be more or less discursive.

One can distinguish, for example, between the following cases: (a) pragmatic identity $M=B$ (the rules/concepts M , B and $M=B$ are implicit), (b) partly discursive identity $M=B$ as an explicit use of the rule/concept M (B) (M (B) specifies the pragmatic identity, the rule/concept B (M) is implicit), (c) discursive identity $M=B$ (M and B specify the common referent), (d) formal identity $M=B$ as an explicit rule.

If the identity $M=B$ is purely pragmatic, a gap or any appearance of a gap between M and B is obviously absent. If it is partly discursive, the gap is absent thanks to the pragmatic relation between the discursive and pragmatic levels. Nevertheless, the appearance of a gap can, for example, be due to the confusion between these levels resulting from neglecting the pragmatic relation between them. This confusion can be viewed as the confusion between reality and its description. I think that the struggle against such confusions is one of the central points of the second philosophy of Wittgenstein.

The discursive pragmatic relation between B and M (case (c) above) is weaker than the pragmatic relation between them within the purely pragmatic identity (case (a)); so it can be easily neglected. This entails the doctrine of dualism and the view about a purely nomological relation between two kinds of properties.

In Brandom's terms, the case (c) is the following pragmatically mediated semantic relation:

$$\text{VBP-suff} \ \& \ \text{VMP-suff}, \tag{2}$$

where the rules/concepts M and B are explicative of, or specify, one and the same practice P .⁸ Note that (2) includes the case $\text{VBP-suff} \ \& \ \text{PVM-suff}$.⁹ In general, the practice (2) does not include the practice P itself – the common referent of M and B .

7 Conclusion

In conclusion, I claim that the referent of the mind-body identity is not a bold physicalist property, but a physicalist property with a justification, or a property as a language game. Such referent is not static, but dynamic. In other words, the identity $M=B$ viewed as a language game is a contextual normative fact or a judgment about it; that is, it must be justifiable in a given context. The formal identity $M=B$ is an epistemic rule. The classical non-normative physicalism is false.

The pragmatically mediated semantic relation allows for the closing of the gap between a phenomenal concept M and a theoretical concept B . The practice of specification is a gapless pragmatic connection between the term M (B) and the referent/practice P , and so between M and B . The explanatory gap is a gap in the pragmatically mediated semantic relation. It is closed if we do not neglect pragmatics. In Brandom's (2008, p. 193) words: what relates M and B "is a *process, a practice, the exercise of an ability, a kind of practical doing*".

References

- Block, Ned. (Forthcoming) "Functional reduction". In a festschrift for Jaegwon Kim, *Supervenience in Mind*, edited by Terry Horgan, Marcelo Sabates and David Sosa.
- Brandom, R. B. (2008) *Between saying and doing*, Oxford, Oxford University Press.
- Varela, F. (1995) "Neurophenomenology: A methodological remedy for the hard problem", *Journal of Consciousness Studies*, 3, pp. 330-49.
- Wittgenstein, L. (2001 [1953]) *Philosophical investigations*, G.E.M. Anscombe (trans.). Malden, Mass.: Blackwell.

⁸ The practice of specification can be more or less close to the specified practice P , so that at a limit when PV_{\max} -suff, $V=V_{\max}$, it coincides with P .

⁹ Here is an analogy to the mind-body identity. According to Brandom (2008, p. 73), the isomorphism ("identity") established by Descartes between algebraic formulae and geometric figures is a pragmatically mediated semantic relation between syntactic properties and semantic ones. Note that mathematics and physics are imbued with different kinds of more or less pragmatic identities and equivalences which, after having been made explicit, become different kinds of "isomorphisms". The "heuristic" pragmatic identities and equivalences are not something approximate. They are authentic practices.

Toward an Analytic-Pragmatist Account of Folk Psychology?

Derek W. Strijbos
Leon C. de Bruin

Dept. of Philosophy
Radboud University Nijmegen
Erasmusplein 1
P.O. Box 9103
6500 HD Nijmegen

Institute for Philosophy
Leiden University
Matthias de Vrieshof 4
P.O. Box 9515
2300 RA Leiden

In this paper we want to establish two points of interest to Brandom's project of analytic pragmatism. First, we aim to show how, in light of this project, Brandom's deontic scorekeeping model can be used as a valuable descriptive tool for characterizing folk psychological interpretation, and how this reveals certain problematic assumptions underlying dominant positions in the debate on folk psychology. Second, we will consider the relevance of empirical study of folk-psychological practice to Brandom's project of analytic pragmatism: to what extent should pragmatic analyses of the sort Brandom proposes be subject to empirical research on what people actually say in the practice of giving and asking for reasons?

1. Introduction

In this paper we want to establish two points of interest to Brandom's project of analytic pragmatism (2008). First, we will show how the project can be relevant for a philosophical debate that is not directly associated with Brandom's work: the debate on so-called 'folk psychology' in the philosophy of mind. The positions that still dominate this debate are inspired by mentalistic and formalist assumptions regarding folk-psychological interpretation (section 2). Getting these assumptions into clear view and assessing their propriety requires an overarching account of interpretation that is at the same time informative in its characterization of what people are doing during interpretation and neutral regarding rivaling explanations of this practice. Brandom's pragmatic analysis of deontic normative vocabulary (in terms of 'commitments' and 'entitlements') explains how his deontic scorekeeping model (1994) can meet this demand and why it should thus be taken seriously as a starting point in the debate on folk-psychology. Doing just that, we shall briefly argue why mentioned assumptions are problematic and sketch the outlines of an alternative account (section 3). Second, we shall consider the relevance of empirical study of folk-psychological practice to Brandom's project of analytic pragmatism. On the one hand, analytic pragmatism appears to have an unprecedented potential of opening up to empirical research. On the other hand, it is not clear in what ways empirical results are to contribute to the formulation of the kind of pragmatically mediated semantic relations that Brandom aims at (section 4).

2. Mentalistic and Formalist Assumptions Regarding Folk Psychology

There has been a tendency among philosophers of mind to use the term 'folk psychology' to refer to the practical lore and know how that 'the folk' puts to use in making sense each other in everyday life. Our focus in this paper is on a fairly 'high-level', sophisticated folk-psychological capacity: the capacity to interpret each other's behavior in terms of *reasons for action*. It is important to be clear from the start about what this capacity comprises: in recent years, a growing number of philosophers and psychologists have convincingly argued that much of our daily social interaction doesn't involve the attribution of full-blown reasons to fellow agents. The lion's share of our embodied engagements with each other doesn't seem to require interpretation at a propositional level (Gallagher 2001).

Furthermore, our social interactions take place in socially structured, normalized environments in which the need for interpretation in terms of the *individual* agent's reasons is often obviated. As long as people do what they are supposed to do, according to the rules of social practice, they can often get along fine without focus on each other's reasons (Bruner 1990, Hutto 2004, 2008). Still, there are occasions on which it is important to find out the reason *in light of which* the agent performed a particular action. While it is true that traditional accounts of folk psychology have significantly overestimated the *scope* of this particular folk psychological competence, this by itself says nothing about the theories they put forward concerning its *nature*. It is the latter issue that is our prime target here.

There are two dominant positions in the debate on folk psychology: theory theory (TT) and simulation theory (ST). According to TT, interpretation proceeds by calling upon an acquired or innate folk psychological theory - a body of generalizations that functionally structures the relation between perceptions, mental states and actions. ST claims that our default procedure in making sense of behavior in terms of reasons for action consists in simulating the mental process responsible for the action by, in some way or other, 'placing ourselves in the shoes' of the agent under consideration.

A remarkable feature of both approaches is that they tend to conceive of reason interpretation as a process of mental state attribution. Inspired by the metaphysical thesis of functionalism in the philosophy of mind, TT maps folk psychological understanding on the alleged functional or causal roles of mental states. On all construals ('modular TT' (Fodor 1992, Leslie et al. 2005), 'scientific TT' (Gopnik and Meltzoff 1997), 'modal TT' (Maibom 2003), 'external TT' (Braddon-Mitchell and Jackson 2007) the folk's understanding of other people's actions in some way or other depends on innate or acquired (tacit) knowledge that specifies the functional roles of mental states. ST's original aim was to fend off such functionalist rendering of folk psychological interpretation (Gordon 1986, Heal 1986). While some ST-ists (most notably Gordon 1996, 2000, 2002) have also resisted the heavily mentalized picture of interpretation that went with it, the dominant cognitivist strand in ST simply replaced the 'theory-driven' interpretation process postulated by TT with a 'process-driven' mindreading procedure (Goldman 1989). On this 'pretense-driven off-line simulation' account (Stich and Nichols 1997), interpretation proceeds by feeding pretend beliefs and pretend desires into one's own offline practical reasoning system and attributing the resulting pretend decisions (and corresponding actions) to the agent who is being simulated (Goldman 1989, 2006).

These mentalistic views are often accompanied by a strong commitment to the belief-desire model of action interpretation. Accordingly, interpreting others in terms of reasons for action requires the (tacit) (re)construction of the action under consideration in terms a constellation of beliefs and desires, primarily a desire toward some goal and a belief regarding the means. The interpretation process thereby obeys the 'central action principles' of belief-desire psychology, perhaps the most salient one being 'if A wants P and believes that doing q will bring about p, then *ceteris paribus*, A will q.' (Borg 2007, p. 6) On TT-accounts, such principles are simply part of people's theory of mind.¹ As such, they must be (tacitly) represented in some way or other. On cognitivist ST-approaches, the process-driven mindreading procedure must at least *mirror* these principles: a pretend-decision can only be reached by feeding a pretend belief and a pretend desire into one's own offline practical reasoning mechanism.²

As a consequence of all this, reason explanations in ordinary discursive practice in terms of facts, values, beliefs *or* desires are treated as essentially *truncated* versions of a kind of practical syllogism that specifies the underlying interpretation process. On these mentalistic and formalist assumptions, such explanations are only regarded as appropriate in virtue of a mindreading procedure filling in the omitted premises.

3. Interpretation as Deontic Scorekeeping

These assumptions, we argue elsewhere, stem from a deep-rooted bias of mentioned accounts in favor of third-person interpretation practices (Strijbos and De Bruin, forthcoming). Here the *caveat* regarding the scope of folk psychology, mentioned at the beginning of the previous section, becomes extremely relevant. For what is at issue, and what has always been at issue in the debate, is the folk's capacity to interpret each other in terms of *their reasons for action*. And what appears to have been

¹ Frith and Happé (1999, p. 2), for example, say that 'in everyday life we make sense of each other's behaviour by appeal to a belief-desire psychology.'

² Thus, Goldman thinks that 'a decision-making mechanism normally takes genuine (nonpretend) desires and beliefs as inputs and then outputs a genuine (nonpretend) decision. In simulation exercises, the decision-making mechanism is applied to pretend desires and beliefs and outputs pretend decisions.' (2006, p. 29)

systematically overlooked is that when and where it really matters to know the *agent's* side of the story, that is: to know *his* reasons, people *don't trust* their own theories and simulations; they tend to ask someone, preferably the agent himself. *Speculating* about the reasons of others is something people normally engage in only when asking is inappropriate, inconvenient or impossible. Reason interpretation, as Hutto (2004, p. 565) puts it, is not primarily a 'spectator sport'. Compared to the answers we get 'from the horse's mouth', the explanations and predictions drawn from reason speculation are generally far less accurate and successful. In short, there are good reasons to think that when and where it is important to get the agent's perspective into clear view, the primary way to achieve this is *by participating in reason discourse*.

Reason discourse being its primary context of application, participants in the debate should reconsider their mentalistic and formalist assumptions regarding action interpretation in terms of reasons. Shifting focus to the second-person practice of giving and asking for reasons, it seems reasonable to start by taking reason discourse at *face value* and closely study what people *are actually saying* when they are trying to make themselves understood. Consider the following brief conversations:

A: "Why are you wearing a tie?"

B: "Because it is important to make a good impression."

C: "Why aren't you coming to the party tonight?"

D: "I want to finish my paper."

E: "Why are you getting up so early?"

F: "It is Monday."

G: "Why did you get up so early?"

H: "I thought it was Monday."

In all four examples the interpreter asks the agent for his reasons for action. Now suppose that the interpreter *accept* the agent's answers in these different scenarios. Taking these examples at face value, we ought to at least consider the agents' *responses* as their *reasons*. That is: to consider their reasons to be 1) facts (that it is Monday), 2) values (that making a good impression is important), 3) desires (towards finishing the paper) or 4) beliefs (that it was Monday).

To many this may seem an outrageous proposal. For how could interpretation of *actions*, even in the context of second-person reason discourse, get off the ground without at least implicit reference to the agent's *motivation*? This rhetorical question is well taken, but hard to pin down. It all hinges on what is meant by 'implicit' and 'motivation'. On TT and (cognitivist) ST proposals, 'implicit' equals 'tacit' and 'motivation' comes down to 'well-formed belief-desire pair'. Taking it this way is surely begging the question at this point: arguments need to be provided as to why this is the interpretation of 'implicit' and 'motivation' that we should endorse. We think there is clearly another sense in which it is indeed nigh inconceivable to understand an agent's reasons without a grasp of what it is that moves him. But making this sense explicit is a complicated affair as long as the terminology available is dominated by a mentalistic, formalist reading.

The challenge is to give a non-question begging characterization of what it is that people are doing when they are participating in the practice of giving and asking for reasons. It is against this background that we should look at Brandom's deontic scorekeeping model (1994), especially in light of the pragmatically mediated semantic relation he proposes for deontic normative vocabulary in chapter 4 of his 2008. On the scorekeeping model, interpreters are *withholding entitlement* to (having non-inferentially responded to) a practical commitment (to act) when they ask an agent for his reasons for action. In giving his reason, the agent is attempting to *vindicate entitlement* for his practical commitment by *acknowledging* (and undertaking) *a commitment* (and claiming entitlement) to his answer, to the effect that he *takes responsibility* for vindicating entitlement to it (and its committive consequences) when challenged. Accepting the agent's answer, the interpreter 1) *grants authority* to the agent in *attributing entitlement* to the commitment expressed in the answer and 2) endorses that answer as providing a reason for action by *attributing entitlement* to the agent's original practical commitment (to act). According to this model, what people do in participating in reason discourse is *keeping score* of what each participant is committed and entitled to say and do, given the features of the world in which the game unfolds, their past claims and actions, social standards and the authority of

others. In saying or doing something, participants *change their deontic scores* by altering the constellation of moves they are committed and entitled to make and the moves they are prohibited from making. Thus, actions (and speech acts) may stand in need of reasons in case they constitute a move to which fellow scorekeepers do not attribute entitlement. Under such circumstances, performing a certain speech act will count as giving a reason if it can vindicate one's entitlement.

Now why should this model provide the neutral description of interpretation practices we are looking for? By characterizing reason discourse in terms of the attribution of *commitments* and *entitlements*, it is left wide open as to whether such attribution further requires the attribution of *mental states* and whether this be done according to the belief-desire model. It seems to us that all contenders in the debate on folk psychology could agree on this: that withholding entitlement to a commitment to act is indeed a way of saying what one does in asking someone for a reason for action. *Mutatis mutandis* for the other characterizations given above.

But looked at in light of the project of analytic pragmatism, there is more to it than this. For why should we be confident with *this* particular vocabulary? With Brandom, we could endorse what he calls the 'normative Kant-Sellars thesis' (2008, pp. 109-116) and say that the normative vocabulary of 'commitment' and 'entitlement' stands in an 'Elaborated-Explicated' (LX) relation to the practice of giving and asking for reasons. What this means is that the linguistic practice that suffices to use this normative vocabulary can be *algorithmically elaborated* from the practice of giving and asking for reasons (which on Brandom's account is moreover a necessary condition for utterances to count as *linguistic* utterances) and in turn suffices to *make explicit* this very same practice by enabling people to *say* what they are *doing* when they are saying things in the course of reason conversation (see figure 1).³ For Brandom, this means that mentioned normative vocabulary is a species of *expressive* (logical) vocabulary that enables one to codify the know-how present in the practice of giving and asking for reasons in explicit know-that.

Normative Kant-Sellars Thesis: Normative Vocabulary is Elaborated-Explicating (LX)

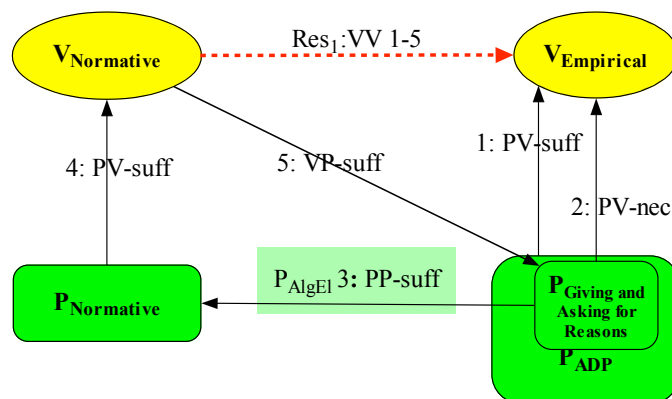


Figure 1

Importantly, in having this specifically expressive function, the normative vocabulary does not alter the reason discourse it is directed at: it does not add anything *substantial* to what people are already saying to each other in giving and asking each other for reasons. Rather it makes explicit what was already implicit in reason discourse. What is implicit is a particular kind of *doing*, not some piece of tacitly represented propositional content. Consider the above example again.

A: "Why are you wearing a tie?"
B: "Because it is important to make a good impression."

³ Figure adopted from <http://www.pitt.edu/~brandom/locke/locke-w4.html>

In accepting B's answer that it is important to make a good impression, A attributes entitlement to B's acknowledged commitment *that it is important to make a good impression* and thereby attributes entitlement to his wearing a tie. Nothing substantial is added to what is actually being said during the conversation. Compare this to a mindreading account, according to which by accepting B's answer, A must be tacitly attributing to B something like a *desire* (felt duty?) *to make a good impression* and a *belief that wearing a tie is a (the?) way to make a good impression*. Characterizing the interpretation process this way significantly expands the content of interpretation beyond what is actually being said, without proper argument.

Using Brandom's deontic scorekeeping model as a neutral descriptive tool to characterize what people are doing in interpreting each other is of course only a first step to providing an explanation as to *how* people are able to do these things. Here we think all options should *prima facie* be taken seriously, including the versions of TT and cognitivist ST criticized above. Considerations of parsimony strongly speak against these formalist mindreading accounts, however (Strijbos and De Bruin, forthcoming). Taking the second-person practice of reason discourse as the primary mode of acquiring and deploying reason interpretation skills, our alternative account uses Sellars' (1953) (and Brandom's) notion of *material inference* to explain how people withhold and attribute commitments and entitlements to each other (*ibid.*).

Consider the examples above once more. Our suggestion, in short, is that in responding to the question why he is wearing a tie by saying that it is important to make a good impression, B is constructing the *practical material inference* "It is important to make a good impression, therefore I shall wear a tie." In attributing entitlement to B's wearing a tie, A herself *endorses* the proposed practical material inference against the background of B's deontic score. Similarly, by answering that he wants to finish his paper, B is constructing the material inference "I want to finish my paper, therefore I shall not go to the party tonight." Etc.

An attractive feature of this account, we argue, is that it puts much less cognitive burden on the shoulders of the interpreter. The interpreter can simply follow the lead of the agent by endorsing (or rejecting) the material inference proposed by the agent. Moreover, the content of the inference is entirely explicit in what is actually said, nothing needs to be tacitly added in order to understand the agent's answer. Of course the material inference needs to be assessed against the background of the agent's deontic score. But the interpreter doesn't have to *guess* at these scores (by means of theory or simulation). In normal intra-cultural practice, the deontic scores of agent and interpreter significantly overlap. And to the extent that they don't, the additional 'personalized' scores will normally have been provided by the agent himself, by his sayings and doings in the past.

4. Empirical Issues

In our view, making progress on mentioned how-question regarding the attribution and withholding of commitments and entitlements requires both philosophical and empirical effort. There is simply no way to decide between rivaling accounts without paying proper attention to what people are actually saying and doing in reason discourse. The question we would like to raise in this last section is to what extent this is also true of the pragmatically mediated semantic analyses that form the centrepieces of Brandom's project of analytic pragmatism.

At first sight, Brandom's project seems to have real potential to reach out to the empirical sciences. By considering the semantic relations between different vocabularies as being *born out* of ordinary discursive practice, study of uses of these vocabularies in discursive practice can be expected to yield important insights into the logical structure of concepts that are of philosophical importance. Taking the semantic relations between different vocabularies to a pragmatic level, the project of philosophical analysis frees itself from its often-criticized commitment to *a priori* methods of inquiry and opens up possibilities for an empirically informed and essentially hermeneutical way of investigation.

Endorsing such a hermeneutical approach would of course make our own adoption of Brandom's deontic scorekeeping model vulnerable to empirical considerations. It is open to philosophers, psychologists and other researchers to refute our claim of neutrality with regard to the normative vocabulary of 'commitment' and 'entitlement' by collecting evidence that speaks against the LX-analysis that backs it up. What in our view really speaks for Brandom's analytic pragmatism is that such considerations can be allotted a proper place within the project itself.

It is not clear to us to what extent Brandom himself would endorse this empirical-hermeneutical approach. In particular, it would be interesting to find out if and to what extent he would want to see LX-relations between different vocabularies reflected in 1) ontogenetic development and 2) the structure of actual discursive practice. With regard to the first: to what extent should the

