UNIVERSITY OF CALIFORNIA, SAN DIEGO

Brandom's Expressive Conception of Logic

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Philosophy

by

Adam Joseph Streed

Committee in charge:

Professor Rick Grush, Chair
Professor Gedeon Deák
Professor Rafael Núñez
Professor Gila Sher
Professor Clinton Tolley

2015

Copyright
Adam Joseph Streed, 2015
All rights reserved.

The dissertation of Adam Joseph Streed is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2015

DEDICATION

To my parents.

TABLE OF CONTENTS

Signature Page iii

Dedication iv

Table of Contents v

List of Figures vii

Acknowledgements viii

Vita ix

Abstract of the Dissertation x

Chapter 1 Introduction 1

Chapter 2 Inference, Historical and Brandomian 7
2.1 introductory 7
2.2 inference 16
2.2.1 the traditional account of inference 16
2.2.2 traditional correctness for arguments and inferences . 20
2.3 Brandom's inversion of the traditional account 22
2.4 deontic scorekeeping: a first pass 26
2.5 looking ahead 32

Chapter 3 Material Inference and the Expressive Conception of Logic 34
3.1 material inference 34
3.1.1 the doctrine of logic as form 35
3.1.2 Sellars on material correctness 38
3.1.3 Brandom on material inference 46
3.2 logical vocabulary and deontic scorekeeping 50
3.2.1 'if' and 'not' 51
3.2.2 quantifiers, identity, and other connectives 53
3.3 the priority of material inference 59
3.3.1 the argument against enthymemes 60
3.3.2 formal inference as definitionally derivative 64
3.3.3 Dennett's principle 65
3.3.4 explaining logical competence 66
3.3.5 what priority amounts to 67
3.4 the expressive conception of logic 68
3.4.1 but isn't logic formal, not material? 71
3.4.2 isn't primitive material goodness unacceptably weird? 73

3.5 further questions 77
3.5.1 are there parallel logics? 77
3.5.2 material inferences are defeasible 80
3.5.3 what about Tarski's problem? 83

Chapter 4 The Expressive Conception as a Form of Psychologism 89
4.1 naturalism, expressivism, psychologism 90
4.1.1 naturalism 90
4.1.2 expressivism, the basic idea 95
4.1.3 two approaches to the meaning of logic 97
4.1.4 psychologism 101
4.2 scorekeeping and psychology 104
4.3 classic objections to psychologism 107
4.3.1 the shakiness of psychology 107
4.3.2 normative and descriptive logical laws 110
4.3.3 reference to psychological objects 116
4.3.4 truth and taking-to-be-true 121
4.3.5 species relativism and universality 126
4.4 naturalism in a foundational theory of logic 129

Chapter 5 The Normativity of Logic 130
5.1 logic's historical arc 130
5.1.1 logic as concerned with acts 131
5.1.2 logic as concerned with contents 133
5.1.3 norms and facts 134
5.2 Ryle and Sellars on norms implicit in logic 142
5.2.1 Ryle 142
5.2.2 Sellars 144
5.3 acts and norms: the expressive conception 148
5.3.1 conveyance as a constraint on bridge principles . . . 148
5.3.2 logic and norms on Brandom's expressive conception 150
5.3.3 Brandomian bridge principles 160
5.3.4 Harman's challenge 162

Chapter 6 Conclusion 164

References 166

LIST OF FIGURES

Figure 2.1: An Incomplete Taxonomy of Meaningful Things 9
Figure 2.2: A Simple Episode of Scorekeeping (from Player B's Perspective) . 27

ACKNOWLEDGEMENTS

I have benefited from the help of many, many people during the completion of this project. Rick Grush, Gila Sher, and Clinton Tolley were patient and encouraging while I struggled to assemble it, and gave me the benefit of the doubt when I first proposed it in its roughest form.

I received crucial feedback from various partners in my dissertation work group over the years, including Tim Jankowiak, Erick Ramirez, Cole Macke, Veronica Pear, Amanda Brovold, Per Milam, and Joyce Havstad; along with audiences at the UCSD Graduate Philosophy Colloquia. I owe a debt to John Dougherty for a helpful recent discussion of proof and inference. My former officemate Matt Brown deserves credit for introducing me to pragmatism, and prompting me to think more deeply about logic than I'd done before.

At UCSD, I am fortunate to have had the guidance of Pippin Schupbach, Nancy Guerrero, and especially Catherine Asmann, all of whom helped me navigate an unfamiliar place.

Finally, I owe my deepest gratitude to Joyce Havstad, without whom I would not have accomplished this at all.

VITA

2003	B.A. in Philosophy and Philosophy-Neuroscience-Psychology, Washington University in St. Louis
2008	M.A. in Philosophy, University of California, San Diego
2015	Ph.D. in Philosophy, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Brandom's Expressive Conception of Logic

by

Adam Joseph Streed

Doctor of Philosophy in Philosophy

University of California, San Diego, 2015

Professor Rick Grush, Chair

This dissertation is an exploration of the *expressive conception* of logic, as found in the work of Robert Brandom. I situate the expressive conception against a historical background of thought about logic, investigate the key notion of *material inference* on which the expressive conception rests, and argue that the expressive conception counts as a form of *psychologism* about logic which avoids many of the difficulties faced by older forms of psychologism. Finally, I argue that the expressive conception can make sense of logic's *normative import*, and offer as a test case the normative significance of validity.

Chapter 1

Introduction

In this project I explore the *expressive conception* of logic set forth in the recent works of Robert Brandom.[1] I think this is worth doing for several reasons.

First, the expressive conception is a rare example of attempting what Gila Sher (2013) has called a "substantive theoretical foundation for logic," a philosophical theory which would enable us to answer fundamental questions about logic such as whether and in virtue of what logic is true, what logic's normative status for thought is (and why), what to make of the variety of formal systems called logic (and which, if any, is the right one), which expressions are the logical constants, what the correct analysis of logical consequence is, and so on. Each of these questions can be, and has been, addressed in a relatively local fashion, independently of others—witness the literature on the demarcation of the logical constants, for example, where many contributors make appeal to intuition regarding what expressions are to count as "logical."[2] A substantive

[1] See especially Brandom (1994, 2000, 2008), along with some discussion and replies in Weiss & Wanderer (2010).

[2] See, for example, McGee (1996), Hanson (1997), Gomez-Torrente (2002). For another example in a slightly different literature, see Beall and Restall's much-discussed defense of so-called logical pluralism (2006), in which they argue that there are at least three genuine deductive consequence relations. Their argument begins by identifying what they call the "conceptual core" of deductive logical consequence, on the basis of appeals to intuition about which consequences are genuinely logical.

theoretical foundation, by contrast, would provide an answer to many at once, and in a unified way which springs from a conception of what logic is. If any proposed substantive theoretical foundation can be made to work, it will be an exciting and deep intellectual achievement.

Second, within the small group of substantive foundations for logic, the expressive conception is an unusual contender. It is—or so I argue—a variety of *psychologism*, the old and now-unpopular idea that logic's foundation is to be found in psychology. If a form of psychologism about logic can avoid the usual suite of problems which led philosophers to disregard it, then we can have a substantive foundation for logic which is *naturalistic*. Although 'naturalism' means many things in philosophy, not all of them equally appealing, there is a broad sense in which naturalism is a laudable philosophical goal, and a novel naturalistic account of logic is a valuable addition to philosophy's corpus. Interestingly, other recent attempts at providing a substantive foundation are either avowedly naturalistic (Maddy 2007) or psychological (Hanna 2006), but not both. Part of the distinctiveness of the expressive conception is thus, as I see it, its combination of both psychologistic and naturalistic elements. But it is also unusual in its espousal of an *expressive*, in contrast to *epistemic*, role for logic. The idea that logic's significance lies primarily in what it lets us say, rather than in its certainty, or its ability to extend our knowledge, is quite unusual.

Third, although the expressive conception of logic features in several of Brandom's works, it does so always as a supporting player in a larger systematic philosophical enterprise, not as a lead. And although Brandom's remarks regarding logic have received a few technical treatments,[3] or applications to cognate fields,[4] there has been very little engagement with the philosophical significance of the expressive conception. Instead, the bulk of philosophical engagement with Brandom's work has been on other aspects

[3]See, for example, Lance and Kremer (1994, 1996), and Lance (2001).
[4]For example, Kibble (2006), Keeler (2004), Whiting (2008).

of his theoretical philosophy, for example whether it can deliver a robust enough sort of *objectivity* in the way it theorizes the conceptual content of claims, or whether the norms which Brandom holds are implicit in practice can really be instituted by attitudes in the way he claims they are.[5] These are interesting questions, but they leave unexamined the distinctive conception of logic on offer. So philosophers of logic lack a dedicated resource for understanding Brandom's proposal, along with any accounts of how that proposal can provide answers to the various local questions in the philosophy of logic. I offer such a resource, and a few of those accounts, in this work.

Finally, from within Brandom's philosophical system, there is something valuable about exploring the expressive conception of logic in this way. For in Brandom's theoretical account of language and thought, what it is to *understand* a claim is to know both what it follows from and what follows from it. On this construal of understanding, the expressive conception of logic offered in *Making It Explicit* is only dimly illuminated.[6] Although the conception does figure in, for example, a demarcation of the logical constants, there are few local problems in the philosophy of logic which Brandom explicitly addresses, and so we don't get to see what follows from the expressive conception in the broader philosophy of logic. Nor do we get much in the way of an investigation of *material inference*, one crucial notion in terms of which the expressive conception is characterized. My remarks here are not intended as a criticism of Brandom, by any means—his primary concerns are elsewhere, in the systematic philosophy of language and mind—but merely to point out that if we want to better understand the expressive conception, then, by Brandom's own lights, what we need is to see what that conception follows from, and what follows from it. So one of the aims of this dissertation is to do just that, by investigating the notion of material inference, and by applying the expressive conception to the broad philosophical problems of psychologism, and of the normativity

[5] See Rosen (1997, 2001), Rosenberg (1997), Fodor and Lepore (2001, 2007), Hale and Wright (2010).
[6] The same is true of Brandom (2000), and also, though to a lesser extent, Brandom (2008).

of logic.

Thus in this work I have both exegetical and exploratory aims. I take myself to be doing something analogous to what contemporary virtue theorists in ethics do when they return to Aristotle's ethical treatises, offering both an interpretation of Aristotle's work but also applying his insights to problems in contemporary moral theory (for example, the question whether acting morally always conduces to a person's well-being). The point is not simply, or perhaps not even, to provide a reading of Aristotle which contributes to Aristotle scholarship, but to articulate the promises of a new way of thinking about ethics and the moral life.[7] Likewise, I do not simply aim to offer a reading of Brandom, but also to articulate the promises of a new way of thinking about logic and some local problems in the philosophy of logic.

The list of local problems in the philosophy of logic is long and open-ended, and, being unable to address them all, I have chosen just three for dedicated treatment in this dissertation: how logic, as a specifically formal discipline, stands to the "matter" of thought and speech; whether a psychologistic foundation for logic is tenable; and what the normative import of logic is for reasoning and thought. I've chosen these three because they are historically important problems for thought about logic, and because I believe Brandom's expressive conception has interesting consequences for them. Some other local problems will be discussed along the way—in particular, the demarcation of the logical constants, the issue of deviant logics and pluralism, and the relative priority of proof over semantics—but they will not receive dedicated treatment. I hope that other philosophers interested in Brandom's work can give these problems and others more attention than I have given them here.

The structure of this work, then, is as follows. In Chapter 2 I present the general

[7]Really it's a very *old* way of thinking about ethics, of course. But against the background of Twentieth-century utilitarianism and deontology, it's effectively new. As Anscombe (1958) says, from the perspective of virtue ethics "the differences between the well-known English writers on moral philosophy from Sidgwick to the present day are of little importance."

aims of Brandom's systematic theoretical philosophy, some historical background against which to set his idiosyncratic conceptions of logic and inference, and the basic model of social practice—deontic scorekeeping—in terms of which the significance of logic is characterized.

The model of deontic scorekeeping is couched in terms of (among other things) a proprietary notion of *material inference*, which Brandom adopts, with modification, from Wilfrid Sellars. So in Chapter 3 I explore material inference. Beginning with a brief rehearsal of the significance of the form/matter distinction in the history of logic, I next present Sellars's conception of material inference, and its origin in his reflection on what a philosophical theory of meaning requires. I then present Brandom's descendant conception of material inference, and several reasons why the notion of material inference is philosophically attractive. With material inference thus explicated and motivated, I turn to the expressive conception of logic, and consider objections to that conception from the standpoint of traditional thought about logic.

Chapter 4 takes up the problem of psychologism. On my reading, Brandom's expressive conception of logic counts as a form of psychologism, and thus has both merits and challenges. The merits of psychologism are due to its relationship with philosophical naturalism, and so I present some background and motivation for naturalism, particularly as it manifests in Twentieth-century Anglophone philosophy. Because Brandom's expressive conception is *also* related to the expressivist program in Twentieth-century metaethics, itself a form of philosophical naturalism, I take some time to present what I see as the basic commitments of, and motivations for, metaethical expressivism. We then have a new vantage point from which to view Brandom's expressive conception of logic: as a substantive foundational theory of logic which is naturalistic, and which can be motivated for some of the same reasons which animate a prominent strand of metaethical thought. However, the psychologism which came to prominence in the late

Nineteenth century faced a number of damning objections, and as a result came to be regarded as a serious misstep for the philosophy of logic. So in the remainder of the chapter I take up several of the most serious objections to psychologism, as articulated in the writing of Frege and Husserl, and show that the form of psychologism manifested in Brandom's expressive conception doesn't make the mistakes made in the bad old days of psychologism.

Finally, in Chapter 5 I turn to the question of normativity: is logic normative for thought, and if so, in what ways, and why? On what is now the dominant conception of logic, logic is a kind of fact-stating discipline or science, and its normative import for thought, if any, comes via the general norm that true statements ought to be believed. The expressive conception opposes this view of logic, however, and in at least a couple ways hearkens back to traditional pre-Nineteenth-century conceptions of logic. So I articulate the traditional conception, showing how it differs from today's dominant conception in that it conceives logic as concerned with *acts* of thinking rather than the contents of thought, and with *norms* for thinking rather than facts about propositions or entailment. One way to ask about the normativity of logic is via so-called bridge principles: given some putative logical entailments, what *may* we do, or what *must* we do, about our beliefs? I approach the question of bridge principles by offering a novel account of what it means to ascribe validity of an argument on the expressive conception.

Chapter 2

Inference, Historical and Brandomian

2.1 introductory

Brandom's general philosophical ambition, as articulated in *Making It Explicit* and developed in *Articulating Reasons* and (to some extent) *Between Saying and Doing*, is to explain the conceptual contentfulness of thought and language in terms of the things thinkers and speakers do. In slogan form, the idea is to explain semantics in terms of pragmatics, or meaning in terms of use, or what we can say in terms of what we can do. In this introductory section, I will explain the basic outlines of Brandom's systematic philosophy, with the aim of making the slogans precise, so that the reader is prepared for the more scrupulous discussion in this and later chapters.[1]

There are many kinds of meaning in the world. We say that smoke in the distance means there's a fire; that *n* rings in a tree's cross-section means it is *n* years old; that a prairie dog's call means a hawk approaches; that the height of a mercury column means rain

[1] For a more abstract but still accessible introduction to Brandom's thought, which situates his views relative to others in logical space, see his own introduction to Brandom (2000).

is likely; that the pointy end of a sign means a destination is along this path rather than that one. Within the genus of such phenomena, the species of meaning which concerns Brandom is *conceptual contentfulness*—the kind of meaning carried by, paradigmatically, the claims expressible in human languages. The word 'claims' here is important; although human linguistic activity takes many forms, including joking, asking questions, making commands, telling stories, and so on, the primary home of the conceptual is in the things we say when we make claims or, to use the term Brandom prefers, *assertions*.[2] (See Figure 2.1 for a diagram of these species/genus relationships.)

Assertion is special among the varieties of human linguistic performance, in that when making assertions, we say *that* things are such-and-so. Our assertions are also *about* particular things (e.g. Benjamin Franklin), or collections of things (18th century inventors), or even fictional, abstract, or indefinite things (Captain Nemo, democracy, the geographical region subject to U.S. hegemony). These two features of assertion, expressed by the propositional '*that*' and the object-directed '*about*', are assertion's distinctive semantic features. To the extent that other linguistic performances can share either or both features—jokes, for example, can be about particular things, such as George W. Bush—they inherit this capacity from assertion, or so Brandom argues. Assertion is thus semantically fundamental.

What explains assertion's special semantic status is, according to Brandom, its special *pragmatic* role. There is a special thing we do when we make assertions: we take up a position in an *inferential social practice*. Assertions are the sorts of things that can stand in need of reasons, and can provide reasons for other assertions. Which is to say that assertions can play the role of premises or conclusions in inferences, a role which other linguistic performances cannot. But this inferential role is also essentially

[2] I am primarily going to stick with the term 'assertion', although the reader should be aware that 'claims', 'judgments', and 'propositions' will sometimes be used interchangeably, especially in the discussion of Frege, who moves freely between such terms (or their German equivalents). Where a distinction between these terms is consequential, it will be explicitly flagged.

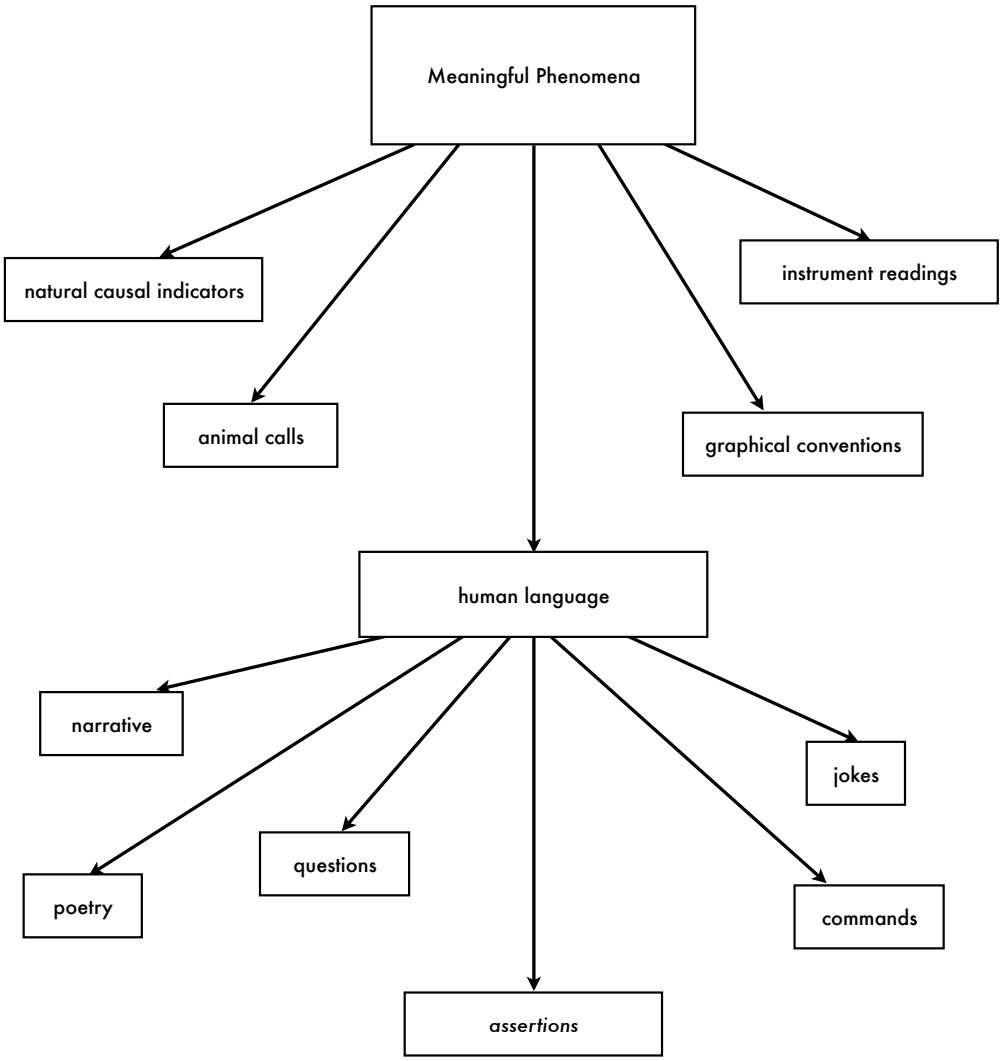

Figure 2.1: An Incomplete Taxonomy of Meaningful Things

socially articulated: making an assertion is something that, once done, a second party can demand a reason for, i.e., demand a premise which justifies the assertion as its conclusion. Likewise, a second party can draw out a *consequence* of an assertion, and hold the speaker liable for that consequence.[3]

Importantly, this social inferential practice is normative. The pragmatic significance of an assertion is not that, e.g., second parties *do* challenge it, but that they *may*. A speaker who makes an assertion undertakes a special sort of responsibility for it, which among other things entitles others to challenge it, incurs an obligation to respond to such challenges as are appropriate, and carries commitment to endorse its consequences. Such responsibility is always to be understood in terms of what is *appropriate* or *inappropriate* for the speaker or interlocutors to do.

Thus Brandom's explanation of conceptual contentfulness—in the first place, the contentfulness of assertions—comes in terms of what he calls a *normative pragmatics*. It is ultimately these things we do, which have normative significance in an inferential social practice, that confer conceptual contentfulness on our utterances, and make us more than mere parrots trained to vocalize 'hello' when someone enters a room, or fancy speaker-equipped thermometers which produce 'hot!' when the temperature crests 90°F. The full model of normative pragmatics, which Brandom dubs *deontic scorekeeping*, is considerably finer-grained than the simple account of assertion recounted above, but that is a difference of degree, not of kind: at bottom, the meaningfulness of human linguistic expressions is a matter of their normative pragmatic significance.

Of course, it is not enough to explain how assertions, or any other group of expressions, come to have conceptual contentfulness in the abstract. We would also like to know what makes one assertion—say, PAISLEY IS ON THE CHAIR—have the particular

[3]Of course, members of this practice can learn to challenge their own assertions, and to think through the possible consequences thereof, but the ability to do so is parasitic on the more fundamental interpersonal practice, according to Brandom. (Although see Brandom's reply to Gibbard in *Reading Brandom* (Brandom 2010a) for a hint that he thinks a fully intrapersonal inferential practice is possible.)

conceptual content that *it* has, as distinguished from all other possible assertions.[4] This is where the inferential nature of the social practice comes in.

Because assertions can stand as premises and conclusions in inferences, we can distinguish them by their *inferential potential*. This idea, which has its roots in Frege's (1879/1967) *Begriffsschrift*, exploits the usually unremarkable fact that in inferences, different premises generally license different conclusions, and likewise, different conclusions require different premises for their justification. For example, PAISLEY IS ON THE CHAIR has SOMETHING IS ON THE CHAIR as a consequence, but PAISLEY IS ON THE ROOF does not. This is one difference between the inferential potentials of the two claims. However, further differences manifest in the presence of auxiliary premises. To continue the example:

PAISLEY IS ON THE CHAIR
PAISLEY IS A CAT

A CAT IS ON THE CHAIR

A different auxiliary premise yields a different conclusion:

PAISLEY IS ON THE CHAIR
PAISLEY IS A PATTERN

A PATTERN IS ON THE CHAIR

Neither conclusion would be licensed without the appropriate auxiliary premise. This is most saliently a result of the inferential potentials of those auxiliary claims—when added to the original premise, each permits the drawing of a new conclusion. But this is *also* a result of the original premise's inferential potential, as well as those of the conclusions.

[4] In what follows, I am going to use small caps to indicate assertion types, in order to distinguish them from the mentioning of mere words or strings, which will be indicated by the usual single-quote device. (The role of small caps is thus analogous to the role of the judgment stroke in Frege's Begriffsschrift.)

The original premise, in conjunction with these auxiliary premises, yields the respective conclusions, and this is true of the original premise even when such auxiliary premises aren't present (hence the term 'potential'). Something parallel is true, of course, of each conclusion.

Frege's insight in the *Begriffsschrift* was to use inferential potential to individuate the conceptual contents of claims:

> . . . [T]he contents of two judgments may differ in two ways: either the consequences derivable from the first, when it is combined with certain other judgments, always follow also from the second, when it is combined with the same judgments [and conversely], or this is not the case. The two propositions "The Greeks defeated the Persians at Plataea" and "The Persians were defeated by the Greeks at Plataea" differ in the first way. . . . Now I call that part of the content that is the *same* in both the *conceptual content*.[5]

Frege's two sample claims do differ, he allows, in some aspect of what could be called their content (we might say, generically, their *meaning*). For example, they might differ in whether they emphasize the prowess of the Greeks or the ineptitude of the Persians. However, with respect to their specifically *conceptual* content, they are identical; they have the same inferential potential.

Brandom takes over Frege's insight, and adds a further theoretical commitment about the direction of explanation: what explains *why* an assertion has the particular conceptual content it has, as distinguished from all other possible assertions, is its inferential potential.[6] This explanatory direction is the reverse of the traditional philosophical view, which explains inferential potential in terms of conceptual content. (We will see more fully across the next two sections just how Brandom's project inverts the traditional

[5] Frege (1879/1967, p. 12)

[6] Although Brandom also attributes this further commitment to Frege, it seems to me that in this passage of the *Begriffsschrift* Frege is simply silent on the direction of explanation. His two claims in the passage appear to be (1) that sameness of inferential potential is sameness of conceptual content, and (2) that the Begriffsschrift will permit the explicit and unambiguous expression of conceptual content. Holding these two claims is compatible with also holding that conceptual content is *not* explained by inferential potential, but rather vice versa (i.e., in the traditional way).

order.)

Thus Brandom's account of the conceptual content of particular claims comes in the form of an *inferential semantics*. In order to understand, on his account, why a claim means what it does, we look to the inferences that the claim can figure in. This holds likewise for the meanings of subsentential terms, like 'California' or 'is bankrupt': we look to the contributions such terms make to the inferential potentials of claims which contain them—and not, notably, to the things which the terms represent or stand for.[7]

However, in order to understand Brandom's views, it is important to remember that the inferences in which claims can figure are things we *do*, and in particular that they are performances within the normative practice outlined above. It is only because the inferences articulate commitment and entitlement relations that the claims can be meaningful at all. One way to think about the relation between Brandom's normative pragmatics and his inferential semantics, then, is that the normative pragmatics tells us the character a performance must have to count as conceptually contentful in the first place—it must license a difference to the normative statuses in the social practice of deontic scorekeeping—and then the inferential semantics tells us, given that a performance is conceptually contentful, which content it has—the inferential potential of that particular performance.

[7]Because the inferential potential of a claim is a matter of how it behaves in all possible inferences, including those with arbitrary auxiliary premises, conceptual content turns out to be radically holistic: a claim's meaning depends on the meanings of all other possible claims. (And because subsentential terms inherit their conceptual content from that of the claims which contain them, holism obtains at the level of subsentential terms as well.) Meaning holism faces some well-known difficulties, including the fact that it seems to violate the widely held requirement that any semantic theory be compositional. More abstractly, holism threatens to make the very idea of individual meaningful claims (or terms) incoherent, since if a claim's meaning depends on every other's, and each of *these* depends in turn on the rest, then any claim means all of them at once. Only the totality of possible claims, considered as a whole, seems to be a stable unit of meaning. (Cf. Quine's (1951, p. 39) remark that " even in taking the statement as unit we have drawn our grid too finely. The unit of empirical significance is the whole of science.") For a sustained discussion of many such difficulties for holism, see Fodor and Lepore (1992).

I make no attempt to solve these difficulties here, since my target is not Brandom's systematic philosophy as a whole, but only the expressive conception of logic; nor am I attempting to defend holism.

It is against this general philosophical background that the expressive conception of logic is situated. To sum up that background, and situate *it* within a larger space of philosophical views, we can characterize it as:[8]

Concerned with conceptual, as opposed to nonconceptual content. Although there are many things we ordinarily speak of as meaningful, it is specifically conceptual content which is of primary interest, and at which the general philosophical explanation is aimed. Assertions, which are both *about* things and say *that* they are such-and-so, provide the paradigm case of conceptual content, and other linguistic performances are to be understood in terms of the account given for assertion. But it is not only non-assertional linguistic performances which are theoretically marginalized—the sort of experiential content philosophers have discussed under the rubric of 'nonconceptual content' is also of secondary concern. To put the point in terms of a related distinction, Brandom's primary philosophical concern is with what it is to be sapient, as opposed to sentient: to be aware of things not simply by having sensations of various kinds, but by having the *knowledge* that things are such-and-so.

Inferential, as opposed to referential. Given that it is specifically conceptual content that is at issue, there are two major ways to approach the explanation of that content. One can take the basic explanatory notion to be either inference or reference. The former adverts to the rational relationships between linguistic items, in particular sentences,

[8]The following is heavily abridged from Brandom's own introduction to his (2000). I have rehearsed only a few of the distinctions Brandom himself makes, in an attempt to cover only those which are preliminary to understanding the expressive conception of logic. Thus I do not discuss the distinction between discursive and nondiscursive practices, between construing the conceptual as assimilated to or differentiated from other mental phenomena, between Platonism and pragmatism about meaning, between taking mental or linguistic phenomena to be primary, between representation or expression as the key to sapience, between the intensional or the inferential as the mark of the conceptual, between traditional or rationalist expressivism, or—because this is really one major task of the dissertation as a whole—between logic as fundamentally epistemic or expressive in its function. Readers interested in those distinctions should consult Brandom's introduction.

assertions, judgments, or claims. The latter adverts to semantic relationships between linguistic items and worldly objects, in particular the relation of a name to its bearer, or a predicate to the property it represents. Brandom cleaves to the former.

Holistic, as opposed to atomistic. The conceptual content of any particular meaningful item might be explained in either holistic or atomistic fashion. This distinction can be illustrated in a number of ways: could there be a language which had only one predicate? could there be a thinker with only one concept? The atomist says 'yes' to these questions, holding that however it is that meaning is conferred on conceptually contentful items, such meaning is conferred upon each item independently of the others. The holist denies this, holding instead that the content of each item is dependent upon the content of all other such items. Brandom's approach is decidedly holistic, since the notion of an assertion's inferential potential relies on what consequences that assertion can have when conjoined with any number of other assertions, and these other assertions can be taken arbitrarily given the expressive resources of the language.

Top-down, as opposed to bottom-up. To a traditional way of thinking, the varieties of conceptually contentful items form a hierarchy: at the bottom are the subsentential items like names and predicates, in the middle are the sentential items—sentences, statements, assertions, or judgments—formed from these subsentential items, and at the top are inferences linking two or more sentential items. A traditional way of explaining the *correctness* conditions for each level is to start at the bottom and explain each higher level in terms of the lower, explaining correct inference in terms of correct sentences, and correct sentences in terms of correct subsentential parts. This traditional direction of explanation could thus be called "bottom-up." Brandom's approach is just the reverse, hence "top-down." But this characterization of Brandom's approach, though accurate, is quite abstract, and deserves more thorough exposition. So the next section explains

the traditional view in greater detail, and then the following section explains Brandom's inversion of the traditional view.

2.2 inference

The philosophical history of inference is at once uniform and multifarious. Although there is a common core of ideas about inference that originates in Aristotle and persists into the 19th and even 20th centuries, the treatment of inference—a thing that we persons can *do*—is always related to (and generally overshadowed by) the treatment of argument, logic, and rationality; and in these respects the philosophical history of inference is anything but unanimous. In this section I will recount some of the historical theses about inference which form the common core of what we can call the "traditional account" of inference, and then articulate the traditional account's explanation of inferential correctness, so that in the next section I can situate Brandom's conception of inference against this tradition.

2.2.1 the traditional account of inference

We should begin with a thin definition of inference, so as not to build substantive theoretical commitments into the definition, and to allow ourselves to construe philosophers (or psychologists) with different conceptions of inference as talking about one and the same phenomenon. Inference, then, is something we *do*, it both relies on and produces *beliefs*, and it is somehow intimately related to the social practice of *argument and disputation*, as well as the academic discipline of *logic*.

Inference, thus thinly defined, has a remarkably steady philosophical history. In outline, the traditional view of inference among philosophers has been that it is a mental act which connects two or more *judgments*. Judgments, in turn, are mental acts which connect (exactly) two *ideas* or *concepts*. These last are the mental items by which we

minded beings relate to the individual objects and properties that constitute the world around us.

Aristotle inaugurates this history. In *On the Soul* iii.6 he distinguishes two fundamental operations of the mind [*nous*]: the thinking of *simple* objects of thought, and the combination of those objects into a "quasi-unity."[9] Thinking of simple objects is the thought of things where questions of truth or falsity are inapplicable, because the simples are not the kinds of thing that can be true or false. Thinking of incommensurability, for example, is not thinking of something that can be true or false, nor is thinking of diagonality. Thinking *that* the diagonal of a square is incommensurable with its sides, however, or (to put it elliptically) *that the diagonal is incommensurable*, is thinking something that is capable of truth or falsity. In Aristotle's terminology this is a "combination" or "synthesis" of the simple objects of thought.

This distinction between the two operations of mind is maintained, and indeed explicitly relied upon, in *On Interpretation*, in which Aristotle calls nouns and verbs the linguistic expressions of mental simples, and sentences the linguistic expressions of mental syntheses.[10] With nouns we express our experiences of boats and steeds and other persons; with verbs we express our experiences of the things "said of or present in" other things.[11] Nouns and verbs are semantically simple, in that they have no independently meaningful parts. Sentences, on the other hand, do have independently meaningful parts—namely, nouns and verbs—and Aristotle is particularly interested in *propositions*, the class of sentences which admit of truth and falsity.[12] Propositions express *judgments*, which, given the parallel drawn between mental and linguistic structures, are naturally

[9] *DA* iii.6, 430^a27–430^b5 (Barnes 1984).

[10] "Just as some thoughts in the soul are neither true nor false while some are necessarily one or the other, so also with spoken sounds. For falsity and truth have to do with combination and separation. Thus names and verbs by themselves—for instance 'man' or 'white' when nothing further is added—are like the thoughts that are without combination and separation . . . ." *DI* i.1, 16^a10–16^a18 (Barnes 1984)

[11] Presumably a reference to the *Categories* i.2 (Barnes 1984).

[12] As opposed to sentences which don't admit of truth or falsity—Aristotle's example of such a sentence is a prayer, but we could include questions, commands, jokes, and so on.

identified as the mental syntheses Aristotle has introduced in *On the Soul* iii.6.

These two linguistic categories—the subsentential terms and the propositions—are discussed in the first two treatises of Aristotle's *Organon*: first the *Categories* concerns the subsentential terms, and then *On Interpretation* concerns propositions. The treatises immediately following are the *Analytics*, which concern the arrangement of propositions into the *syllogism*, a kind of reasoning or speech [*logos*] which lays out one proposition as following from two others. If we think of linguistic phenomena on the model of physical structures, this is a natural order in which to treat these topics: first the simplest items—the subsentential terms—then the propositions which are formed from those terms, and finally the reasoning which can be formed from propositions.[13] However, although Aristotle identifies mental operations corresponding to subsentential terms and propositions, he does not explicitly identify inference as a third mental operation, corresponding to syllogism. In his discussion of syllogisms in the *Analytics*, Aristotle speaks freely of what we may infer from various premises (or what we may not so infer),[14] but there is no explicit distinction between the mental and linguistic aspects of the syllogism, and inference is entirely absent from the discussion of the psyche in *On the Soul*, where one would naturally expect to find it alongside the thinking of simples, and the thinking of judgments.

By the time of Aristotle's medieval commentators, inference is recognized to be a third operation of the mind. In his commentary on *On Interpretation*, Thomas Aquinas says

> There is a twofold operation of the intellect, as the Philosopher says in III *De Anima*. One is the understanding of simple objects, that is, the operation by which the intellect apprehends just the essence of a thing alone; the other is the operation of composing and dividing. There is also a third operation, that of reasoning, by which reason proceeds from what is known to the

[13]Indeed, the naturalness of this order is probably why Aristotle's ancient editors arranged the *Organon* in this way.

[14]See especially *APr* ii (Barnes 1984).

> investigation of things that are unknown. The first of these operations is ordered to the second, for there cannot be composition and division unless things have already been apprehended simply. The second, in turn, is ordered to the third, for clearly we must proceed from some known truth to which the intellect assents in order to have certitude about something not yet known.[15]

Thomas thus makes explicit the underlying rationale for the arrangement of the *Organon*, and draws out the obvious corollary for Aristotelian psychology: inference is a third operation of the mind, depending on the second as the second depends on the first.[16]

In the early modern period, Arnauld and Nicole's popular *Logic or the Art of Thinking* ("the Port-Royal *Logic*") retains these three operations of the mind (and includes one more):

> Logic is the art of conducting reason well in knowing things, as much to instruct ourselves about them as to instruct others.
>
> This art consists in reflections that have been made on the four principal operations of the mind: *conceiving*, *judging*, *reasoning*, and *ordering*.[17]

Conceiving is the forming of *ideas* to represent experiences of simples; judging is the bringing together of ideas to affirm or deny one idea of the other. Reasoning, then, is "the action of the mind in which it forms a judgment from several others . . . ."[18] Arnauld and Nicole, in their enumeration of mental operations, thus explicitly count inference as the mental operation by which a person arrives at new beliefs on the basis of others. However, in the portion of the Port-Royal *Logic* devoted to reasoning, they take no greater pains than does Aristotle to distinguish that mental operation from its linguistic expression in the syllogism. (Although perhaps, given the instructional aims of the *Logic*, the authors don't consider the distinction necessary at this point in the text.)

---

[15](Oesterle 1962)

[16]"And since the three operations of reason are ordered to each other so are the books: the *Categories* to *On Interpretation*, and *On Interpretation* to the *Prior Analytics* and the books following it [the *Posterior Analytics*, *Topics*, and *Sophistical Refutations*]." (Oesterle 1962)

[17](Arnauld and Nicole, 1683/1996, p. 23)

[18]*Ibid.* Arnauld and Nicole's *ordering*, also called *method*, is the large-scale organization of a whole body of knowledge on some subject, arranged so as to facilitate learning.

In every traditional account of inference, very little if anything is said about what inference, considered as a kind of mental act, really *is*. Instead, the lion's share of philosophical thought is devoted to inferences' linguistic expressions, particularly the syllogism and its numerous subspecies. Inference *qua* mental activity is either conflated with its linguistic expression, or it is presumed to be some kind of syllogism done in the head.[19] What the traditional philosophers appear to have been most concerned with is not any analysis of inference, nor any candidate mechanisms which produce it, but rather an account of what makes *arguments* good. This account of good argument can then be read backwards into the mental as an account of inferential correctness. It is to this account that we now turn.

2.2.2 traditional correctness for arguments and inferences

There is a traditional account of inferential correctness which runs parallel to the structure of the traditional operations of the intellect. That is, first correctness is defined for the subsentential items, which can then be used to define correctness for sentential items, which can in turn be used to define correctness for inferences. This is typically done on the linguistic side of the mind/language distinction, via correctness conditions for *terms*, *statements*, and *arguments*.[20]

Terms come in two varieties, singular and general. Singular terms, or names,

[19]This latter view is articulated by Arnauld and Nicole, who hold logic to be a *post hoc* reflection on things our minds already do, and who claim that "because we can make our thoughts known to others only by accompanying them with external signs, and since this habit is so strong that even when we think to ourselves, things are presented to the mind only in the words in which we usually clothe them in speaking to others, logic must examine how ideas are joined to words and words to ideas" (Arnauld and Nicole, 1683/1996, pp. 23–4).

Cf. Hume in the *Enquiry*, who says, of the claim that inductive generalizations are the result of inferences, that "[t]here is required a medium, which may enable the mind to draw such an inference, if indeed it be drawn by reasoning and argument" (Hume 1748/1999, p. 114). Hume's demand that an inference requires a medium appears to come from Aristotelian logic, in which every well-formed syllogism has a *middle term* common to both premises. (Note also Hume's casual use of 'inference', 'reasoning', and 'argument'.)

[20]In what follows I will use a contemporary idiom to explain the traditional account, but nothing substantial should turn on the use of that idiom.

purport to represent individual objects; general terms, or predicates, purport to represent properties that objects can have, or relations that objects can stand in. When singular terms represent correctly, they *refer to* or *designate* their purported objects. When predicates represent correctly, they refer to their purported properties or relations.[21] Both kinds of term can fail to refer, as with names like 'Santa Claus' or predicates like 'is a unicorn'. So the mark of representational correctness for terms is reference. However, this correctness is taken as primitive on the traditional account; there is no attempt to explain how successful reference happens.[22]

Given this account of correctness for terms, there is a traditional account of correctness for statements. A statement is correct just when the object designated by its name possesses the property designated by its predicate (or, in the case of relational predicates, when the objects designated by names stand in the relation designated by the predicate). Statements that meet this standard of correctness are *true*; those that don't are false.[23]

Finally, given this account of correctness for statements, there is a traditional account of correctness for arguments. An argument is correct just when the truth of its

[21]There is an issue here about whether to construe predicates as referring terms. Certainly not all philosophers have thought of predicates, on the model of names, as referring; some have thought of predicates as *characterizing* or *describing*. See, for example, Strawson (1959).

[22]Within naturalistic 20th-century philosophy of mind and language there are many attempts to explain correct reference in terms of putatively nonsemantic properties such as information, biological function, causal history, and brain structure (or combinations thereof). (See, for example, Dretske (1981), Fodor (1987), Churchland (1989), Millikan (1984, 1989), Papineau (1987), among others.) These can be seen as attempts to naturalize the framework already present in the traditional view.

The fact that successful reference is primitive on the traditional account can *also* be discerned in Hartry Field's essay "Tarski's Theory of Truth" (1972), in which he points out that although Tarski's famous definitions for truth in formalized languages do tell us how to compute the truth-values of compound sentences from the truth-values of atomic sentences, those definitions do *not* tell us what it is for an atomic sentence to be true—and hence in an important sense do not constitute a "theory of truth"—because what makes atomic sentences true is treated as primitive.

[23]Cf. Aristotle's celebrated characterization of truth from the *Metaphysics*:"To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, and of what is not that it is not, is true; so that he who says of anything that it is, or that it is not, will say either what is true or what is false . . . " (*Met.* iv.7, 1011^b23–1012^a17, trans. Ross (Barnes 1984)). This is precisely the passage which Tarski cites approvingly in his famous (1933) paper "The Concept of Truth in Formalized Languages", bringing Aristotle's ancient and homely definition all the way into 20th-century model theory.

premises requires the truth of its conclusion. One typical way of putting this condition is to say that an argument is *valid* just in case whenever the premises are true, so is the conclusion. However, the same general idea appears in no less than (Tarski 1936), under the guise of logical consequence: "The sentence X follows *logically* from the sentences of the class K if and only if every model of the class K is also a model of the sentence X" (p. 417).

Although the traditional account of inferential correctness has just been presented in terms of linguistic items, it can also be taken as an account of the correctness of mental items. Terms correspond to concepts, statements correspond to judgments, and arguments correspond to inferences. The foregoing account of correctness applies, *mutatis mutandis*, to these corresponding mental items: concepts are deployed correctly when they successfully represent objects or properties, judgments are correct when they correctly represent objects as falling under the deployed concepts, and inferences are correct when they lead from true judgments to further true judgments.

2.3 Brandom's inversion of the traditional account

In its broad outlines, Brandom's conception of inference is identical to the traditional conception: inferences are doings that relate sentential items which are themselves composed of subsentential items. However, as mentioned briefly in §1, the order of understanding is reversed: inference is taken as the basic notion in terms of which the others are understood. For Brandom, the sentential items—on the linguistic side, assertions, and on the mental side, judgments[24]—are essentially what appear in inferences, and the meanings of such items are in turn to be understood in terms of the inferences they participate in. For example, the meaning of

[24] In the official Brandomean idiom, the term for such sentential items is 'doxastic commitments'; this will be introduced in §4.

(1) California is in the United States

is partly due to its supporting the further claims

(2) California is on Earth
(3) California is north of Guatemala
(4) California has a state legislature

but also partly due to its being supported by the claims

(5) San Diego is in the United States
(6) California is a member of a federation that includes Wisconsin, Michigan, and Missouri

and so on. It is *because* claim (1) participates in inferential relationships that it counts as an item at the sentential level, and it is because it participates in *these* particular inferential relationships—supporting (2)–(4), supported by (5)–(6), and all the other such claims—that it has the particular meaning it does.

Likewise, the subsentential items—concepts or terms—are essentially things that make systematic contributions to the inferential potential of the judgments or assertions they occur in. 'California', for example, means what it does in part because its occurrence makes the inferences from (1) to (4) and from (5) to (1) licit, as well as any inferences between claims which substitute 'California' for terms like 'the 31st state admitted to the Union' and 'the most populous U.S. state in 2015'.

As with the traditional account, Brandom's conception of inferential *correctness* runs parallel to his conception of inference, and is thus the reverse of the traditional account of such correctness. Instead of building up a notion of inferential correctness from subsentential and then sentential correctness, Brandom treats inferential correctness as the primitive notion, and then defines sentential and subsentential correctness in terms of it. The full story will only unfold across Chapters 3–5, but the basic idea is already here: the foregoing account of the meanings of sentential and subsentential items presupposes

that the inferential relationships in question be *correct*, for it is the correct inferences that determine what a claim means, and thus what its subsentential constituents mean.[25] (If the account permitted bad inferences to determine the meanings of claims, then every claim could participate in any arbitrary inference, and there would be no distinguishing claims according to inferential potential.)[26]

However, this strategy depends on an unusual notion of *material* inferential correctness, which (once again) contrasts with a feature of the traditional view. On the traditional view, inferential correctness is construed as a formal property. Aristotelian logic, for example, has it that the correctness of the inference

EVERY TRIANGLE HAS TWO RIGHT ANGLES
EVERYTHING WITH TWO RIGHT ANGLES HAS 180°
———————————————
EVERY TRIANGLE HAS 180°

is due to its having the logical form *Barbara*:[27]

EVERY *X* IS *Y*
EVERY *Y* IS *Z*
———————————
EVERY *X* IS *Z*

In contemporary logical language, we would instead use a quantified version of the logical form *hypothetical syllogism*:

[25]The full story of how subsentential items come to correctly represent objects and properties takes up the entire second half of *Making It Explicit*. I will give an abridged exposition of this story in Chapter 4, where it will be important to understand how Brandom is able to distinguish between merely *taking* an inference to be correct and its actually *being* correct.

[26]Cf. the deviant connective *tonk* in (Prior 1960).

[27]The name 'Barbara' is due to Aristotle's medieval commentators, who assigned names to syllogistic forms based on the type of categorical statements involved. Each categorical statement receives a letter designation ('Every *S* is *P*' = *A*; 'No *S* is *P*' = *E*; 'Some *S* is *P*' = *I*; 'Some *S* is not *P*' = *O*), and so every possible syllogistic form can be represented as a triple of vowels (*AAA*, *EAE*, etc.). These triples were then given mnemonic names to match the vowels: 'Barbara' for *AAA*, 'Celarent' for *EAE*, and so on.

$$\frac{\forall x(Fx \to Gx) \quad \forall x(Gx \to Hx)}{\forall x(Fx \to Hx)}$$

But no matter which logical language we are using, the traditional account construes the correctness of any particular correct inference as resulting from its being an instance of a correct logical form. That is, each claim in the inference is taken to be a combination of logical and nonlogical terms; the nonlogical terms are treated as instantiations of variables; the arrangement of logical terms and variables determines each claim's logical form; the arrangement of the claims' logical forms in the inference determines *its* overall logical form; and then if that overall logical form is a good one—say, if it is deductively valid—the original inference is correct. The task of the academic subject of logic, then, is to investigate not individual inferences but *logical forms*, whose goodness (or badness) is inherited by the inferences of which they are instances. On the traditional view, then, the correctness of inferences is a derivative matter, and the correctness so derived is formal correctness.

For a number of reasons, the conception of inferential correctness Brandom opts for is material, rather than formal.[28] This means, at a first pass, that the correctness of an inference is essentially a matter of which *nonlogical* terms it contains. (Hence the name 'material correctness': on the traditional view, the logical terms articulate an inference's form, and the nonlogical terms provide its matter.) For example, the inference

$$\frac{\text{THIS FIGURE IS A TRIANGLE}}{\text{THIS FIGURE HAS } 180^\circ}$$

[28] A detailed engagement with Brandom's reasons for relying on material correctness can be found in the next chapter.

is correct, but not because it is an instance of *Barbara* or any other logical form. Instead, its correctness is due, roughly, to what it's about—triangles and their properties. The nonlogical terms are essential to the correctness of this inference, as can be seen by the fact that some uniform substitutions of nonlogical terms will yield an incorrect inference:

$$\frac{\text{THIS FIGURE IS A STATUE}}{\text{THIS FIGURE IS MADE OF BRONZE}}$$

(Contrast the formally correct inferences above, in which uniform substitution of nonlogical vocabulary preserves inferential correctness.)

There is a sense, however, in which it's not right to say that the correctness of a material inference is "due to what it's about," or "due to the presence of its nonlogical terms," because these ways of glossing material correctness suggest that inferential correctness is a function of subsentential correctness, and hence posterior to it in the order of explanation. But that is the traditional direction of understanding inferential correctness, which Brandom is determined to invert. So there must be a way of understanding material correctness which doesn't make surreptitious appeal to subsentential correctness, and indeed there is; it is to be found in the model of deontic scorekeeping.

2.4 deontic scorekeeping: a first pass

We are now in a position to pick up the promissory note from the end of §1, to see how Brandom's inferential semantics is integrated with his normative pragmatics via the model of deontic scorekeeping, the inferential social practice responsible for the conceptual contentfulness of our speech and thought.

Deontic scorekeeping can be thought of as a game in which the players, whom we can for obvious reasons call *scorekeepers*, attribute a set of statuses to themselves and

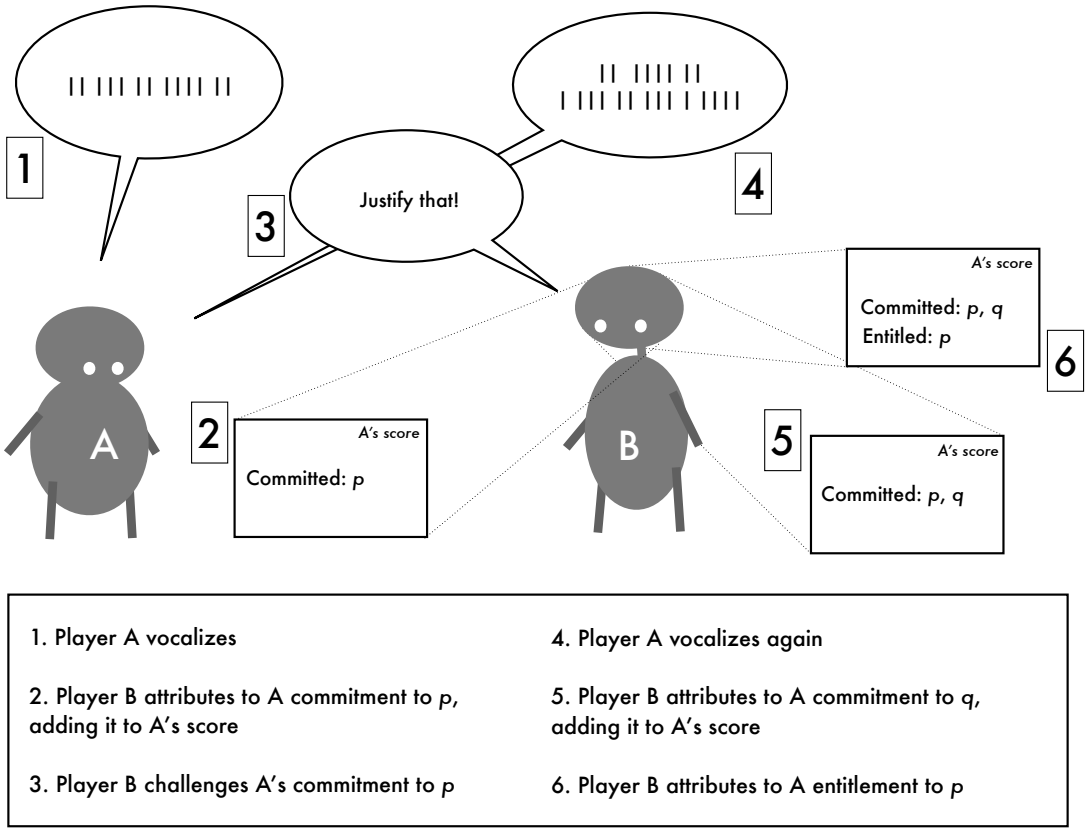

Figure 2.2: A Simple Episode of Scorekeeping (from Player B's Perspective)

one another. These statuses constitute each player's score (hence 'scorekeeping'), and each status is a normative status (hence 'deontic'). A simple illustration is provided by a two-player game in which one player makes an assertion:

At the beginning of this game, both players' scores are empty. When Player A makes a certain noise in Player B's presence, Player B updates A's score by attributing to her the deontic status of *commitment to p*—which we can think of, proleptically, as A's believing the proposition *that p*. Acquiring this status is a normative pragmatic change for A, because it changes what she is responsible for, and how others may appropriately treat her—in the scorekeeping game, it changes the permissible moves available to each

player. For example, B may now question whether A's commitment to p is legitimate, and A is responsible for vindicating such a challenge. If A vindicates the challenge, say by making more noise which B takes to be a further commitment q, B can update A's score to include the deontic status of *entitlement to* p.

Entitlement need not be attributed as it is in the example, however—namely, only when a scorekeeper has provided explicit justification. Player B might attribute to A the status of entitlement to p even without A's justifying performance, perhaps because B himself takes p to be true, or because of a policy to award entitlement to all assertions by default.[29]

Likewise, both entitlement and commitment can be attributed *consequentially*, on the basis of existing entitlements or commitments, without explicit performance by the interlocutor. To continue the illustration above, suppose B takes commitment to p to carry commitment to r (perhaps p = "California is to the left of Nevada" and r = "Nevada is to the right of California"). B can then update A's score to include commitment to r, even though A hasn't asserted it. Similarly, if B takes entitlement to p to carry entitlement to r, B can update A's score accordingly. And B can even update A's score to include entitlement to still other claims, including not only those which A hasn't asserted, but even those to which B doesn't count A as committed. (Take, for example, some further claim s such as "A randomly selected voter from Nevada probably opposes gun control" which B takes to inherit entitlement from entitlement to p, but commitment to which hasn't been earned by A, even as an inheritance from other commitments A has.)

In any such updating of score by inheritance of normative status—attributing entitlement to p on the basis of commitment to q, or commitment to r on the basis of

[29]Indeed, in Brandom's full model of scorekeeping, all assertions come with entitlement by default, and scorekeepers can lose that entitlement only by having it challenged, or by possessing other commitments incompatible with the asserted claim. See (Brandom 1994), pp. 176–178.

commitment to p, etc.—player B is not only altering player A's deontic score, but also making and endorsing an *inference*. B takes it that q justifies p, hence vindicates the challenge, and thereby takes it that the inference from q to p is a good one. B takes it that r is required by p, and thereby the inference from p to r is a good one. In this way, the normative pragmatics of deontic scorekeeping—the attribution of deontic statuses, which attributions alter the deontic score and thereby license further moves in the game—underwrites the inferential semantics outlined in §1, by making possible the moves which have the significance of inferences.

We can see now how A's initial vocalization satisfies the conditions, outlined in §1, for a performance to count as an assertion: it is an act which has the pragmatic significance of changing the deontic score, in a way which incurs responsibility on A's behalf and permits interlocutors like B to hold A responsible, for example by challenging A's entitlement to the claim. And the normative status incurred by the performance is one that stands in inferential relationships of commitment and entitlement to other such statuses, in just the way that items with the significance of sentences need to do.

The final point to make in this section is how the model of deontic scorekeeping underwrites an explanation of material-inferential correctness. But first I want to make an explicit disclaimer: the illustration of the deontic scorekeeping model I've just given is simplified, and the simplification is broadly along two dimensions.

First, it is coarse-grained; the full model of deontic scorekeeping comprises much more than just the moves which correspond to simple assertions and the attribution of individual normative statuses. For example, one important kind of scorekeeping move is the creation of a *socially hybrid* deontic status—the attribution of knowledge, to take one example of this kind, is a move in which a scorekeeper both attributes commitment and entitlement to a claim, *and* undertakes commitment to that claim himself. And similarly, although there is a crucial role in the full model for *anaphoric* connections

between claims—for example, between one scorekeeper's assertion of "That's a horse" and another's of "It's a cow on a dark night," where 'it' is anaphoric on 'that'—I've said nothing here about how anaphora is construed in scorekeeping terms. Nor have I said anything about the deontic status of *incompatibility*, which is directly relevant to the expressive conception of logic.

The second way in which this illustration is simplified is that it makes use of straightforward descriptions of what player B does, for example that B challenges A's assertion of p. In the official vocabulary by which the model of deontic scorekeeping is characterized, scorekeepers do not simply *do* things with normative significance. Rather, they do things which other scorekeepers *take* to have normative significance. This is why in the illustration above, I have not portrayed player A as simply asserting p, but instead described A's performance in nonnormative, nonsemantic terms—as a vocalization—and depicted it in the diagram with non-script symbols. It is B's scorekeeping that takes A's performance to have the normative significance it does, and hence why the propositional variable 'p' appears only on B's scoreboard. But by the same token, we shouldn't simply describe B as *challenging A's assertion*, since the challenge is itself a move with normative significance, and officially should be described as *a performance which is correctly taken by another scorekeeper to have the significance of a challenge*.

These are, I think, acceptable simplifications, for the purposes of this dissertation and at this stage in the exposition of Brandom's views. For one, insofar as my goal is to explore the expressive conception of logic, and not the entirety of the systematic theoretical philosophy of which it's a part, we need not have the full model of deontic scorekeeping, with all its involutions, in view. The core idea of the expressive conception can be explained with far less; the only thing we'll need beyond what the illustration already provides is the notion of incompatibility, to be covered in the next chapter.

As for the second dimension of simplification, it's usually just fine to talk straight-

forwardly of scorekeepers asserting claims and so on—after all, although the official theoretical idiom talks only at bottom of what scorekeepers *take* each other to do, the whole point of the theory is to account for homely phenomena like assertion. People do in fact assert things, and there's no harm in describing them as doing so. Moreover, a rigorous adherence to the official theoretical idiom makes for difficult reading, especially in early exposition. However, there is a real philosophical issue which can be obscured unless the official idiom is used: officially, scorekeepers do things which are correctly taken to have normative significance, but the only resource for explaining what this means—what it is to be correctly so taken—is just further scorekeeping. How are we to understand the difference between the significance a scorekeeper *in fact takes* a thing to have, and the significance it is *correctly* taken to have? Suppose a scorekeeper endorses a particular inferential move. According to that scorekeeper, of course, the move is correct, but what really makes it correct? Surely not that scorekeeper's taking it to be so, for that would entail a radical relativism of conceptual content to scorekeepers. And appealing to the perspective of some second scorekeeper who assesses the first seemingly won't help, for what makes the second's assessment correct? The problem recurs.

For now, I will leave this problem standing. It is a very general challenge to Brandom's systematic philosophical ambition, not specific to the expressive conception of logic; and it will return in Chapter 4, as a version of one of Frege's challenges to psychologism, so I will address it there. However, because we are here concerned with Brandom's conception of inference and its opposition to the traditional conception, and because that opposition requires a notion of material inferential goodness which is not cashed out in terms of an inference's subsentential components, and also because that notion of material inferential correctness is thus cashed out in terms of deontic scorekeeping, and hence what scorekeepers *take* to be correct, I do need to say something here about Brandom's general strategy for securing a notion of correctness by which

scorekeepers' doings are to be assessed which is independent of any particular individual's perspective.

The rough idea behind the strategy is to explain, internal to the practice of deontic scorekeeping, how scorekeepers are able to undertake normative statuses which defer authority to how things are in the world, *not* to another scorekeeper's perspective. However, that contrast—between how the world is and how another scorekeeper takes it to be—is itself drawn in scorekeeping terms. If the strategy is successful, there is a notion of perspective-independent correctness available to all scorekeepers, to which they can subject their own and each others' doings as liable. But this notion is instituted by their own practice, which is ultimately just a bunch of individual acts of taking things to be correct (or not) from individual perspectives.

So an inference, when it is materially good, satisfies the perspective-independent correctness conditions instituted by the practice of scorekeeping, which are not to be identified with any particular scorekeeper's taking that inference to be correct. Nonetheless, it is the individual takings—each scorekeeper permitting or requiring various updatings of score—which constitute the practice. This is, undoubtedly, a difficult idea, and I will devote much more space in Chapter 4 to making it intelligible. But what is important for the purposes of this section is just that there is a way to understand the notion of material inferential goodness which doesn't advert to the subsentential components of the claims involved. Instead, material inferential goodness is instituted by the social practice of scorekeeping, and the moves its practitioners permit and require.

2.5 looking ahead

The goal of this dissertation is to explore Brandom's expressive conception of logic. This chapter was a brief overview of Brandom's systematic theoretical philosophy,

focusing on its key elements: normative pragmatics, inferential semantics, the inferentialist order of explaining conceptual content, and the game of deontic scorekeeping, in which the elements are combined. The upcoming chapters will be dedicated to working out the expressive conception of logic in three main ways:

- An investigation of *material inference* and the conception of logic as material, rather than formal
- An assessment of the status of Brandom's expressive conception as a kind of *psychologism*
- An investigation of the *normativity* of logical rules and truths

Each chapter builds on the work of the previous chapter. It is thus to material inference, the master notion of Brandom's theoretical philosophy, that we turn next.

Chapter 3

Material Inference and the Expressive Conception of Logic

Now that the basic shape of Brandom's views on inference—both its place in the order of understanding, as well as the role for materially correct inference in deontic scorekeeping—is laid out, it is time to look more closely at what material inference is, and explain the expressive conception of logic built on it.

3.1 material inference

This section introduces the distinction between material and formal correctness, so that the expressive role of logical vocabulary can be introduced later. First there is a brief recap of matter vs form in inference, and then accounts of both Sellars's and Brandom's conceptions of material inference.

3.1.1 the doctrine of logic as form

The traditional paradigm of inferential correctness is *deductive validity*: an inference is good if its premises necessitate its conclusion, or perhaps if it's impossible for the premises to be true and the conclusion false.[1] And, as recounted in the last chapter, the usual way of cashing out this notion is in terms of the *form* of the inference, as opposed to its *matter*, what its sentences are about. For example, the validity of the Aristotelian syllogism

> All vertebrates are mortal
> All cats are vertebrates
> So, all cats are mortal

is not due to any particular facts about cats, vertebrates, or mortality, but instead that it has the form *Barbara*:

> All A are B
> All C are A
> So, all C are B

As far as validity is concerned, the items which fill the variable places are irrelevant—they constitute the matter of the inference, the stuff one happens to be reasoning about. It is form which determines validity, and inferences get their validity (if indeed they are valid) by being an instance of a valid form.

This general story about validity—what Etchemendy (1983) calls "the doctrine of logic as form"—remains unchanged even when one moves from Aristotelian to

[1] These are deliberately loose ways of characterizing validity, meant to capture the general notion which philosophers and logicians have sharpened over the last few thousand years. Both of these ways of characterizing validity are problematic from certain points of view today. (For example, "Pegasus is a horse" necessitates "Pegasus is an animal", but few philosophers today would say that the inference from the former to the latter is deductively valid. Similarly, the idea that a valid inference is one which cannot have true premises and false conclusion has been criticized by advocates of relevant logic.) However, the rationales for thinking these characterizations inadequate stem from theoretically motivated considerations, such as the conviction that validity ought to be due to form. These theoretically motivated considerations are posterior to the notion of validity itself, since they are part of a theoretical attempt to say more precisely what validity is.

contemporary logic.[2] The valid inference

> If it's raining, then the streets will be wet
> It's raining
> So, the streets will be wet

gets its validity not from saying anything true about rain or streets, but from being an instance of the form *modus ponens*:

> If P, then Q
> P
> So, Q

Here the variables stand for complete sentential expressions, rather than (as in Aristotelian logic) subsentential terms, but the account of validity remains the same: an inference which is an instance of a valid form inherits that validity, irrespective of its matter. Thus an inference with the very same matter but which is not an instance of a valid form, like

> It's raining
> So, the streets will be wet

is not valid. If deductive validity is the only kind of inferential correctness, then inferences like this will not even be correct, their intuitive appeal and ubiquity notwithstanding. On the other hand, if we want to ascribe to such inferences a kind of goodness that's not rooted in their form, we can consider the possibility of *material* inferential correctness. This would be a kind of correctness rooted in the matter of an inference, which is to say its content, the things it is about, or the meanings of its constituent concepts.

It is just this notion of material correctness to which Brandom adverts in his inversion of tradition. Because Brandom inherits this notion from Sellars but says very

[2]The way Etchemendy characterizes the doctrine is as the claim that "two sentences cannot differ logically if they do not also differ *formally* or *structurally*" (p. 320). But by 'logically' Etchemendy means to indicate an intuitive notion of entailment or consequence, so that according to the doctrine all of a sentence's "consequential" properties supervene on its formal properties. Thus according to the doctrine there can be no such thing as material consequences. Etchemendy would put it, in his terms, as the point that there can be no non-formal logical consequences. But we shouldn't follow him in using 'logic' and 'logical' in this way.

little about it himself, I will say a bit about Sellars's treatment first, and then turn to Brandom's. (Detailed discussion of Brandom's conception of material inference will have to be postponed until §4, because it is in the context of deontic scorekeeping that that conception is fully developed.)

However, before proceeding to discussion of Sellars, there is one last point which must be made about form and correctness: there is a close correspondence between an argument's form and *rules for inferring*. Every valid form can be formulated as a rule, along the lines of "inferences which conform to this form are permitted." For example, the form of *modus ponens* can be formulated as the rule "Q may be inferred from P and $P \to Q$." Indeed, forms and rules are so closely related that names like 'modus ponens' are used to refer to the form and the rule indiscriminately.[3]

Sellars's discussion of material inferential correctness proceeds in terms of rules. If there is indeed such a thing as material correctness, then there are material rules for expressing that correctness, saying that such-and-such materially correct inference is permitted. Of course, material rules could not be expressed formally, since they do not rely on inferential form.[4] But they would stand to materially correct inferences as formal rules stand to formally correct inferences, and their expression would make essential use of the inferences' matter. For example, if the inference

[3]It is important here to distinguish the rules which concern Sellars from the "inference rules" in a formal proof system. The former are rules for what a thinker may or must do when performing the cognitive act of inferring; the latter are rules for syntactic manipulation of formulas within the activity of doing a proof. Although these two kinds of rules might well be related, they are conceptually distinct. So, for example, when I claim that for every valid form there is a corresponding rule for inferring, I am *not* claiming that for every valid form there is a proof of its conclusion from its premises. That would be a standard kind of completeness claim for a formal logic—that every semantic consequence can be captured by the proof system—and we know that there are incomplete logics. But the correspondence between forms and rules should not be taken to mean that all logics are complete.

[4]Perhaps a more careful way to put this point is that material rules cannot be expressed *entirely* formally, and may have some formal and some material components. For example, the rule " 'x is in North America' may be inferred from 'x is in the United States' " is plausibly a material rule, even though it has a schematic formality on account of the variable place. However, identifying such a rule as material rather than formal requires identifying inferential form in general with specifically *logical* form, and although this identification is somewhat traditional, accounting for it requires introducing a distinction between logical and nonlogical vocabulary, which will not be done until §3.

It's raining
So, the streets will be wet

is materially correct, then the corresponding rule would be " 'the streets will be wet' may be inferred from 'it's raining'."

We are now ready to turn to Sellars's discussion of material inferential correctness.

3.1.2 Sellars on material correctness

Sellars (1953) argues for the importance of material rules of inference in the context of considering what a theory of *meaning* requires. This is in fact exactly the purpose for which Brandom needs the notion of materially correct inference; recall from the last chapter's discussion of deontic scorekeeping in §1.3 that it is correctness of inference which determines the meanings of items at the sentential and subsentential levels. In the following I will lay out the background for Sellars's discussion, his conception of material rules, and finally his argument for the claim that material rules have an authority for inference that is not derived from formal rules.

setting the stage for material rules

It's widely agreed, Sellars thinks, that a theory of meaning requires the sort of formal rules of inference embodied in classical logic. But this position has to have something to say about inferences which are not valid in terms of those rules—are such inferences good? The traditional answer given, Sellars thinks, is that such inferences are not good. Where there are intuitively compelling inferences which are not ratified by formal rules, these must be supplemented in order to be good, typically by a conditional which links premise(s) to conclusion. The only genuine rules of inference are formal rules, and all such inferential goodness as there is is formal goodness. On this view, then, intuitively compelling but formally invalid inferences are to be treated as

enthymemes: inferences which aren't good as formulated, but which can be made good by the appropriate supplementation.

Sellars has a counterproposal: suppose that there are not only formal but material rules of inference, which license inferences based on the particular concepts that are constitutive of the premise and conclusion. We then have the question whether these rules should be given a role in the theory of meaning, and if so, how they stand relative to the formal rules. According to Sellars, the possible views on this question can be stated in terms of authority: material rules either have or lack authority, and what authority they do have can be either derived from formal rules, or underived. His considered position is that material rules do have authority, that this authority is underived, and hence that the theory of meaning must appeal to material rules as well as formal rules. (The argument for this position is given an abbreviated formulation in §2.2.3.)

Sellars's conception of material rules situates them in the following structure: there are first of all *inferences*, which are mental or psychological activities that move from claim to claim. Inferences are liable to evaluation. Then there are the *rules* by which inferences are evaluated. And finally there are two kinds of relevant linguistic expression: *arguments*, which use particles like 'so' to indicate that an inference is being made, and *subjunctive conditionals*—of the form 'were it the case that . . . then it would be the case that . . . '—which express the rules which would, if correct, ratify the inferences which move from antecedent to consequent. So here, then, are the primary features of material rules on Sellars's conception of them.

- Material rules are normative, and in particular they license inferences and the arguments which express them,[5]
- the reasoning thus licensed is content-based,

[5]To be precise, material rules license arguments only to the extent that those arguments express rational support for a conclusion. Arguments, *qua* linguistic performances, are of course subject to norms of other kinds, e.g. norms of etiquette regarding when it is or isn't acceptable to make an argument.

- the rules themselves are expressed by subjunctive conditionals, and
- these rules are essential to meaning (whether mental or linguistic).

One thing to keep in mind as Sellars's argument unfolds is that material rules, if they exist, are independent of any particular linguistic expression of them. Although he thinks they are expressed by subjunctive conditionals, there need be no explicit such expression in order for a rule to obtain; the rules are in force even if not expressed. Sellars's own way of framing the discussion—in terms of whether "there are material rules"—perhaps obscures this point. However, I take it that whether "there is a rule" to ϕ is just the matter of whether in fact one ought to ϕ.[6] And so in the case of material rules, the question is whether we ought in fact to make inferences which are materially good, and if so whether the obligation is derivative from obligations specifiable in formal terms—that is, without reference to content.

what counts as a *material* rule?

In Sellars's exposition, the category of formal rules for inference is introduced ostensively: formal rules are familiar rules from logic textbooks like *modus ponens*. So the distinction between material and formal inference is first made against the background of whatever the general understanding of formality was in the 1950s—probably schematicity, following Quine, abstraction from semantic content, following Tarski, or topic-neutrality, following Ryle.[7] But while Sellars leaves his conception of formal rules at this more or less commonsense level of detail, this meager exposition invites possible misreadings of his position. It will be important in what follows to head off these misreadings.

[6]Hence the predicate 'is a rule' functions as a pleonastic judgment-forming operator, in a manner analogous to the judgment stroke in Frege's Begriffsschrift. "...[T]he subject contains the whole content, and the predicate serves only to turn the content into a judgment" (Frege 1879/1967, p. 12).

[7]See Quine (1950/1982), Tarski (1936), Ryle (1954).

As mentioned in above, Sellars conceives of inferential rules as normative for reasoning, as essential to meaning, and as expressible by subjunctive conditionals. What makes formal rules different from material rules is simply that material rules turn essentially on the matter of sentences, rather than their form. How is the venerable form/matter distinction to be understood here? Given the attention to enthymemes and their supplementation by conditionals, it's tempting to think that form for Sellars is a sentential matter: formal inferences turn on sentential connectives like the conditional, while material inferences turn on subsentential content.

But this isn't quite right. If we take his examples as a clue, Sellars includes under the rubric of 'formal' not only *modus ponens* and the rule of conjunction–elimination, but also cousins of these rules which depend on the first-order predicate calculus, for example "$\forall x(\phi x \supset \psi x)$, but ϕa, therefore ψa." These rules require their components to have the right sort of subsentential configurations in order to be correct, so Sellars must think that formality is present within sentences, so to speak, as well as between them.

Now it might be that the hallmark of material content is *being about the empirical world*. Much of the later part of "Inference and Meaning" is devoted to showing that mastery of material rules is required in order for our speech to get a grip on the things presented to us in experience, with the result that empirical conceptual content necessarily has inferential antecedents and consequences. And Sellars's own examples of material inferences tend to concern empirical causes: rain causing wet streets, lightning causing thunder, gravity causing chalk to fall. If this is right, then the way to understand Sellars's rather opaque metaphor—material rules have a role in meaning by "contributing the architectural detail of its structure within the flying buttresses of logical form"[8]—is in terms of empirical content. Formal rules are necessary for meaningfulness at all; material rules are necessary for meaningful judgments about the world.

[8](Sellars 1953, p. 317)

But this reading can't be right, either. Sellars frames much of his discussion as a response to Carnap, and often uses Carnap's own terms in making his points. What Sellars first distinguishes as 'formal' and 'material' rules are later called by Carnap's 'L-rules' and 'P-rules', and the illustration of these terms is as follows. Where

$\forall x(\phi x \supset \psi x)$, but ϕa, therefore ψa

is an L-rule,

ϕa, therefore ψa

is a P-rule. Thus the difference between the two is simply the presence of the generalized conditional (the very statement which appeared earlier as the enthymeme's supplementation), and there is no restriction on the type of predicates which may be taken as values of ϕ and ψ. In particular, there is no requirement that these predicates express empirical concepts, and so an inference like "Seven is prime, so it has no factors other than itself and one" would be a material inference. This means that Sellars's conception of formality is something like Ryle's topic-neutrality—formal rules aren't about anything in particular, but can be instantiated by sentences which are about particular things. Such authority as formal rules have comes from the ways that sentences can be related in abstraction from anything they happen to be about. And thus, by complementation, material rules get their authority from the content of the sentences involved.

Aside: 'material' in logical writings. Crucially, material rules as Sellars conceives them *do* have rational authority—given some material rules, it is rationally correct to make inferences in accordance with them. So we should not confuse Sellars's notion of materially correct inference with other things called "material" in the neighborhood of logic: statements formed by the truth-functional *material conditional* connective, the relation Tarski called *material consequence*, and the sort of inference which would

naturally correspond to either. A material conditional is true just when it has false antecedent or true consequent. A material consequence in Tarski's sense holds between a class of sentences K and a sentence X just when either K contains at least one false sentence, or X is true.[9] And so there is a notion of inference corresponding to both of these, which we might naturally call "materially correct" if we weren't careful to keep Sellars's terminology separate: such inferences would be correct just in case they were to have at least one false premise, or a true conclusion.

But this sort of inference is not what Sellars would call materially correct. As is well-known, the truth conditions for the material conditional have no restriction on the relevance of antecedent to consequent—"if roses are red, then water is denser than air" is a true material conditional. Likewise, 'water is denser than air' is a Tarskian material consequence of {'roses are red'}. But "roses are red, so water is denser than air" is *not*, as Sellars uses the term, a materially correct inference; it is not rational to make on the basis of the content of the claims. In general, inferences from arbitrary premises to arbitrary conclusions will not be materially correct in Sellars's sense, even if they happen to have false premises or true conclusions—there must in addition be rational connection between the contents of the premise(s) and conclusion.

why material rules are underived (and hence necessary for the theory of meaning)

For Sellars, the rationale for material rules of inference is given within an approach from the philosophy of language and mind: what sorts of rules do we need to countenance in order to explain the meaningfulness of thought and talk? Given the necessity of formal rules (more on which in a moment), Sellars has a convincing argument to the effect that material rules are also needed. Here's how that argument goes:

If formal rules but not material rules were needed, we ought to be able to express

[9] See Tarski (1936, p. 419).

any purported material rule by the use of formal rules alone. The most obvious way to attempt this is by construing alleged expressions of material rules as enthymemes, which when properly supplemented yield inferences whose goodness is formal. Sellars thinks this strategy fails: either we fail to reconstruct the force of the original material inference, or we reconstruct it only by smuggling in the rule of inference we started with. For example, suppose we supplement the enthymeme like so:[10]

(A) Every time it rains the streets are wet.
So, if it were to rain the streets would be wet.

If the logical rule at work here is Universal Instantiation, and if the supplementary premise is a material conditional, then the inference would be symbolized as:

(A_1) $\forall x(rain(x) \rightarrow streets(x))$
So, $rain(a) \rightarrow streets(a)$

But this symbolization makes it clear that supplementation by material conditional is inadequate: UI merely instantiates variables, leaving sentential connectives unchanged. So if the inference is valid, then the conclusion must likewise be a material conditional, and the original inference, expressed as a subjunctive, has not been captured. This failure suggests, then, that we supplement the enthymeme with an entailment statement (read 'entails' for the turnstile $\vdash$):

(A_2) $\forall x(rain(x) \vdash streets(x))$
So, $rain(a) \vdash streets(a)$

Now the conclusion is expressively equivalent to the initial inference—as Sellars puts it, the entailment statement "conveys the same information as" what the subjunctive says.

[10]This argument and the ones that follow are after Sellars (1953, p. 324).

But this accomplishment has been achieved only by employing a general entailment statement, which is effectively the material rule of inference from rain to wet streets. Hence one supplementation (A_1) is austere but inadequate, and the other (A_2) is adequate but makes use of a material rule.

Sellars tries out one final supplementation (the '$\Box\!\!\rightarrow$' expresses the subjunctive conditional):

(B) $[\forall x(rain(x) \rightarrow streets(x)) \wedge rain(a)] \Box\!\!\rightarrow wet(a)$

This formulation has two problems. First, what were in previous formulations premises for inference rules ('$\forall x(rain(x) \rightarrow streets(x))$' and '$rain(a)$') now occur in the antecedent of a conditional. This means that they are "unasserted," and hence this formulation fails to express the original inference "It's raining, so the streets will be wet," where '$rain(a)$' is asserted. Second, if formulating things this way is a general strategy to make formally good inferences out of subjunctively-expressed enthymemes, it will do so for even false subjunctives, i.e. bad inferences.

Sellars takes these three failures to show that material rules can't be captured in terms of formal rules, and hence that there must be a device available *somewhere* for the expression of material rules:

> ...[E]ven though material subjunctive *conditionals* may be dispensable, ... it may nevertheless be the case that the *function* performed in natural languages by material subjunctive conditionals is indispensable, so that if it is not performed in the object language by subjunctive conditionals, it must be performed by giving direct expression to material rules of inferences in the meta-language [as in formulation (A_2)]. In other words, where the object language does not permit us to say "If *a* were ϕ, it would be ψ" we can achieve the same purpose by saying " 'ψa' may be inferred from 'ϕa'."[11]

Thus both material and formal rules are necessary for the meaningfulness of thought and language, even if material rules are expressed as normative statements in a metalanguage,

[11](Sellars 1953, p. 326), emphasis added.

rather than subjunctive conditionals in the object language. The necessity of formal rules is not argued for; this is a view Sellars takes his interlocutors to share, so he spends no time motivating it. And since material rules can't be derived from formal-rules-plus-material-premises, their normative authority for reasoning is likewise underived.

Hence for Sellars, material and formal rules stand alongside each other, having equal rational status. This is a view that Brandom, upon taking up Sellars's notion of material rules, revises, holding instead that formal rules have their authority derivative upon material rules.

3.1.3 Brandom on material inference

Brandom's exposition in *Making It Explicit* appears to simply endorse wholesale Sellars's own argument to the effect that material rules are essential to meaning. But the dialectic runs differently in Brandom's work; one salient difference is that Brandom has much more to say about what inference is, and how it's related to rules for inference. So Brandom's discussion proceeds in terms of *material inference*, and that discussion has subsequent consequences for what he thinks about rules for inference. In this section I will sketch the major features of Brandom's conception of material inference; a more detailed account must wait until §4.

Here is a definition, which will require a bit of unpacking:

> Material inference is a response to one normative status attribution with another status attribution, where these statuses are related to claims, the second is rationally "on account of" the first, and this correctness turns essentially on the material content of the claims.[12]

Material inference is a response to one normative status attribution with another: According to Brandom, what rational subjects like us are able to do—for each other as

[12]This definition (which is mine, not a quotation from Brandom) omits complications which arise from Brandom's systematic theoretical philosophy, in particular the inferential moves he calls (following Sellars) *language-entry* and *language-exit* moves (which we might more naturally call *perception* and *action*).

well as themselves—is attribute the *normative statuses* of entitlement and commitment, which are analogous to the familiar statuses of, respectively, being permitted and being obligated. Such attributions can be made for any number of reasons; inference is the particular type of attribution which is made in response to a prior attribution.

These statuses are related to claims: A rational subject might be entitled or committed to many things, most familiarly to particular actions (such as entitlement to shoot a free throw), or action-types (such as commitment to fighting government corruption). In the special case of inference, subjects are entitled or committed to *claims*, where entitlement to P corresponds intuitively to being warranted in the belief that P, and commitment to P corresponds to having P entailed by one's beliefs.

The second attribution is "on account of" the first: Two sequential status attributions might be merely sequential, as when a speaker utters P, and later Q, and a hearer attributes commitment to P, then commitment to Q. Inference is not this sort of merely sequential attribution; the second attribution must be done because the attributor takes it to be appropriate given the first.

Inferential correctness turns essentially on the material content of the claims: The four conditions above are sufficient to define Brandom's conception of inference in general;[13] this is the condition which makes inference *material*. Brandom's way of accounting for the difference between formal and material content will be discussed below. One consequence of talking in terms of material *inference*, rather than material *rules* as Sellars does, is that we can speak of material inferences, as things scorekeepers do, being correct or incorrect, good or bad, appropriate or inappropriate.

[13] Again, omitting complications that arise from language-entry and -exit moves.

Unlike Sellars, Brandom does not presume any particular notion of form against which to contrast the material content of claims. At the outset, all content is material, and no particular subsentential terms—logical or otherwise—are picked out as articulating the form of an inference. So although both Sellars and Brandom are interested in the philosophical status of materially correct inference, and in the role such inference plays in a theory of meaning, the question which occupies Sellars—whether material rules have any rational authority, and if so, whether it is derivative on formal rules—doesn't arise for Brandom. There is no antecedently intelligible way of picking out the formal rules, such that we could then inquire into the status of material rules relative to them.

Instead, Brandom thinks that formal rules are derivative on material rules, and has a recipe for so deriving them: beginning with a collection of good material inferences, we distinguish some vocabulary that occurs in them as *fixed* vocabulary, the rest being non-fixed. Then we pick some materially good inference, which we can call a *basis inference*. Now it follows that any basis inference which remains good under all substitutions of non-fixed for non-fixed vocabulary is an inference that's good in terms of its fixed vocabulary.[14] For example, if our collection of material inferences includes some color vocabulary, then we can pick out 'red', 'green', 'colored', and the like to be the fixed terms. Then if our basis inference is "This banana is green, so this banana is colored," all substitution-instances are materially good, and hence are good in terms of their color vocabulary. (If the basis inference were "This banana is green, so this banana is unripe," but the fixed terms remained the same, some substitution-instances would be materially bad, and so even though the basis inference is good, it is not good in terms of its color vocabulary.)

The recipe is familiar from similar procedures in Bolzano and Tarski, but differs in that it takes material-inferential goodness as primary, and in particular not as a species

[14] Assuming the additional restriction that substitutions don't mix up grammatical classes, so that e.g. singular terms don't get substituted for general ones.

of truth-preservation.[15] Nonetheless, Brandom's recipe, like theirs, yields a relativized notion of form: if a class of inferences is good in terms of its K-vocabulary, then we have identified a good K-form. The good form identified in the previous paragraph is the schema "x is green, so x is colored."

One important consequence of the relativity of form to vocabulary is that it obviates a certain type of worry about form, matter, and content. The way I introduced the traditional take on form and matter above, as it arises in the context of logic, was to say that formal inferential goodness in no way depends on the *content* of the claims involved, but only on their form. When form is introduced this way, as the complement of content, then it's tempting to think that the formal terms in an inference—whatever's left over once the content is pulled out—must somehow express no content. But surely the traditional logical terms express a kind of content, even if they do articulate the form of a statement, and so we appear to have a puzzle. What the relativized notion of form makes clear is that there is no in-principle problem with the formal terms in an inference also having content. In the above example, 'is green' and 'is colored' are the formal terms, but are clearly contentful; the situation is likewise for many other possible choices of fixed vocabulary.

However, another consequence of the relativized notion of form is that, if it is to underwrite an account of specifically *logical* consequence, then we will need to have a way of picking out the specifically logical terms to hold fixed, such that, when applying the substitutional method, we have a resulting class of inferences good in terms of its logical vocabulary. But because the class of logically good inferences will be a species of the formally-relative-to-K good inferences, the specifically logical vocabulary has to be demarcated without reference to form, on pain of circularity. So an independent criterion

[15] Because of concerns about limitations resulting from the language on which the substitutional procedure is conducted, Tarski (1936) rejects this particular procedure, opting instead for the similar but more powerful method of models (pp. 415–16). Tarski's concerns apply likewise to Brandom's substitutional procedure, and will be addressed in §7.4.

for logical vocabulary is needed.

3.2 logical vocabulary and deontic scorekeeping

The idea that there is a class of terms to be demarcated as the specially logical vocabulary is very old. In Aristotelian logic, there is a distinction between categorematic and syncategorematic terms; categorematic terms are those capable of playing the role of subject or predicate in a statement (i.e. they may take the place of schematic variables in the syllogistic forms above), while syncategorematic terms "bind" or connect the categorematic terms. Syncategorematic terms can be thought of as the special logical vocabulary; they include 'all', 'some', 'not', and the 'belongs to' verb Aristotle uses as a copula, as in "*A* belongs to every *B*".

Much more recently, there has been debate over the demarcation of the *logical constants*, which serve as the fixed terms in the substitutional definition of logical consequence (and in Tarski's more sophisticated model-theoretic definition, along with its descendants). There is general agreement that the logical constants include 'all', 'some', and 'not', as well as binary propositional connectives like 'if' and 'and', although there are plenty of disputed cases, along with disagreement about how to characterize even the agreed-upon cases within a formal language.[16]

What unites the syncategorematic terms and the logical constants, as two species of the genus of logical vocabulary, is that each is responsible for articulating the logical form of a statement. Hence one's demarcation of logical vocabulary has immediate consequences for which arguments are pronounced logically valid, and this has been the

[16]On the need for a demarcation of the logical constants, see Tarski (1936). Proposals for demarcations include Peacocke (1976), Hacking (1979), Tarski (1986), Sher (1991), McGee (1996), Warmbrod (1999), and Gomez-Torrente (2002). For arguments about how to formally construe the logical constants, see Routley, Meyer, Plumwood, & Brady (1983), Read (1988), Priest (2008), and Dummett (1973/1978), among others.

primary motivation for the 20th-century debate over demarcating the logical constants. Brandom's way of making the demarcation is both interestingly different from its rivals, and it also obviates some of the concern over getting the demarcation *right*, so as to pronounce the correct set of arguments logically valid. But in order to see these points, we must first understand Brandom's way of demarcating logical vocabulary.

3.2.1 'if' and 'not'

As the model of deontic scorekeeping has so far been described, scorekeepers can undertake and attribute deontic statuses, some instances of which count as making (and hence endorsing) inferences, but they have no distinguished way of communicating to other scorekeepers that they are doing so. This is where logical vocabulary comes in—it allows the expression, in the form of an assertion, of the underlying moves that scorekeepers can make.[17]

The conditional is a paradigm case. Without it, a scorekeeper can attribute commitment to some claim p to an interlocutor, and, taking that commitment to carry commitment to the further claim q, update the interlocutor's score to include commitment to q; but it is only by employing the conditional to say "If p, then q" that the scorekeeper can express that updating of score. Merely saying "q" shortly after the interlocutor has said "p" will not carry the same significance.[18]

[17]In MacFarlane's (2009) terminology, this makes Brandom's approach to logical vocabulary a "pragmatic" rather than an "analytical" demarcation of the logical constants. Analytical demarcations seek "to identify some favored property ... as a necessary and sufficient condition for an expression to be a logical constant." By contrast, pragmatic demarcations "start with a *job description* for logic and identify the constants as the expressions that are necessary to do that job." Brandom's approach is unusual among pragmatic demarcations, however, which generally hold logical vocabulary to be whatever is minimally required to systematize some discourse or theory—scientific theories or mathematical ones, for example. (See e.g. Warmbrod (1999), Shapiro (1991).) These pragmatic demarcations are thus guided by a principle of parsimony. But Brandom's approach, which specifies the particular expressive role of logical vocabulary, has no such parsimony consideration; *anything* which plays this role counts as logical vocabulary. So there is still a way in which Brandom's expressive conception counts as an analytical demarcation, since it provides a necessary and sufficient condition for logicality.

[18]Or at least, it isn't guaranteed to carry the same significance, because it doesn't have the right

This is most obvious if we bring in negation (the expressive significance of which will be explained shortly). Suppose the interlocutor has claimed p, and then q. Our scorekeeper takes this to be an argument on the interlocutor's behalf, to the effect that the interlocutor takes q to follow from p. And our scorekeeper also disagrees that q follows from p. How is this disagreement to be indicated? Saying "Not q" doesn't do the job, particularly if our scorekeeper thinks that q is true but simply doesn't follow from p. And saying simply "No" is ambiguous between (at least) denying q and denying the inference to q from p. It is only by employing the conditional, and saying "Not: if p then q" that our scorekeeper can explicitly register disagreement with the *inference* rather than with one or the other assertion.

Explaining negation's expressive significance requires elaborating the model of deontic scorekeeping. We have already seen that a scorekeeper can, in response to attributing some commitment p to an interlocutor, make a material inference by attributing a second deontic status: commitment or entitlement to another claim q. However, a scorekeeper can also, upon attributing commitment to p, *rescind* the interlocutor's entitlement to q. Doing so is taking p and q to be materially incompatible—another material-inferential relationship that can hold between deontic statuses. For example, commitment to "The Cure plays maudlin, mopey music" might preclude entitlement to "The Cure is an excellent band."

Material incompatibility allows scorekeepers to disagree with one another in advance of having the expressive resources of negation: to disagree with someone else's assertion of p, simply assert some q you take to be incompatible with it. However, there are a couple difficulties with doing so. The first is that for any p, there is a range of possible claims q which one might choose to assert, but each carries its own material-

expressive role. The interlocutor in this situation might still *take* the utterance of q to be the expression of an inference from p, and in some contexts, doing so might be correct (suppose, for example, that the two scorekeepers are working together on a problem, and thinking out loud about how to solve it).

inferential status. Both “cats are robots” and “cats are invertebrates” are incompatible with “cats are mammals,” but the significance of asserting the one is quite different from that of the other. If scorekeepers’ only means of disagreement is the assertion of incompatible claims, then they need to navigate the varied scorekeeping significances of those claims in order to select one for assertion.

The second difficulty with using materially incompatible claims to express one’s disagreement is that there is no guarantee that one’s interlocutor will also take the claims to be incompatible. For example, an interlocutor who believes that some mammals lack spines won’t take “cats are invertebrates” to be incompatible with “cats are mammals,” and hence won’t register the assertion of the former as disagreement with assertion of the latter. This phenomenon is endemic to material incompatibility; for any two materially incompatible claims we can always find some third which would make them compatible.

Adding negation to a language circumvents both these difficulties. By negating a claim—saying “Not p”—a scorekeeper can express p’s minimal incompatible: the unique claim which is compatible with everything incompatible with p. Thus “Not p” carries no extra material-inferential significance beyond its incompatibility with p, and is guaranteed to be taken as incompatible by any (well-behaved) interlocutor.

3.2.2 quantifiers, identity, and other connectives

By anyone’s accounting, there is more logical vocabulary than just the conditional and negation, and this is no less true on Brandom’s view. However, there is reason to single out ‘if’ and ‘not’ at first—they are paradigms of logical vocabulary, and their satisfaction of the general criterion by which Brandom demarcates logical vocabulary (that they express underlying deontic scorekeeping move-types, which types confer conceptual contentfulness on scorekeepers’ token performances) is reasonably easy to illustrate. But having done so, we might want to know how the general criterion admits

other traditionally logical terms, such as the universal and particular quantifiers, the identity relation, and sentential connectives other than 'if' and 'not'.

Explaining the scorekeeping significance of the quantifiers requires elaborating the deontic scorekeeping model to account for subsentential structure, in particular the categories of singular term and predicate. The key idea is that of a *substitution inference*, which is any single-premise inference in which premise and conclusion are related by the substitution of one expression for another.[19] The semantic significances of singular terms and predicates can be distinguished by the classes of materially good substitution inferences they figure in—while predicates can have asymmetric substitution-inferential potential, singular terms have only symmetric substitution-inferential potential. For example, in the substitution inference

STING PLAYED BASS

GORDON SUMNER PLAYED BASS

not only is the inference from premise to conclusion materially good, but so is the inference vice versa. This holds true for 'Sting' and 'Gordon Sumner' generally. But that's not the case with materially good substitution inferences in which the predicate is substituted for. The inference

STING PLAYED BASS

STING WAS A MUSICIAN

[19] See (Brandom 1994), chapter 6. In the official theoretical idiom, it is important to distinguish between two types of expression that occur in a substitution inference, and a third which is a result of discerning a substitutional pattern: *substituted-for*s, which are the expressions replacing or replaced by another; *substituted-in*s, which are the expressions remaining constant across substitutions, and *substitutional frames*, which are the schemata resulting from putting a variable in the place occupied by a substituted-for. These types are used to determine syntactic differences between substituends. (See (Brandom 1994, p. 368).) However, since my present purpose is only to explain the expressive role of the quantifiers, and doing so doesn't require keeping track of these three types, I'll avoid keeping track of them for the sake of simplicity.

is materially good, but isn't vice versa. Brandom's idea is that this is what distinguishes singular terms from predicates: "Predicate substitution inferences may be asymmetric, while singular-term inferences are always symmetric" (Brandom 1994, p. 372).

Once we have the distinction between singular terms and predicates in terms of substitution inferences, the expressive role of the quantifiers becomes clear. A universally quantified claim like "Everything is a musician" expresses commitment to all substitution instances of the form "*x* is a musician", while "Something is a musician" expresses commitment to some such substitution instance. These can of course be combined with other bits of logical vocabulary, for example in universally quantified conditionals like "Everything that plays bass is a musician"—claims which display the inferential relationships between predicates, and hence make (some of) the content expressed by those predicates explicit.[20]

The expressive role of identity is also now clear: identity claims express the endorsement of symmetrical substitution inferences involving the expressions on either side of the identity sign. So, for example, asserting "Sting is identical to Gordon Sumner" expresses endorsement, wherever a substitution inference which replaces 'Sting' for 'Gordon Sumner' is materially good, of the symmetrical inferences, and vice versa.[21] When combined with an account of demonstratives and their anaphoric dependents (which I will not recount here), identity claims also enable the expression of the *reidentification* of objects, as in "That one is the same as the one yesterday."[22]

[20] See §5 for more on this content-displaying power of logical vocabulary on the expressive conception.

[21] Hence this account of identity does not run aground in opaque contexts, since it is not tied to intersubstitutibility *simpliciter*. Asserting "Clark Kent is identical to Superman" doesn't express endorsement of

$$\frac{\text{LOIS LANE BELIEVES CLARK KENT IS HER COWORKER}}{\text{LOIS LANE BELIEVES SUPERMAN IS HER COWORKER}}$$

unless the converse is also taken to be materially good (in the paradigm case, because one also believes that Lois believes that Superman = Clark Kent).

[22] See chapter 7 of (Brandom 1994), particularly pp. 459–473.

So much for the quantifiers and identity. What about other sentential connectives traditionally taken to be logical, like conjunction and disjunction?

In asking and answering this question, I think it becomes clear just how unusual Brandom's approach is within the philosophy of logic. After all, it is well-known that the usual set of classical sentential connectives can be defined in terms of just one or two truth-functions, such as the Sheffer stroke (NAND), or the material conditional together with negation. Brandom devotes no particular attention to conjunction or disjunction in *Making It Explicit*, not even to say that they can be defined in terms of negation and the conditional, the two connectives for which he does account.[23] Why not?

The reason, as I see it, is that Brandom is not in the first place offering an account of any particular formal logical system, but rather of the logical vocabulary that occurs in natural languages and which is—genealogically and conceptually—prior to any attempt at formalizing a logical system. (This is one of the reasons materially good inference shouldn't be conflated with true material conditionals, or with Tarskian material consequence.) The logical vocabulary, on Brandom's account, is all and only those terms which express the underlying content-conferring scorekeeping moves. So it might well turn out that 'and', for example, just isn't a piece of logical vocabulary, if it doesn't play the right expressive role—even though many formal logical systems include conjunction as one of the connectives. Thus what we need, in order for conjunction and disjunction to qualify as logical vocabulary on Brandom's conception, is to find the right kind of expressive role for them to play. And I think it can be done, in both cases, by attending to a very general feature of inferential relationships; namely, that what follows from what depends on which collateral premises are available.

[23]There is a brief mention on pp.117–118 of MIE that conjunction can be given an inferential characterization: "to define the inferential role of an expression '&' of Boolean conjunction, one specifies that anyone who is committed to P , and committed to Q, is committed to $P\&Q$." But this is in the context of a discussion of Gentzen's approach to defining the logical constants, not in the context of Brandom's own logical theory.

Take conjunction first. In many cases, adducing additional premises increases the material-inferential power of a claim. For example,

THAT'S A RIBBON OF COLORED LIGHTS IN THE NIGHT SKY

I'M IN THE NORTHERN HEMISPHERE

THAT'S THE AURORA BOREALIS

is a materially good inference, but it wouldn't be if either premise were missing. However, sometimes adducing additional premises also decreases the material-inferential power of a claim. (Which is to say that material goodness is in general nonmonotonic.) If we add another premise to the previous inference, like so:

THAT'S A RIBBON OF COLORED LIGHTS IN THE NIGHT SKY

I'M IN THE NORTHERN HEMISPHERE

I'M UNDER THE EFFECTS OF A HALLUCINOGENIC DRUG

THAT'S THE AURORA BOREALIS

the resulting inference is no longer materially good. So in general, it matters for the material goodness of an inference just which premises are involved, and this feature is quite general in our production and assessment of material inferences, which is to say that it's a general feature of the practice of deontic scorekeeping. And we can make this feature explicit by using 'and' to indicate that we are appealing to two or more premises at once, and that such an appeal is crucial to the material goodness of the inference.[24]

What about disjunction? Unlike conjunction, it doesn't obviously correspond to

[24]In his response to Bernhard Weiss, who asks how conditionals should be used to express multipremise inferences, Brandom says "…I would say that making it possible to make explicit multipremise inferences is the principal expressive role characteristic of conjunction" (Brandom 2010b, p. 354). However, in order to show that conjunction qualifies as logical vocabulary, it must be shown that what is expressed is a feature of deontic scorekeeping which is necessary for the practice to confer content on its individual moves, and Brandom does not do this. I take myself here to have completed the required account for conjunction.

adducing more premises, even though in formal settings both conjunction and disjunction concatenate sentences. But I think we can take a clue from the symmetry of conjunction and disjunction in the sequent calculus, where the introduction rules for each connective are the same, except that where conjunction appears on the left-hand side of the turnstile, disjunction appears on the right, and vice versa. So instead of thinking about disjunction as in the first place something with a primary role among the premises, let's think about it as something with a primary role on the side of the conclusion:

THAT'S A RIBBON OF COLORED LIGHTS IN THE NIGHT SKY
I'M IN THE NORTHERN HEMISPHERE
I'M UNDER THE EFFECTS OF A HALLUCINOGENIC DRUG

THAT'S THE AURORA BOREALIS, OR A DRUG-INDUCED VISION

Now we once again have a materially good inference, which we got from *weakening* the conclusion by disjunction. In this way, we can think of disjunction as expressing the importance of multiple premises, and the nonmonotonicity of material inference generally, but as doing so within the conclusion. In this way, it is doing something like the job '*ceteris paribus*' does on the premise side of things—saying "*Ceteris paribus*, if *P* then *Q*" is a way of saying "Unless there are no defeaters, if *P* then *Q*," which is just to recognize that a material inference is nonmonotonic.[25]

On this way of thinking about things, conjunction and disjunction are both ways of expressing the same underlying feature of deontic scorekeeping, but on different sides of the inference. What about the other fourteen binary sentential connectives—do they have expressive roles of the same kind? In general, I should think not. The set of sixteen binary sentential connectives is an artifact of truth-functional logical theory, and not the sort of "natural" logic Brandom is in the first place accounting for. It might turn out that

[25] On the expressive role of *ceteris paribus* clauses, see (Brandom 2008, p.107).

some of the truth-functional connectives do a decent job at expressing something like what is expressed by the Brandomean logical vocabulary, but then again it might not. (Truth-functional 'and' seems pretty good, the truth-functional material conditional is poor, as we've already seen in §2.2.2, and there's no reason to expect that truth-functional connectives like

P	Q	$P \triangleleft Q$
T	T	T
T	F	T
F	T	F
F	F	F

P	Q	$P \triangleright Q$
T	T	T
T	F	F
F	T	T
F	F	F

which are redundant with P and with Q, respectively, will have any particular expressive role, even though from a truth-functional perspective they're as logical as conjunction.)

3.3 the priority of material inference

The above account of logical vocabulary in the context of deontic scorekeeping requires that material inference is prior to formal inference. But what are the reasons to think that material inference has this priority? As I count them, the reasons are: (1) Brandom endorses Sellars's argument in "Inference and Meaning" against construing material inferences as enthymemes, (2) formal correctness can be defined in terms of material correctness, but not vice versa, (3) an interpretive principle of Dennett's, according to which "imperfectly rational" creatures should be ascribed mastery of content-specific rules for inference, pushes in favor of the priority of material inference, and (4) the virtue of being able to explain logical competence in terms of more primitive abilities. Each of these reasons will be examined in turn, followed by a brief section on

what ‘priority’ means.

3.3.1 the argument against enthymemes

What does Sellars’s argument (recounted in §1.2 above) really establish? The first thing is that the usual idea of an enthymeme—an argument invalid as it stands, but valid with its suppressed premise included—is in some sense wrong, owing to the difficulty to be had in adequately supplying the suppressed premise. But the difficulty arises only because we want to be able to say validly exactly what we had been saying enthymemetically; and while this shows that we can’t simply eliminate material rules in favor of formal ones without loss of expressive power, it doesn’t show that we *can’t* do without material rules, or, for that matter, without formal rules. This point raises two natural questions.

First: why do we need formal rules? Their necessity was something assumed by Sellars from the outset, presumably because that view was shared by his empiricist interlocutors, but we might reasonably ask for some justification on this point. The closest Sellars comes to providing such justification in “Inference and Meaning” is in a passage in which he imagines what his interlocutors would say about formal rules:

> Formal rules of inference are essential to the very possibility of language; indeed, of thought. Kant was on the right track when he insisted that just as concepts are essentially (and not accidentally) items which can occur in judgments, so judgments (and, therefore, indirectly concepts) are essentially (and not accidentally) items which can occur in reasonings or arguments. Without formal rules of inference there would be no terms, no concepts, no language, no thought.[26]

It is not clear whether Sellars himself endorses this particular rationale for the necessity of formal rules, although it is reminiscent of arguments he provides elsewhere.[27] If he

[26](Sellars 1953, p. 314)

[27]For example, in the section on foundations for empirical knowledge in “Empiricism and the Philosophy

does endorse this argument, however, it seems like a bad one. The idea is that judgments are essentially the kinds of things that occur in reasoning, but this requires only that there be *some* rules or other which link judgments in reasoning. Sellars himself thinks we need material rules, but perhaps material rules are the only rules we need; in fact, once we've admitted material rules (on the grounds that they can't be reduced to formal ones), we might then ask whether formal rules can be reduced to material ones. If so, then material rules by themselves look sufficient.

Second: why do we need material rules? It's conceivable, at least, that we could have a language with only formal rules, and in place of material rules have corresponding general premises. Sellars's argument shows that this equipment won't yield the same expressive results as it would with the addition of material rules, but we might wonder whether, in constructing a theory of meaning, we could do just as well without material rules. Indeed, Sellars acknowledges that the special grammatical category of subjunctive conditionals—the paradigmatic expression of inference rules—could be optional; perhaps the expressive role of subjunctives is likewise dispensable.

What Sellars says to address this point is that "the function performed in natural languages by material subjunctive conditionals is indispensable" (Sellars 1953, p. 326). The function in question is the expression of inferential correctness, the indication that one claim may be inferred from another. What is the evidence that this function is indispensable?

One piece of evidence comes from the diagnosis Sellars makes of Carnap's program for artificial formal languages. Carnap attempts to eliminate subjunctive conditionals from the object language on the grounds that they are not extensionally tractable. Sellars argues that the attempt succeeds only because it displaces the role of asserting in-

of Mind", in which Sellars argues that an observation report can count as knowledge only if the observer has other bits of knowledge which warrant it; i.e. stand to it as premises to conclusion in an inference (Sellars 1956, §36).

ferential correctness from the object language to the metalanguage.[28] That is, in Carnap's system, it is metalanguage sentences like '$\psi a \vdash \phi a$' which say what may be inferred from what, and they *must* express this correctness if the whole system is to do its job. If they are seen as merely presenting some syntactic relation between items in a system of formal signs, then the system has no relevance for reasoning or argument.[29]

But if this diagnosis is correct, it works on the principle that rules for inference must be expressed somewhere in the system; it doesn't require that these rules be *material* rules. So the diagnosis needs to be conjoined with a second piece of evidence—the prior argument against enthymemes—in order to show that the function performed by specifically material subjunctives is indispensable.

I think the diagnosis of Carnap's system (and formal languages in general) is correct—a formal language needs to have its derivation rules be genuine rules for inference if it is to be used in reasoning or argument.[30] However, I am not sure about the more general position (which is not obviously Sellars's) that usability in reasoning or argument is the *sine qua non* of contentfulness in general. Nonetheless, we may not need to settle this matter here; the language we actually speak certainly does make use of subjunctive conditionals, and I believe Sellars has adequately shown—via the argument against enthymemes—that material rules are essential to such languages.

So what about the argument against enthymemes? I think Sellars is convincing on this point, but there are two further reasons to be skeptical of the idea that a material inference is good only when supplemented by its suppressed premise.

First, there is a crucial difficulty with the notion of an enthymeme.[31] Some examples will help show this. Consider first an enthymeme of the sort that Aristotle

[28] See the exposition of this argument in §1.2.3.

[29] Indeed, if it's the case that judgments are essentially things that occur in reasoning or argument, then on the merely syntactic reading of metalanguage derivation rules it follows that the formal system cannot express conceptual content at all.

[30] Cf. (Dummett 1959) on the use of truth-tables.

[31] The following point is made in (Kapitan 1982).

might have examined:

(E_1) Some birds are flightless.
So, some birds are not delicious.

If our theory of deductive inference is the syllogism, then we know exactly which premise is suppressed. The syllogism permits only four kinds of sentence, which may be arranged only in premise-premise-conclusion form, yielding 64 possible arrangements, only a subset of which are valid.[32] If we are to make (E_1) valid by supplying a suppressed premise, it will have to look like this:

(E_1*) *No flightless thing is delicious.*
Some birds are flightless.
So, some birds are not delicious.

Now that we moderns have left the syllogism behind, the paradigm of an enthymeme looks more like this:

(E_2) It's raining.
So, the streets will be wet.

And supplying the suppressed premise looks, again, something like this:

(E_2*) *If it's raining, the streets will be wet.*
It's raining.
So, the streets will be wet.

But even if we bracket Sellars's concerns about reconstruction (from section 1.2), there is still a problem. Unlike the syllogism's missing premise, the conditional premise here is not *required* to validate the inference; we can equally well validate it in this way:

(E_2**) *It's not raining, or whales are mammals, or the streets will be wet.*

Whales are not mammals.

It's raining.

So, the streets will be wet.

[32] In the traditional terminology, these 64 arrangements are moods. Each mood admits of multiple ways of arranging the terms in the syllogism, called figures. (There was dispute among logicians whether there were three figures or four.) Taking each figure to give rise to a different syllogism yields more than 64 syllogisms.)

Modern logic permits any number of premises of various forms, and as long as there is some route to the conclusion via legitimate inference rules, the inference is good. The lesson to draw from this is twofold. First, modern logic requires us to revise the classical notion of an enthymeme so that possibly multiple premises are suppressed, and no unique set is required. Second, the very idea of an enthymeme is relative to a background logical system—an argument might be enthymemetic relative to one logical system but not another. So, since any argument can be made valid according to some system, or by adding an arbitrary number of "suppressed" premises, the enthymeme strategy loses much of its motivation.

Second, the reasoning according to which good material inferences ought to be construed as enthymemes is exactly the reasoning that leads the Tortoise to generate a paradox for deduction in "What the Tortoise Said to Achilles".[33] If $\ulcorner P \Rightarrow Q \urcorner$ is an enthymeme, and is good only when reformulated as $\ulcorner P \wedge (P \rightarrow Q) \Rightarrow Q \urcorner$, then by parity of reasoning we ought to demand the same supplementation for the new formulation, requiring yet another suppressed premise, and so on to infinity. So insisting that inferential correctness requires the conditional premise amounts to an infinite deferral of that correctness.

3.3.2 formal inference as definitionally derivative

As recounted above (§1.3), it is possible to define a notion of formally correct inference via substitution of non-fixed vocabulary in materially correct inferences. It is not likewise possible to define formally correct inference in terms of materially correct inference. This asymmetry is a reason in favor of thinking that material inference is prior to formal inference, and indeed Brandom counts it as such (1994, p. 104).

[33](Carroll 1895)

3.3.3 Dennett's principle

In "Intentional Systems", Dennett offers the following principle for attributing mastery of inference rules to intentional systems:

> If we found an imperfectly rational creature whose allegiance to *modus ponens*, say, varied with the subject matter, we could characterize that by excluding *modus ponens* as a rule and ascribing in its stead a set of nonlogical inference rules covering the *modus ponens* step for each subject matter where the rule was followed.[34]

Brandom endorses Dennett's approach of attributing mastery of inference rules to creatures by interpreting their behaviors, but objects to Dennett's "formalism" (Brandom 1994, p. 101). Dennett's principle by default attributes mastery of fully formal logical rules like *modus ponens*, and retreats to attributing mastery of nonlogical (i.e. material) inference rules only when forced to do so by a creature's piecemeal, subject-specific aptitudes. Such a formalist approach to interpretation needs justification, but Dennett doesn't provide any, instead treating rationality as essentially a matter of logical competence.

Brandom quotes Dennett's principle as an illustration of a mistake—the unjustified adoption of a formalist interpretive principle—and then moves on to elaborating his own competing "materialist" stance. So he (Brandom) cannot be said to endorse the principle. Nevertheless, we can focus on Dennett's principle a bit longer than does Brandom, and see that it has a *prima facie* difficulty not faced by the materialist approach.

Suppose we adopt Dennett's principle, and use it to go about attributing mastery of inferential rules. How likely are we to find "imperfectly rational" creatures of the sort he refers to? Many, if not all, nonhuman animals will qualify. But it turns out that humans, too, have their allegiances to logical rules swayed by various subject matters, and by logically irrelevant features such as the way a problem is framed. Since groundbreaking psychological work in the '60s on reasoning (Wason) and in the '70s on judgment and

[34](Dennett 1971, p. 95), quoted in (Brandom 1994, pp. 100–101).

decision-making (Kahneman and Tverksy), a cottage industry in demonstrating human "irrationality"—namely, deviation from standard rules of logic and probability—has flourished.[35] Particularly relevant for Dennett's principle are the results from variations on Wason's Selection Task, which appear to show that humans' allegiance to *modus tollens* is highly dependent on the subject being reasoned about—we do very well at applying the rule when it concerns social norms, and very poorly otherwise. Indeed, the disparity is so striking that some psychologists (notably Cosmides and Tooby (1992)) have proposed that we have an evolved proprietary mental module for reasoning about social norms.

This fact about humans' reasoning abilities places a burden of plausibility on Dennett's formalist principle: why should we suppose that the default sort of rationality to attribute to a creature is mastery of fully formal logical rules, if even human beings—rational creatures if anything is—have subject-specific allegiances to those rules? It is not merely that formalism ought, as Brandom points out, to be justified. Formalism also faces an implausibility which materialism does not face, and so materialism—the idea that mastery of material inference rules is prior to master of formal logical inference rules—is in better standing.

3.3.4 explaining logical competence

The priority of material inference makes possible an explanation of competence with logical vocabulary in terms of more primitive abilities. That is, supposing we take creatures to be competent with at least some material inferences, we have a natural way to explain logical competence. For example, a creature can learn to use 'if P then Q' as appropriately deployed in circumstances where that creature already endorses the

[35] See (Wason 1966, 1968) and (Tversky and Kahneman 1974) for early presentations of key findings. For major criticisms, see (Gigerenzer 1991, 1996), (Cohen 1981).

inference from P to Q. And that endorsement, in turn, is just an ability to treat the move to Q from P as licensed in the game of deontic scorekeeping.[36]

On the other hand, to take formal inference to be prior is to eschew this sort of explanation. And, as Brandom says, "It is a short step from treating mastery of [logical] concepts as implicit in inferential abilities to treating it as an innate presupposition of them. This sort of thing gave the classical rationalists a bad name" (1994, pp. 101–102). If taking material inference to be prior avoids the nativist pitfall, and permits the explanation of logical competence, then the materialist approach is hugely attractive. (Which is not to say that Brandom's materialist approach doesn't take *some* things to be innate—basic organismal abilities to respond differentially to stimuli, for example. But crucially, none of these innate abilities is specifically conceptual. Young humans acquire mastery of concepts by being trained into the practice of scorekeeping.)

3.3.5 what priority amounts to

There are at least two senses in which material inference is, in Brandom's work, prior to formal inference. The first is priority "in the order of explanation," a sense which Brandom explicitly endorses.[37] When we take up the position of a theorist trying to understand and explain how creatures traffic in conceptual content, material inference has explanatory priority over formal (and hence over logical) inference. When we want to understand what it is for a creature to treat an inference as formally good, we explain it in terms of the creature's treatment of inferences as materially good.

But there is another sense in which Brandom accords material inference priority, although it is (I believe) implicit in his writing. This is a kind of *genealogical* priority, according to which competence with material inference emerges in the development of a

[36]The elaboration of logical abilities from mastery of material inferences is described more formally in chapter 2 of (Brandom 2008).

[37]See, for example, (1994 p. 134).

species before competence with formal inference does (and likewise for the development of an individual of a species in which formal competence has developed). This is the sense of priority expressed by Gilbert Ryle (1950) when he claims that statements employing conditionals are "sophistications upon arguments." The ability to use logical vocabulary can only be acquired by a creature which can *already* successfully treat some statements as following from others. But this latter ability is just the ability to treat some inferences as materially good.

3.4 the expressive conception of logic

The previous sections have clarified the expressive role of conditionals and negation, and the role of logical vocabulary as such: to express in the form of a claim the underlying and prior material inferences which scorekeepers can make, the making of which confers conceptual content on individual utterances. However, we still need to investigate the resulting conception of logic as such; the discipline or body of knowledge. Following Brandom, I'll refer to this as *the expressive conception of logic*.[38]

One terse way to characterize logic's status on the expressive conception is as "the organ of semantic self-consciousness" (Brandom 1994, p. xix). This characterization is intimately related to Brandom's reading of Sellars on "Socratic method:"

> In dealing with such situations [attempts to justify acceptance of lawlike statements such as *'Clouds of kind X cause rain'* by means of an argument from instances], philosophers usually speak of inductive arguments, of establishing laws by induction from instances. . . . I am highly dubious of this conception. I should be inclined to say that the use Jones will make of instances is rather in the nature of Socratic method. For Socratic method serves the purpose of making explicit the rules we have adopted for thought and action, and I shall be interpreting our judgments to the effect that *A*

[38] Brandom uses this label to refer to not just his own view, but also those of (early) Frege and Sellars, each of whom he sees as articulating, darkly, the view he himself holds. In my discussion I will reserve the name for Brandom's view, remaining agnostic about whether attributing the view to Frege and Sellars is apt.

> causally necessitates *B* as the expression of a rule governing our use of the terms '*A*' and '*B*'.[39]

Think of Socrates buttonholing Euthyphro, asking him "And what is piety, and what is impiety?"Euthyphro first offers instances of piety, but Socrates wants to elicit the "general idea" of piety, which Euthyphro then offers: "Piety, then, is that which is dear to the gods, and impiety is that which is not dear to them." This can be interpreted, à la Sellars, as a proposal for a rule about how to use 'pious' and 'dear to the gods', such as

> *Whenever you characterize something with 'pious', you may (/must) characterize it with 'dear to the gods'.*

In the idiom of deontic scorekeeping, the rule is about the appropriateness of making a particular material-inferential move:

> *Whenever you attribute or undertake a claim of the form 'x is pious' to a scorekeeper, you may (/must) attribute the claim 'x is dear to the gods' to that scorekeeper.*

The key to this conception of Socratic method is to see the generalizations it produces as expressing material rules for inference, which thereby articulate the content of subsentential concepts. When such a rule is expressed schematically, as above, the content-displaying role of logical vocabulary becomes particularly clear. Because applying '*x* is dear to the gods' is a materially appropriate consequence of applying '*x* is pious', it is part of the content of calling something pious. And by using the conditional to say 'If something is pious, then it is dear to the gods', we can make that part of the content visible, by making explicit what we are committing ourselves to when we call something pious.

Hence another of Brandom's ways of characterizing the expressive conception:

[39](Sellars 1949), quoted by Brandom in (1994, p. 105).

> . . . the philosophical significance of logic is not that it enables those who master the use of logical locutions to *prove* a special class of claims—that is, to entitle themselves to a class of commitments in a formally privileged fashion. The significance of logical vocabulary lies rather in what it lets those who master it *say*—the special class of claims it enables them to express.[40]

What is so special about the class of claims whose expression is made possible by logical vocabulary? The content-displaying function adumbrated above is one such special thing; ordinary claims which lack logical vocabulary can't perform this function.[41] Another special feature of logical claims, enabled by their content-displaying function, is their ability to underwrite the disputation and sharpening—what Brandom would call "grooming"—of concepts. Take Socrates again: after he has elicited an account of piety from Euthyphro, he proceeds to challenge that account by raising problematic cases. (For example: what should Euthyphro say about things which are dear to some but not all of the gods?) This disputation would be impossible without logical vocabulary. Once it becomes possible, we scorekeepers can take control of what we (take ourselves to) commit ourselves to when we employ concepts, by revising or rejecting the conditional claims we endorse.

These ways of characterizing the function of logic articulate the expressive conception more finely than the slogans above. But to understand something is on Brandom's view a matter of knowing what commitments it follows from, and what commitments follow from it; thus understanding is achieved by degrees as the web of such commitments is brought into view one filament at a time. The account of the expressive conception begun here is necessarily merely partial. The task of the rest of this work is to make visible more of the web surrounding the expressive conception of logic. In the rest of this chapter I take up some basic objections to the expressive conception, so

[40](1994, p. xix)

[41]This is bracketing the roles of modal and normative vocabulary, which on Brandom's view bear similarities to logical vocabulary (they are species of the same genus).

as to show that common commitments in the philosophy of logic which might appear to require its rejection are in fact quite compatible.

3.4.1 but isn't logic formal, not material?

As indicated above, there is a long tradition of thinking about logic as a discipline concerned with something formal. Kant (1787/1999, A54/B78) explicitly claims that logic "has to do with nothing but the mere form of thinking,"[42] and by the time Tarski wrote his famous paper on logical consequence in 1936, the idea that logic is specially concerned with form had become so commonplace that he was able to casually identify logical with formal consequence.[43] But Brandom's conception of logic takes material-inferential goodness to be primary, and logical-inferential goodness to be parasitic on this. Isn't this entirely out of step with traditional thought about logic?

The short answer is: no. In fact, there are two ways in which Brandom's conception of logic is seen as a natural companion of traditional conceptions, one relatively shallow and the other relatively deep.

The first way in which Brandom's conception is a companion of the tradition is that although he appeals to a novel notion of material-inferential goodness, the definition of deductive validity built on this is formal in exactly the way Bolzano's and Tarski's substitutional definitions are.[44] Form results from taking certain terms to be fixed in the substitution and others to be variable, and specifically logical form results from holding all and only the logical terms to be fixed. Logically valid forms are then those in which all instantiations of variable terms yield materially good inferences. Every step

[42] In Kant's taxonomy, this is logic as "pure general logic," abstracted from thought's relation to objects, and from any particular subject-matter.

[43] "...since we are concerned here with the concept of logical, i.e. *formal*, consequence ..." (Tarski 1936, p. 414)

[44] Of course, Tarski rejects the initial substitutional definition offered in (Tarski 1936), but not because it isn't formal—instead it's a concern for the limitations of the formal language in which the substitution is done. On that concern, see §7.3.

in this account of logical validity is the same as Bolzano-Tarski except the last—where they would rely on the truth-preservingness of the instances, Brandom instead relies on material-inferential goodness. But this difference in the basic notion of inferential goodness, though possibly problematic for independent reasons, doesn't make Brandom's account of validity any less formal.

The second, deeper, way in which Brandom's conception is a companion of the tradition is that the criterion for logicality, or rather for which vocabulary counts as logical vocabulary, is itself a matter of form. It is not, as Kant would have it, the form of thinking, i.e. the different forms of judgment the mind can make, but rather the form of the scorekeeping practice itself. As Brandom puts it in an early formulation, logical vocabulary "is distinguished by its function of expressing explicitly *within* a language the features of the use of that language that confer conceptual contents on the states, attitudes, performances, and expressions whose significances are governed by those practices" (1994, pp. xviii–xix). Most vocabulary in a language will not have this character. Even though it's true that, e.g., asserting something expresses an underlying scorekeeping commitment, what is expressed will in general *not* be a feature of the use of the language in virtue of which the particular performances, etc. are contentful. Asserting "the cat is on the mat" expresses something about a cat and a mat, not about content-conferring features of the practice of deontic scorekeeping. Only the use of logical vocabulary can do this.

Which features of scorekeeping, exactly, are those that "confer conceptual contents on the states, attitudes, performances, and expressions whose significances are governed by" the scorekeeping practice? All those features which can be thought of, I suggest, as the form of scorekeeping. For example, scorekeepers can make material-inferential updates to their own and each others' scores, and each such update will be from particular statuses to other particular statuses. But the expressive role of the

conditional abstracts from the particular statuses involved, and simply expresses the doing—the inferential update—independently of the objects of the doing. Thus we can think of inferential updating as a formal feature of the practice, and of the conditional as expressing that feature. Likewise, negation expresses the scorekeeping move-type of taking two claims to be incompatible, another form of a doing. If my accounts of the expressive roles of conjunction and disjunction above are correct, they express the taking of multiple claims to be essential for the goodness of an inference, and of a material inference to be defeasible.

It is in this sense, then, that logic is on Brandom's conception formal, and this is a deeper sense of 'formal' than simply that the terms are held fixed in a substitutional notion of inferential goodness. *Any* terms can be held fixed in that sense. But logical vocabulary expresses the very form of the scorekeeping practice, and no other terms do. Moreover, since scorekeeping is the practice which makes conceptual content thinkable and sayable in the first place, there is a sense in which Brandom really does hold that logic articulates the form of thought, or of speech—it is just that this traditional idea is transposed into a pragmatic key.[45]

3.4.2 isn't primitive material goodness unacceptably weird?

From the perspective of traditional representationalist thought about logic and inference, Brandom's appeal to primitive material-inferential goodness seems bizarre or incoherent. How could it be that the inference from, say, "Paisley is a cat" to "Paisley is

[45] MacFarlane (2008) makes a similar point about logic and pragmatic form, but in terms of Brandom's slightly different theoretical apparatus presented in (Brandom 2008). One might reasonably wonder whether, given this notion of pragmatic form, Brandom is really taking the material to be "prior to the formal" after all. This is probably just a terminological matter, since Brandom's views as recounted above are directed at the issue of form versus matter on the side of *content*, not of practice, and the priority of the material holds there, regardless of whether we also want to say that, on the side of practice, form and content have equal priority. However, for an argument that Brandom ought to reject priority even on the side of content, see (McDowell 1997).

a mammal" is good, but *not* because of what 'cat' and 'mammal' mean? (Indeed, that things are the other way around, and that these subsentential terms get their meanings from the goodness of the inference (and others like it)?) The idea is difficult, to say the least, and even Brandom himself has difficulty introducing it without lapsing into the old representationalist idiom:

> The kind of inference whose correctnesses essentially involve the conceptual contents of its premises and conclusions may be called, following Sellars, "*material* inference." [footnote excised] As examples, consider the inference from "Pittsburgh is to the West of Philadelphia" to "Philadelphia is to the East of Pittsburgh," the inference from "Today is Wednesday" to "Tomorrow will be Thursday," and that from "Lightning is seen now" to "Thunder will be heard soon." It is the contents of the concepts *West* and *East* that make the first a good inference, the contents of the concepts *Wednesday*, *Thursday*, *today*, and *tomorrow* that make the second inference correct, and the contents of the concepts *lightning* and *thunder*, as well as the temporal concepts, that underwrite the third.[46]

Reading this passage, it is easy to get the picture of material inferences as *deriving* their goodness (or badness) from the antecedently intelligible subsentential expressions in the premises and conclusions (and hence something like analytic inferences). But this would be a representationalist order of understanding, to which Brandom's inferentialism is opposed. Officially, it must be the case that the material-inferential goodness is primitive. But speaking in the official idiom turns out to be very difficult.

One might have two sorts of complaint about the claim that material-inferential goodness is primitive, the first broadly philosophical and the second from the perspective of logical theory.

On the first, the issue is that if the goodness of inference is not to be explained by appeal to subsentential components, then there is simply nothing to explain what makes an inference good, materially or otherwise. Brandom's own strategy is to explain material-inferential goodness in terms of scorekeepers *taking* inferences to be good,

[46](Brandom 1994, pp.97–98)

but on its face this seems to explain not material goodness as such, but only material goodness-according-to-scorekeeper-so-and-so. Hence Brandom's account would seem to entail a kind of subjectivism about inferential norms, and all of the philosophical edifice built on the scorekeeping model—not just the account of logic, or logical validity, but also the whole explanation of conceptual content—would inherit this subjectivism, to its detriment.

This is indeed a serious issue, and Brandom treats it as such. One way in which the issue arises is via the distinction between how things are in reality, and how they merely appear to us; or to put it another way, the distinction between being true and merely being taken to be true. Because this is also an issue that arises in the classical disputes over psychologism, however, and because psychologism—and the status of the expressive conception as a form of psychologism—is the subject of the next chapter, I will not address the issue here. (Next chapter's §4.5 is the place to look for discussion of this problem.)

The second possible complaint with primitive material-inferential goodness is that it is the wrong sort of thing on which to ground a notion of logical validity. In the standard semantic account of validity, we at least know what the property is we're grounding the account on—truth, or truth-preservation from premises to conclusion. But what is this thing called material-inferential goodness, such that validity can be grounded in it?

A short answer is that it's the rightness, or normativity, of the inference. A longer answer comes across the next two chapters. But for now I want to point out that material inference's counterpart in the standard semantic account of validity—truth—is not exactly unproblematic. Although there have been various more or less well-received accounts of truth in formal settings, philosophical disputes about truth—what it is, and even whether we can say anything informative about it—are still ongoing. And as Field (1972) pointed

out about Tarski's definition of truth for formalized languages, the representationalist order of explanation needs some account of what it is for an object to satisfy a predicate, or to fall under a concept, for only then can we understand what it is for a logically atomic sentence to be true. If we *assume* satisfaction is understood, then the rest of Tarski's truth-definitions will follow. But we shouldn't assume satisfaction is understood. Indeed, an ongoing and sprawling enterprise in the philosophy of mind is to make sense of mental content at the level of individual representations like concepts, such that we can make sense of what it is for a logically atomic judgment to be true.

But just as truth is not entirely straightforward, neither is material-inferential goodness entirely mysterious. Brandom aims to explain the norms of material inference as a result of *social institution*. In a way much like we socially institute the value of money (which is otherwise just paper slips or electronic data states) or the entitlement provided by a driver's license (which is otherwise just a plastic card), we institute the norms by which material inferences are to be assessed. One crucial difference between these examples is that in the first two cases, we can appeal to language in explaining how the institution comes about—for example, we can explain the significance of a driver's license in part by pointing to the written language in a state's vehicle code. But of course in explaining the social institution of inferential norms, Brandom cannot appeal to any prior linguistic expression or practice, since it is exactly these which are to be explained by his inferentialist account. This is a real difficulty, to be sure. But it puts Brandom's inferentialist account, and its reliance on material-inferential goodness to underwrite logic, on a par with the standard representationalist account, and its reliance on truth—each account faces a difficulty in making its basic notion intelligible, but neither difficulty is obviously insuperable. There is simply more philosophical work to be done.

3.5 further questions

In the rest of this section I raise and address more specific objections to the expressive conception.

3.5.1 are there parallel logics?

As recounted above, Brandom's substitutional definition of logical consequence has it that X is a consequence of Γ iff every substitution of nonlogical for nonlogical vocabulary yields a value of " Γ, therefore X" which is a materially good inference. And up to now, I've been talking generically of material-inferential goodness or correctness. But there are at least three varieties of material-inferential goodness: commitment preservation, as in

$$\frac{\text{FALAFEL IS A CAT}}{\text{SOMETHING IS A CAT}}$$

entitlement preservation, as in

$$\frac{\text{FALAFEL IS A CAT}}{\text{FALAFEL MEOWS}}$$

and incompatibility-entailment, which can be defined as follows: a claim P incompatibility-entails a claim Q "just in case everything incompatible with Q is incompatible with P" (Brandom 1994, p. 160). An example would be

$$\frac{\text{FALAFEL IS A CAT}}{\text{FALAFEL IS A VERTEBRATE}}$$

since everything incompatible with being a vertebrate (being a robot, being a rock, being

a sea slug . . .) is also incompatible with being a cat. But if the conditional expresses endorsement of a material inference, and there are three types of material inference, then does each sort of material-inferential goodness give rise to its own version of the conditional, hence its own logic? And if so, how do these parallel logics relate to one another?

Here again it is important to distinguish between the account of logical vocabulary in natural languages which Brandom in the first place provides, and formalized logical systems which are created by theorists who have already mastered one or more natural languages. The conditional 'if' which Brandom introduces in *Making It Explicit* prescinds from the particular *reasons* a scorekeeper might have to endorse an inference. This conditional is thus relatively schematic, since it expresses any sort of inferential updating, including the above three. But for the primary discursive job conditionals are supposed to do—"making *implicit* inferential commitments *explicit* in the form of declarative sentences"—this is schematicity is fine. Endorsing any of the three types of material inference is possessing an inferential commitment, and the conditional will make that commitment explicit regardless of which type is involved. So the 'if' of Brandom's expressive conception doesn't pay attention to which variety of material-inferential goodness is being expressed, and hence doesn't give rise to different logics.

When we consider the construction of particular formal systems, however, there are two possibilities to consider. First, we might deliberately introduce formal notions meant to capture more fine-grained notions of material-inferential goodness. This is what, for example, Lance and Kremer do in a pair of papers which treat '$A \rightarrow B$' as intended to capture the meaning of *anyone committed to A is, in part, committed to B*, and they consider various ways of developing this formally.[47] Likewise, Brandom himself develops a formal logic by treating incompatibility as the basic notion, drawing out a

[47] See Lance and Kremer (1994, 1996).

variety of technical results (for example, since incompatibility is a modal notion, a modal logic is generated, and it turns out to be the well-known system S5).[48] But all of these formalizations are sophisticated systems which rely for their construction on a host of conceptual resources beyond Brandom's basic logical notions—for example, they need normative terms in order to specify in the metalanguage what may be done in a proof.

Still, given the ability to create such systems, one might still ask which of them, if any, *really* captures the conditional as Brandom introduces it, along with the rest of the inferential structure of scorekeeping—commitments, entitlements, incompatibility, and all the rest.[49] Surely there is a correct formalization of logic on Brandom's expressive conception? In a reply to Weiss, Brandom demurs:

> On [my] account, the question of correctness [i.e., which is the right logic?] lapses. *No reasonably well-behaved logic is any more correct than any other* (though some – such as classical logic – have other distinctive virtues such as being able to specify in their own terms the inferential roles of their own vocabulary). The right question is rather which aspects of inferential role do the various kinds of vocabulary serve to make explicit. Thus the classical two-valued conditional lets us say that an inference is good in the sense that it does not have true premises and a false conclusion. (Admittedly, this is a pretty weak endorsement of an inference, but it is still an important good-making property of inferences. Those that do not have it are bad.) The intuitionistic conditional lets us make explicit assessments of an inference as good in the sense that there is a recipe for turning a proof of its premises into a proof of its conclusion. C. I. Lewis's hook of strict implication lets us claim that an inference is good in the sense that it is impossible for its premises to be true without its conclusion being true. And so on. I take it that *there is no definite totality of dimensions along which we might want to assess the goodness of inferences, and so no definite totality of possible logical vocabulary.*[50]

Why is there "no definite totality of dimensions along which we might want to assess the goodness of inferences . . . ?" This is Brandom's (characteristically Brando-

[48] See chapter 5 of Brandom (2008).

[49] For example, Lance (2001) claims to provide the correct semantics for Brandom's logical vocabulary in terms of inferential potential, but only as regards commitment (and not entitlement).

[50] Brandom (2010b, p. 353)—emphasis is mine.

mian) way of saying that inferences can have a variety of features, any one of which might at some point be relevant to material-inferential goodness. In saying so, Brandom hearkens back to pragmatism in the sense of Peirce and Dewey. The idea is that we cannot say in advance what future inquiry will require. Just as Dewey urged us to see that *method* (scientific or philosophical) should be seen as an accumulation of successful past practices and revisable in light of future ones, Brandom, as I read him, thinks that the explicit standards to which we hold inferences are the result of the history of logical thought, and that future assessment of inferences might well bring out new standards. This is an instance of the more general idea of pragmatism as opposed to *intellectualism*, which sees behind every practice a set of rules being enacted or followed. By contrast, pragmatism sees the practices as primary, and rules as post hoc attempts to formulate those practices intellectually.

So it appears that on the expressive conception of logic there is no particular formal logical system which is the uniquely correct formulation of the natural logical vocabulary, since that vocabulary is explicative of a practice with indefinitely many aspects that might become of interest to us. And so the hope of formalizing "Brandom's logical theory" is doomed. On the other hand, the way is open for indefinitely many formal logical systems, each of which could have something to illuminate about our inferential practice.

3.5.2 material inferences are defeasible

As we've seen, Brandom's substitutional definition of logical consequence makes logical consequence depend on assessing material inferences as good or not. However, the goodness of a material inference is a fuzzy matter; any material inference is defeasible by some claim (or set of claims), and so material-inferential goodness comes in degrees of robustness. For example, both

THIS CHUNK OF METAL IS COPPER

THIS CHUNK OF METAL WILL MELT AT 1083°C

and

THAT TRACTOR IS GREEN

THAT TRACTOR IS A JOHN DEERE

are materially good inferences, but plausibly the first is more robust than the second—we might say that the class of additional premises which would defeat the first inference (such as "The pressure on the chunk of metal is unusually high") is smaller than the class of additional premises which would defeat the second (such as "That tractor was painted green by its owner"), or we might say that the premises in the first class are less probably than those in the second. In any case, materially good inferences can be more or less robust, and then the question arises whether there is some threshold of robustness at which a materially good inference becomes bad. (And if so, how would we know?) For example,

YOSHI HAS A PET

YOSHI HAS A CAT

is sort of appealing, in that many pets are cats. Supposing for the sake of the example that cats make up a plurality of pets, then the conditional probability that Yoshi has a cat, given that Yoshi has a pet, is greater than the conditional probability of Yoshi having an x given Yoshi having pet, for any non-cat animal x. But is this enough to make the inference materially good? Surely it's more likely, given he has a pet, that Yoshi has a non-cat than that he has a cat. So what are we supposed to say about the material goodness of the inference? Any fuzziness or uncertainty here is inherited by Brandom's definition

of validity, since that definition requires all substitution instances to be materially good inferences.

This would indeed be a problem for Brandom's account of validity if it turned out that the verdict on a particular form turned on materially suspect inferences like the one just above. But I think we have no reason to expect that things will turn out that way. The logical form of the above inference, for example, is "*P*, so *Q*"—and although there are many substitution instances the goodness of which will be equally hard to determine, there are plenty of substitution instances which are clearly bad, and only one is needed to show that the form "*P*, so *Q*" is invalid.

On the other hand, the expressive role of the logical terms will ensure that problematic substitution instances don't arise. Take *modus ponens* as an example. Its substitution instances will include, among others,

IF YOSHI HAS A PET, THEN YOSHI HAS A CAT
YOSHI HAS A PET

YOSHI HAS A CAT

This counts as a materially good inference, but *not* because it has somehow made the previous example unproblematic. Rather, the first premise expresses endorsement of exactly the material inference from 'Yoshi has a pet' to 'Yoshi has a cat', so that inferring the conclusion from the premises is materially good. The conditional is ensuring that no matter what claims are substituted for in the inferential form, a good material inference results.

Here is another way to put the point. In Brandom's normative pragmatics, the semantic significance of a move is the difference it makes to the deontic score. When I assert 'If *P* then *Q*', it expresses endorsement of the inference from *P* to *Q*, and hence changes the deontic score in the following way: it makes it appropriate for any

scorekeeper who attributes to me *commitment to P* to also attribute *commitment to Q*. But as a matter of deontic scorekeeping this is exactly the same circumstance which obtains when the scorekeeper herself takes the inference "*P*, so *Q*" to be materially good. Since the conditional has this effect on deontic score, every instance of *modus ponens* is guaranteed to be a materially good inference.[51]

Logical vocabulary thus obviates worries about the defeasibility of material inferences, or the robustness thereof. This is not to say that defeasibility never matters, but only that it doesn't matter for assessing the substitution instances of logically valid forms, since in these forms the expressive role of logical vocabulary makes defeasibility irrelevant.

3.5.3 what about Tarski's problem?

In his classic paper "On the Concept of Logical Consequence" Tarski identified a problem with the substitutional definition of logical consequence: it accounts for logical consequence in a way limited by the expressive power of the language. So, for example, if the only two names in the language are 'Plato' and 'Aristotle' and the only predicate is 'is a philosopher', then 'Plato is a philosopher' is a logical consequence of 'Aristotle is a philosopher'. Tarski took this as reason to reject the substitutional definition and bring in the apparatus of models. However, Brandom's substitutional definition appears subject to exactly the same difficulty, and he explicitly eschews construing validity as set-theoretic inclusion, as the model-theoretic definition construes it. But surely 'Plato is a philosopher' isn't a logical consequence of 'Aristotle is a philosopher', even if the language is impoverished—how can Brandom's substitutional definition avoid this result?

Once again, we need to distinguish between the formal languages like those under consideration in Tarski's paper, and the natural languages about whose logical vocabu-

[51] I am indebted to Rosenberg (1997) for helping me think of conditionals in this way.

lary Brandom is theorizing. Tarski's counterexample to the substitutional definition is generated by considering an expressively limited formal language, and—crucially—is generated against the background of the actual natural languages we speak, which are expressively much richer, and from which we draw our linguistic intuitions, since these are the languages we inhabit.

So suppose, for example, that there were some creatures—call them the Eleatics—whose discursive practice had only the names 'Plato' and 'Aristotle' and the predicate 'is a philosopher', and that this is not some formal language isolated or constructed from a larger practice, but really is the Eleatics' natural language. We should see at once, even if we're not Brandomeans, that what they mean by these terms isn't what we mean by them. But supposing we do take up Brandom's theoretical perspective, we see that 'Plato' and 'Aristotle' designate the same individual, since, assuming the predicate can be made sense of, all possible substitution inferences replacing one by the other are materially good. Then again, the predicate can't really be made sense of, since it isn't inferentially articulated. In the languages *we* speak, both 'Plato' and 'Aristotle' satisfy the predicate 'is a philosopher'; but this depends, among other things, on being able to distinguish philosophers from the many other sorts of things in the world. On Brandom's view, being able to so distinguish things is a matter of the inferential connections 'philosopher' participates in that other predicates don't. But the Eleatics have no other predicates which could differ in their inferential connections. So not only can they not distinguish philosophers from other things—they can't distinguish anything from anything else.[52] Apparently, the only thing the Eleatics could express in their language is (translated into ours) something like "Universe is a thing" or "Existence is thing-y."[53]

[52] At least, not conceptually. They might be able to discriminate some things from other things via their senses, but that's a different matter.

[53] It might even be worse than this. For reasons such as those Strawson (1959) raises in the first chapter of *Individuals*, beings who lacked the resources to reidentify particular objects might be unable to distinguish themselves from the rest of the world around them. If so, we shouldn't regard the Eleatics as expressing anything like what we can say in our languages.

This really would be a strange and impoverished language, if it could exist at all. But so when the Eleatics infer 'Plato is a philosopher' from 'Aristotle is a philosopher', we could capture the form of this inference with "*P*, so *P*." This is a valid inference in most logical systems, and if it constitutes a counterexample to any account of logical consequence that ratifies it, then most of modern logic will have to go.[54] But on the other hand, given that "*P*, so *P*" is generally accepted, perhaps we don't really have a counterexample to Brandom's substitutional account of validity.

Now the fact that this simple proposed counterexample doesn't really work against Brandom's substitutional account doesn't at all show that there aren't any such cases which could be genuine counterexamples. But consideration of the Eleatics is instructive in at least two respects.

First, Brandom's substitutional account of validity is supposed to be an account of validity for inferences expressible by deontic scorekeepers participating in a full practice. So any proposed counterexample should attempt to situate itself within such a practice and the language it would afford, not some restricted artificial language whose meanings we can stipulate.[55]

Second, there is a dialectical problem with framing an objection in this way—namely, as a challenge to the *extensional adequacy* of the account. On this way of framing things, we ask whether the account captures all and only the genuinely valid arguments. If not, the account is extensionally inadequate since it pronounces some invalid arguments valid, or some valid arguments invalid.

[54]Interestingly enough, Aristotle seems to have ruled out inferences like this in his definition of the syllogism, since he stipulates that "A syllogism is discourse in which, certain things being stated, *something other than what is stated* follows of necessity from their being so." *APr* i.1 24^b18–22 (Barnes 1984)—emphasis is mine. (See also *Top.* 100^a25–27.) Some modern logicians have devised systems meeting this irreflexive condition on logical consequence, e.g. the system *S* of Martin and Meyer (1982).

[55]For a related point about expressively impoverished languages, see Brandom's response to Michael Kremer in (Brandom 2010c, pp. 350–51). Kremer raises a challenge about the nature of singular terms on Brandom's account, e.g. whether it can distinguish singular terms from lowest-species predicates (those at the bottom of a classificatory hierarchy).

The trouble with the extensional adequacy question is that it is not independent of more general philosophical theses about the nature of logic. So, for example, Brandom has his own preferred way of picking out the specially logical vocabulary, according to his idiosyncratic philosophical take on logic—it is the vocabulary by which we make explicit the content-conferring features of our discursive practice. Does the vocabulary so chosen yield an extensionally adequate set of logical forms? The answer to that question varies exactly with one's conception of logic: *yes* if you like Brandom's conception, almost certainly *no* otherwise. And the difficulty remains for all other conceptions of logic which have any bearing on the demarcation of logical vocabulary; the logical forms derived from those conceptions will look adequate if and only if the conception is endorsed in the first place, which is to say that the question of extensional adequacy cannot be evaluated in a philosophically neutral way.

Take the proposed counterexample above, and suppose it works as intended, with the terms meaning what they do in English. Given Brandom's conception of the nature of logic, there is nothing unusual about calling such arguments as "Plato is F, so Aristotle is F" valid. If the role of logical terms is to express features of discursive practice, then any form which does so successfully fulfills its role and thereby counts as logical. 'If' is, according to Brandom, the term which (in English) expresses the correctness of a material inference. Since, in the impoverished language we're looking at, any available substitution for 'F' in this schema yields a good material inference, it's a good logical form. The fact that we speakers of English can substitute predicates for 'F' which thwart the inference does not entail that the schema codifies what is, for speakers of the impoverished language, a pair of good inferences.

The situation is somewhat analogous to that found in normative ethics, where putative counterexamples to normative ethical theories carry no dialectical force against committed proponents of those theories. A committed Kantian can insist that one

shouldn't lie to the murderer at the door, since doing so involves a contradiction of practical reason. A committed utilitarian can insist that one ought to push the fat man and stop the runaway trolley, since doing so will save five lives at the cost of only one—squeamishness shouldn't prevent us from doing the morally right thing!

In normative ethics, it is sometimes thought that these putative counterexamples do carry dialectical force, precisely because the status of a claim like "It is morally wrong to push the fat man in front of the trolley" is epistemically or cognitively independent of any ethical theory. This independence might be because we have non-theoretical *intuitions* that such claims are true, or it might be because these are facts of moral life which we know *before* we engage in ethical theory, and to which subsequent ethical theorizing is beholden. In either case, we can have genuine counterexamples to normative ethical theories, since the source of our confidence in those counterexamples is independent of the ethical theories we adjudicate by them.

But no matter how plausible this view might be in normative ethics, the analogous position in philosophy of logic is highly implausible. Although we might well have something like intuitions about some bare notion of consequence or "following from," specifically *logical* consequence, or validity, is a philosophers' notion which one acquires by initiation into the tradition. As MacFarlane (2004, p. 2) puts it,

> ...our intuitions about logical validity, such as they are, are largely the products of our logical *educations*. Anyone who has taught elementary logic will know that it can take a week or more to get students to distinguish questions of validity from questions of soundness. Even after they've caught on, their intuitions about validity are likely to depend heavily on the glosses they have been offered. ...Professional philosophers and logicians tend to have more settled intuitions. But that is because they have had more time to become indoctrinated.

So proposed counterexamples which depend for their force on intuitions of logical (in)validity are dialectically quite weak. To take a particularly striking example, the

inference

$$\frac{\begin{array}{l}\textsc{Lois believes that Clark Kent is her coworker}\\ \textsc{Clark Kent} = \textsc{Superman}\end{array}}{\textsc{Lois believes of Superman that he is her coworker}}$$

is on Brandom's view valid, since it is good in terms of the form

$$\frac{\begin{array}{l}S \text{ believes that } F(a)\\ a = a'\end{array}}{S \text{ believes of } a' \text{ that } F(it)}$$

and this form is articulated by the vocabulary expressing *de re* and *de dicto* belief attribution, plus identity, and these three count as pieces of logical vocabulary on Brandom's criterion.[56] But the idea that logical vocabulary includes expressions for belief attribution is quite unorthodox, and even identity has its detractors.

Rather than assess Brandom's (or anyone's) account of logical vocabulary and validity in terms of their extensional adequacy, then, it's best, I think, to see whether philosophically satisfying results can be gleaned from the account. The task of the next two chapters is to look at two broad questions in the philosophy of logic, and see how Brandom's expressive conception can provide satisfying new answers to those questions.

[56]The two kinds of belief attribution express the scorekeeping significance of *S*'s deontic status relative to *S*'s collateral commitments (*de dicto*) and relative to the attributor's collateral commitments (*de re*). Identity expresses endorsement of certain substitution inferences, as recounted above in §3.2.

Chapter 4

The Expressive Conception as a Form of Psychologism

In the following I will argue that Brandom's conception of logic is a form of *psychologism*, a family of views about logic which was widely criticized in the late 19th and early 20th centuries, and is now generally considered a nonstarter. I will also argue, however, that the form of psychologism manifested by Brandom's conception avoids the most damning problems from the bad old days of psychologism, and is thus viable. This viability is important, for it opens up space for a kind of naturalistic account of logic.

The structure of the chapter is as follows. §1 briefly makes some background distinctions relevant to philosophical naturalism, expressivism as one among many naturalist proposals, and logical psychologism as another, older, form of naturalism. §2 then explains why Brandom's expressive conception of logic counts as a form of psychologism. §3 presents several of the deepest objections to psychologism, and shows how Brandom's view avoids or defuses those objections. Finally, in §4 I draw out the significance of the results from the previous section for philosophical naturalism.

4.1 naturalism, expressivism, psychologism

4.1.1 naturalism

'Naturalism', like many philosophical -isms, is a single label which stands for a multitude of possible views, different in detail but united around some general principle. I will not here attempt to define the label. Instead, I will make explicit some of the guiding aims which (as I see it) various naturalist proposals intend to honor.[1]

Naturalism is most helpfully seen, I think, as a *metaphilosophical* position regarding which (first-order) philosophical proposals are acceptable. It is endorsed by most Anglophone philosophers today, and by those working in so-called analytic philosophy during the 20th century—as Kim puts it, "if current analytic philosophy can be said to have a philosophical ideology, it is, unquestionably, naturalism" (Kim 2003, p. 84). That is, philosophical proposals which qualify as naturalistic earn, by virtue of that qualification, a kind of approval (though of course the substance of any such proposal will still be contested). Explicitly *non*-naturalist proposals, by contrast, face a corresponding burden: they must show why they deserve to be taken seriously despite their non-naturalistic elements.

Is there a good rationale for this disciplinary asymmetry? I think there is—it's a matter of epistemic modesty—but to say why I will need to first say a bit more about the substance of naturalism.

We can think of naturalism broadly as having two facets, one metaphysical and one epistemological.[2] On the metaphysical side, naturalism says: any philosophical

[1]For a brief historical take on naturalism in the United States, as well as its late-20th-century manifestations in Anglophone philosophy, see Jaegwon Kim (2003). Kim attributes the genesis of American naturalism to philosophers like John Dewey, Ernest Nagel, and Roy Wood Sellars—although he acknowledges that, in general, "philosophical naturalism of course goes back further, at least to Hume" (p. 84).

For a more systematic and less historical (but still brief) take on naturalism, see the opening essay from Huw Price's (2011), which situates various types of naturalist proposal in logical space.

[2]Here I follow Kim in distinguishing these two facets, whose historical paper identifies evidence of both

theory ought to be consistent with the fact that *all that exists is the natural world*. On the epistemic side, naturalism says: any philosophical theory ought to be consistent with the fact that *the methods of the natural sciences are the only means to knowledge*.[3] Of course, many of the load-bearing components of this broad characterization—'natural world', 'method', 'science', 'natural science', even 'exists'—can be made precise in many ways, and different varieties of naturalism result from doing so, yielding more or less stringent stances toward the acceptability of first-order philosophical proposals.

At its most stringent, naturalism demands that philosophical proposals appeal to only such notions as are employed by the natural sciences (at least physics, and possibly chemistry and biology), or are reducible to such notions. On the ontology required by a stringent naturalism, the natural facts are of two types: either (1) those discovered by the natural sciences directly (and hence expressible in the vocabularies of those sciences) or (2) those whose home is outside the natural sciences, but which are "nothing over and above" facts of type 1. Facts of type 2, as stringent naturalists have it, are *reducible* to facts of type 1.[4]

But what about facts—call them, predictably, facts of type 3—whose home is

---

in the writing of the early American naturalists. However, Kim also identifies a third, "methodological" component, which I take to be just a special case of the epistemological. The characterization which follows is my own, not Kim's.

It is possible to combine the two facets, as does e.g. Price: "The Naturalist's mantra goes something like this: The only facts there are are the kind of facts recognized by natural science" (Price 2011, p.4). But such a formulation clearly retains both ontic and epistemic dimensions.

[3]It is of course possible to construe naturalism as a first-order philosophical view, consisting of commitment to the propositions just italicized. (In which case, of course, any naturalist will aim to hold other views consistent with those two propositions.) But construing it as a metaphilosophical view makes two things clear: first, that among practitioners of philosophy, naturalism functions as a kind of tribunal for proposed theories, and second, that naturalism's directives are meant to apply to the practice of philosophy itself.

[4]Of course there is controversy about how to characterize the relation between what I'm calling facts of type 2 and facts of type 1, as well as about how to characterize the vocabularies we use to talk about these facts. 'Reduction', for example, is usually taken to express a very strong relation, and perhaps something weaker, like supervenience, will do (although supervenience itself comes in weaker and stronger forms, the strongest of which arguably constitute reduction). I am here bypassing such controversy, and I will talk about reducibility in a deliberately hand-waving fashion, to give the outlines, but not the details, of an extant naturalist position.

outside the natural sciences but which cannot be reduced to facts of type 1? Here the stringent naturalist must claim that such so-called facts are merely apparent, and that in reality there are no facts of type 3. But this claim will, for some purported facts of type 3, face a *prima facie* implausibility; some of the prime candidates for facts of type 3 are facts of morality, numbers, aesthetic values, the mind, linguistic meaning, and modal facts about what's necessary or possible. Are we really to accept, as the stringent naturalist would have it, that there are no facts of the matter whether causing gratuitous harm is wrong, whether 4 is the square of 2, or whether *Sgt. Pepper's* is better than *Help!*, or whether it's possible for me to have studied economics instead of philosophy?

Here there are three general options for a naturalist. One is to relax the naturalist criteria in some way, perhaps by widening the scope of what counts as natural science, or by weakening the relation required to hold between type 1 and type 2 facts. In this way, some purported type 3 facts can be captured as type 2 facts.[5] The second general position is to simply stick to one's stringent-naturalist guns and adopt an *error theory* about "facts" of type 3. The error theorist holds that there are no such facts, and that our discourse which superficially seems to be about them is simply false.[6]

The third general option for a naturalist to take is to say that facts of type 3 can be accounted for in some nuanced way—not in a way which awards them real existence, as possessed by facts of types 1 and 2, but in a way which avoids flatly banishing them from our ontology, as the error theorist does. Naturalists have articulated this third

[5]For example, consider the position Nicholas Sturgeon takes in "Moral Explanations" [CITE], in which moral facts turn out to be ontologically respectable because they can figure in explanations of historical events and individual behavior. Sturgeon can be construed, in the terms used here, as broadening the ambit of natural science to include history (and hence enlarging the set of type 1 facts), or, more plausibly, as weakening the relation required of type 2 facts to be something like *figuring in an explanation of phenomena which are themselves characterizable in type-1 terms*. In either case, moral facts turn out to be facts of type 2.

[6]Here the most salient exemplar might be J.L. Mackie's error theory about objective values in (Mackie 1977), or perhaps Paul Churchland's error theory about propositional attitudes in (Churchland 1981). For an error theory of just about everything, adopted as a consequence of stringent naturalism, see (Rosenberg 2011).

general option in a variety of ways, among them *fictionalism* (claims about type-3 facts are true-in-a-fiction), *emotivism* (claims about type-3 facts are really just expressions of speakers' emotions), *prescriptivism* (claims about type-3 facts are disguised imperative statements), and *(philosophical) behaviorism* (claims about type-3 facts are really claims about behavior).[7]

Expressivism is another member of this family of naturalist views, and one of the most recent. Since it is the subject of §2.2, and receives more detailed treatment in Chapter 5, I won't say more about it here. However, before changing focus to expressivism in particular, I first want to say a bit more about naturalism in general, its aims, what's at stake in the attempts to articulate first-order philosophical views which respect the naturalist constraints, and why we should care.

For, in addition to the three general naturalist positions just adumbrated, there is another philosophical option: reject naturalism. And indeed, explicitly non-naturalist views have been articulated even during Anglophone philosophy's late-20th century naturalist heyday, about God, souls, qualia, moral facts, and mathematical objects.[8] So it's not as if non-naturalist views are *untenable* by the lights of the discipline; they're simply unpopular. Moreover, some self-described non-naturalist views look, under a certain light, very similar to what a relaxed naturalist would happily endorse. For example, Russ Shafer-Landau (2003) defends what he calls a non-naturalist moral realism, according to which moral facts are necessarily coextensive with, but not reducible to, natural descriptive facts. On his account, although moral facts figure in explanations, they are not causally efficacious. This position is very close, *mutatis mutandis*, to nonreductive physicalism in the philosophy of mind, as Shafer-Landau fully acknowledges (pp. 72–74). However, nonreductive physicalism is typically considered a naturalistically respectable view!

[7]For exemplars of fictionalism, see (Field 1980), (Rosen 1990), (Kalderon 2005). For emotivism, see (Stevenson 1944), (Ayer 1952). For prescriptivism, see (Hare 1952). For behaviorism, see (Ryle 1949).

[8]See, for example, (Swinburne 1977), (Unger 2006), (Enoch 2011), and (Linsky and Zalta 1995), respectively.

Perhaps a stringent naturalist would refuse to countenance any purported facts which are not reducible to physical facts, but a relaxed naturalist who favors *supervenience* for the relation between facts of types 1 and 2 could be quite happy with nonreductive physicalism, whether about mental or moral facts.[9] If naturalist and non-naturalist views can converge so closely, what does it matter whether a view is naturalist or not? At worst, we are faced with a merely verbal disagreement, or a ridiculous tussle over who gets to affix a badge reading 'naturalist' to their own favored positions.

So should we care about naturalism and the constraints it imposes on first-order philosophical proposals? I think the answer is yes, because there is ultimately a worthwhile motivation behind the naturalist trend, and that is a version of epistemic modesty.

This might sound like an odd virtue to attribute to a movement which is inclined to deny existence to things with which we seem intimately familiar, if not acquainted. But to the extent that naturalism is motivated by awareness of the history of human inquiry and the radical successes of science, I think it bespeaks a respect for our cognitive limitations, admiration for the difficulty of finding things out, and circumspection about the beliefs we find easy, natural, or necessary. The naturalist has no shortage of examples of beliefs humans once held which were not only ill-founded, but genuinely harmful. For example, if today's moral discourse is on a par, naturalistically speaking, with 16th-century witch discourse, then the sooner we jettison moral discourse, the better.

So let's take a closer look at expressivism, which aims to preserve the meaningfulness of a naturalistically suspect target discourse without invoking a naturalistically disreputable ontology.

[9]Indeed, a former colleague of mine remarked, after we'd both read Shafer-Landau's book, that he'd "really hoped for something a little more . . . queer" (referring of course to Mackie's claim that any objective values would be both metaphysically and epistemically "queer," i.e. uncanny). I shared the slight disappointment that a view advertised as non-naturalism turned out to be so familiar.

4.1.2 expressivism, the basic idea

As a first pass, let *expressivism* about some vocabulary K be the idea that K's semantics is to be explained by recourse to the mental states it expresses. Hence one of expressivism's cardinal features is that, since its explanatory substrate is a class of mental states, it requires no appeal to the objects or properties that K superficially appears to talk about.This makes expressivism an attractive option where there is reason to be shy of ontological commitment; an expressivist semantics can allow for K-discourse to be meaningful without thereby endorsing a suspect ontology.[10] That is, expressivism appears to avoid entirely a serious *prima facie* problem faced by a representationalist approach to semantics: where a linguistic item has meaning, it's because it represents some *thing*, but not everything we can talk about is something the existence of which we wish to affirm. Instead of looking outward, to what is represented, expressivism looks inward, to what is expressed.

Indeed, this is one of the motivations for the original noncognitivist programs in metaethics—emotivism and prescriptivism. The classical analytic philosopher's naturalistic and empiricistic scruples made moral properties look suspect, and expressive accounts of moral language offered a way to preserve moral discourse without countenancing the existence of apparently nonnatural properties. Much the same is true of the second-generation expressivisms of Blackburn (1984, 1998) and Gibbard (1992, 2003). The general idea is to explain the meaning of normative claims in terms of the mental states those claims express, for example attitudes of approval, in just the same way that descriptive claims express beliefs. Mark Schroeder calls this the *basic expressivist maneuver*: treat all assertions as expressive of mental states, explain the meaning of the

[10]There might also be the following motivations for expressivism: (1) some domain of discourse seems to be closely tied to certain mental states, as moral discourse does to attitudes of approval and disapproval; (2) theories of meaning—meaning in general, not meaning for some restricted vocabulary—might be necessarily expressivist, for independent reasons; (3) the best cognitive science or linguistics might rely on expressivism. I will not address these motivations here.

sentences in terms of the meaning of the mental state expressed, and locate the difference between descriptive and normative claims as the difference between the states of mind being expressed. Thus expressivism about normative language "gives an account of moral language *by* giving an account of moral thought" (Schroeder 2008, p. 4).

So motivated, expressivism is a well-received (if controversial) member of the philosophical landscape, and as Schroeder notes, expressivist views have been advanced for vocabularies beyond the narrowly moral (2008, p. 6). But it has hardly been noticed that logical vocabulary affords an expressivist treatment for exactly the same reason: what objects, or what properties, could logical vocabulary like 'not' possibly be *about*?[11] While difficulties making sense of discourse about abstract objects are well known, and for the case of mathematical objects have been forcefully articulated by Benacerraf (1965, 1973), the fact that Benacerraf's problems might correspond to Mackie's charges of "queerness" against objective values has not, to my knowledge, been appreciated.

Indeed, logic might seem to be an especially apt target of Mackie-type concerns, due to its arguably normative character. Mackie's charges of queerness break down along two lines,

> one metaphysical, the other epistemological. If there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe. Correspondingly, if we were aware of them, it would have to be by some special faculty of moral perception or intuition, utterly different from our ordinary ways of knowing everything else.[12]

Benacerraf's discussion of the difficulties with mathematical objects corresponds only to the epistemological side of Mackie's argument. On the metaphysical side, Mackie is worried that objective values would somehow have to be inherently motivating entities—things which, when we know about them, provide us with not only knowledge of what's

[11] Indeed, even Brandom's expressive conception of logic, the only one extant, is not motivated in this way, but rather as part of his systematic philosophy of language, which is motivated by the more general concern to explain meaning in terms of use.

[12] (Mackie 1977, p. 38)

right and good, but motivation to act in a certain way. But nothing else in the known world seems to be like this, and so objective values, were they to exist, would be metaphysically queer. This concern doesn't apply to mathematics, since mathematical objects, if they exist, don't need to motivate us any more than ordinary concrete objects do. But logic, one might think, is less like mathematics and more like morality, in that the objects, facts, or relations of logic are rationally compelling. Knowing that a set of one's beliefs is logically inconsistent is, generally speaking, a motivation to revise one's beliefs. (Of course, not everyone is motivated equally by such knowledge, and such motivation doesn't necessarily result in actual belief revision, but the same is true, *mutatis mutandis*, of moral knowledge and its relation to action.)

So it would seem that logic raises some of the same metaphysical and metanormative issues that ethics does, and yet 20th century philosophy of logic set before itself an entirely different set of questions. In any case, no analogous expressivist program has been proposed for logical vocabulary. Why not? I suspect that there are two main reasons. First, philosophers' way of thinking about the meaning of logical language developed in a different context from the discussion of moral language. Second, expressivism would appear on its face to be a form of psychologism about logic, a view which was already moribund by the early 20th century. I'll discuss these two issues in turn, though detailed discussion of psychologism's problems will come later, in §4.

4.1.3 two approaches to the meaning of logic

The history of philosophical logic saw a great advance with Tarski's famous pair of essays (1933, 1936), which formed the template for so much further work in formal semantics. From our post-Tarski vantage point, it might seem that since logic is the very birthplace of model-theoretic semantics, logical vocabulary is the thing that's least semantically problematic. Moreover, logic appears to traffic in truth, and to do so quite

promiscuously, or even essentially; so if expressivism is received as a kind of flight from truth-conditions, along the lines of Ayer's emotivism, then logic might look like the last thing in need of expressivist semantics.

But just as Field (1972) observed about Tarski's "semantic" conception of truth—namely, that it's not semantic at all—we ought similarly to notice that Tarski's celebrated semantic account of logical consequence provides nothing like a semantic account for logical vocabulary, if semantics is supposed to relate words and things. In this sense, all of the genuinely semantic machinery in Tarski's account is working in service of the *nonlogical* vocabulary: it is the names and predicates which are associated with the various objects and sets that belong to models. The logical connectives and quantifiers, on the other hand, are given their interpretations via clauses in the metalanguage, like so:

[Conjunction]: '$(\alpha \wedge \beta)$' is true iff 'α' is true and 'β' is true

Now this does give us the truth-conditions for conjunctions, but what it absolutely does not do is associate the *term* '$\wedge$' with any kind of nonlinguistic *thing*. Indeed, what sort of thing could '$\wedge$' possibly stand for? Only one answer has been given, proposed independently by Sher (1991) and Tarski himself (1986): there are special abstract objects, the *formal operators*, which logical vocabulary names. (For example, '$\wedge$' stands for the formal operation of set theoretic intersection.) Doubts have been raised about the adequacy of this proposal,[13] but none of that discussion mentions the fact that, to an empirically-minded naturalistic philosopher, such formal operators ought to look at least as queer as normative properties like rightness, or normative objects like vice.

A second reason for the lack of expressivist theories of logic might be that, since at least the work of Gentzen (1935), there has been a second way of understanding the meaning of logical vocabulary: via the syntactic rules that govern a term, in particular the

[13]Notably by McGee (1996), Feferman (1999), Etchemendy (2008), Hanson (1997), and MacFarlane (2000).

term's introduction and elimination rules. Conjunction, for example, gets the following rules:

$$\wedge\text{-}\textit{Introduction}: \quad \frac{\begin{array}{c}\Phi \\ \Psi\end{array}}{\therefore\ \Phi \wedge \Psi}$$

$$\wedge\text{-}\textit{Elimination}: \quad \frac{\Phi \wedge \Psi}{\begin{array}{c}\therefore\ \Phi \\ \therefore\ \Psi\end{array}}$$

Together, these rules tell us when we may conclude a conjunction, and what we may conclude from it; and this, one might think, is all we need to know in order to understand '$\wedge$'. What's more, this approach to elucidating the meaning of logical terms (which can, of course, be extended to the other connectives and the quantifiers) makes no appeal to any special logical objects, or indeed any nonlinguistic objects at all. In this respect, such an "inferential role" approach might seem to capture the expressivist motivation, namely, to preserve the meaningfulness of K-discourse (in this case, language involving logical vocabulary) without commitment to a suspect K-ontology (in this case, abstract logical objects). And indeed, for just this reason the inferential role approach has been recruited by philosophers with anti-realist sympathies.[14]

But the inferential role approach faces its own problem, which we can see by looking to what it shares with the Tarskian approach. What both approaches do in the abstract is associate some piece of *interpreted* vocabulary with some other thing, the *interpretant*. For Tarski, the interpretants are objects (for names) and properties or relations (for predicates). For Gentzen, the interpretants are pairs of syntactic rules. Gentzen's interpretants might be more ontologically respectable, in that they appear to

[14] See, for example, Dummett (1978). Even Brandom's inferentialist semantics, very much inspired by Dummett, is partly due to his conviction that representationalism's appeal to *representation of objects* as a semantic primitive is unacceptable. But it would be inaccurate, I think, to call Brandom an anti-realist, given his aim to habilitate the notion of representation from the inferentialist standpoint.

be just more bits of language, but they make plain a pressing problem: the syntactic rules are merely stipulations about how to manipulate a formal language, and insofar as we're interested only in such manipulation, one pair of rules is as good as any other.

But of course we're interested in more than simply manipulating a formal language; the whole point of logic is to somehow help us with the inferences we can actually make in our own language. So if a formal language is to qualify as a *logic* in any interesting way, it has to be *relevant*—not in the sense of what are known as relevant (or relevance) logics, but in the sense that the interpretants associated with logical vocabulary have to meet some standard that's present in our antecedent inferential practice. Thus we must always ask, for any candidate piece of logical vocabulary: which rules are the right ones?

One way to read Prior's "The Runabout Inference Ticket" (1960) is as directing us to just this question. Suppose we wanted to include in our formal language the connective *tonk*, which behaves like this:

Tonk Introduction: $\dfrac{\Phi}{\therefore\ \Phi \star \Psi}$

Tonk Elimination: $\dfrac{\Phi \star \Psi}{\therefore\ \Psi}$

In a language that contains tonk, we can infer anything we like from anything we like—every inference is permitted. But our actual inferential practices are more discriminating, which shows that we don't actually employ anything that's correctly translated by tonk. Or, to put it another way, tonk's interpretant is defective from the point of view of our inferential practice.[15]

[15]This is not the only way to read Prior's paper, which is perhaps more naturally seen as directed against the very idea that inference rules can determine the meaning of a logical constant. However, the ensuing discussion with Belnap (Belnap 1962, Prior 1964) shows that the issue of relevance—what Belnap would put as the comportment with the "antecedently given context of deducibility"—is central to Prior's paper.

We can understand Brandom's expressive conception of logic, then, as offering a theoretically motivated account of relevance to the Genzten-style inferential role semantics for logic, thus meeting the main challenge to this way of doing semantics. But in so doing, Brandom's account appeals to what are *prima facie* psychological properties, and this way lies a notoriously disreputable naturalistic program for how to understand logic: psychologism.

4.1.4 psychologism

In order to see how Brandom's conception of logic counts as a form of psychologism, we first need to understand what psychologism is. I will here say only enough about psychologism and its challenges to set up the later sections of the chapter. However, detailed accounts of the history of psychologism are available elsewhere.[16]

The label 'psychologism' is used for philosophical theorizing which attempts to explain the facts of some target domain in terms of psychological facts, or the science of psychology. Its use is almost always pejorative, indicating that its user regards the appeal to psychology as a mistake.[17] The term comes to us from a protracted debate in German-language philosophy during the late 19th and early 20th centuries—the *Psychologismusstreit* or "psychologism conflict"—during which the role of psychology in philosophical theorizing was disputed. As Kusch (1995) notes, a variety of proposals for the relevance of psychology to philosophy were floated, but a chief bone of contention concerned the relevance of psychology to logic in particular: whether logic is a part of psychology, or based upon it, or studied by it; or whether, on the contrary, logic and

[16]For a fascinating history and sociology of psychologism's rise and fall, see Martin Kusch's (1995). See also Kusch (2014); Pelletier, Elio, and Hanson (2008); and Stelzner (2005).

[17]As Brentano put it in 1911, 'psychologism' "is a word which has lately come into use and when it is spoken many a pious philosopher—like many an orthodox Catholic when he hears the term Modernism—crosses himself as though the devil himself were in it" (From Appendix 11 to Brentano 1911, as cited in Jacquette, 1997).

psychology are independent disciplines with distinct subject-matters.

A key figure in the psychologism debates was John Stuart Mill, whose (1843) *System of Logic* was a paradigmatic and widely read work of psychologistic logic. (Indeed, Husserl credits Mill with having made psychologism such a widely held view.)[18] At the beginning of the work, Mill frames the subject thusly:

> Logic has often been called the Art of Reasoning. A writer [fn: Archbishop Whately] who has done more than any other person to restore this study to the rank from which it had fallen in the estimation of the cultivated class in our own country, has adopted the above definition with an amendment; he has defined Logic to be the Science, as well as the Art, of reasoning; meaning by the former term, the analysis of the mental process which takes place whenever we reason, and by the latter, the rule, grounded in that analysis, for conducting the process correctly. There can be no doubt as to the propriety of the emendation.[19]

For Mill, half of logic (the Science) consists in an analysis of mental processes, and the other half (the Art) consists in prescriptions for reasoning which are "grounded in" that analysis. The Science of logic is thus a specialized part of psychology, and the Art of logic is beholden to it:

> [Logic] is not a Science distinct from, and coordinate with, Psychology. So far as it is a science at all, it is a part, or branch, of Psychology; differing from it, on the one hand as a part differs from the whole, and on the other, as an Art differs from a Science. Its theoretic grounds are wholly borrowed from Psychology, and include as much of that science as is required to justify the rules of the art.[20]

This conception of logic as (on the one hand) a part of psychology shows itself at work in Mill's account of such venerable logical rules as the Principle of Non-Contradiction—

[18] "Owing particularly to the influence of the distinguished thinker just mentioned [i.e. Mill], the first of the three main tendencies that we find in logic, the psychological, has definitely come to prevail over the formal and the metaphysical tendencies, both as regards the number and the importance of its exponents" (Husserl 1900/2001, p.11).

[19] (Mill 1843, pp. 17–18)

[20] (Mill 1865, p. 388)

> I consider it to be, like other axioms, one of our first and most familiar generalizations from experience. The original foundation of it I take to be, that Belief and Disbelief are two different mental states, excluding one another. This we know by the simplest observation of our own minds.[21]

—and the Law of Excluded Middle, which in its usual form Mill actually *denies* on psychological grounds, claiming that it holds only for propositions whose predications "can in [some] intelligible sense be attributed to the subject" (*ibid.*). Mill then endorses Herbert Spencer's claim that

> The law of the Excluded Middle, then, is simply a generalization of the universal experience that some mental states are directly destructive of other states. It formulates a certain absolutely constant law, that the appearance of any positive mode of consciousness can not occur without excluding a correlative negative mode; and that the negative mode can not occur without excluding the correlative positive mode: the antithesis of positive and negative being, indeed, merely an expression of this experience. Hence it follows that if consciousness is not in one of the two modes it must be in the other.[22]

These are paradigm cases of psychologistic claims. Longstanding principles of logic are construed as claims about psychological phenomena, and their justification is taken to be on a par with other justification in psychological inquiry: inductive generalization from experience.

This sort of thing eventually fell to sustained (and often vituperative)[23] criticisms from Frege and Husserl, among others, which criticisms will be raised in detail in §4. But to paint them quickly, with a broad brush: psychologism mistakenly makes logic

[21](Mill 1843, p. 205). In this section of the book, Mill is disputing Hamilton's claim that such fundamental logical principles are true of not just Phenomena, but also Noumena, of which we are otherwise quite ignorant.

[22](As cited in Mill, 1843, p. 206)

[23]For example, here is Husserl's reaction to Mill's psychologistic account of noncontradiction: "Where the fundamental principles of his empiricistic prejudices are at stake, all the gods seem to abandon Mill's otherwise keen intelligence. Only one thing is hard to understand: how such a doctrine could have seemed persuasive" (Husserl 1900/2011, pp. 56–57). Regarding such vituperation, Husserl, who began his philosophy of mathematics as a psychologistic thinker, says "As regards my frank critique of the psychologistic logic and epistemology, I have but to recall Goethe's saying: There is nothing to which one is more severe than the errors that one has just abandoned" (*ibid.*, p. 3).

beholden to contingent, individual mental events, and as a consequence gets all the important stuff about logic—its necessity, universality, normativity, basicness—wrong. Says Frege:

> [P]sychology should not imagine that it can contribute anything whatever to the foundation of arithmetic.[24]

Logic is the foundation of arithmetic, according to Frege's logicism. And psychologism's key mistake is that it confuses what people *take* to be true with what really *is* true.

4.2 scorekeeping and psychology

As outlined in Chapter 2, Brandom's general project in *Making It Explicit* is to explain the meaningfulness of thought and speech by beginning with a primitive notion of inferential correctness implicit in social practice, and working from there to a derived notion of representational correctness. To understand his expressive conception of logic, the key part of this project is the account of the inferential foundations: the game of deontic scorekeeping.

To briefly recap: Brandom, following David Lewis (1979), thinks about what we do when using language as keeping a kind of score on one another. For example, while in conversation you might utter some sounds, and in response I update your score—you're now *committed* to the claim that p. But this piece of scorekeeping might have further effects; I might think that, once you're committed to p you're thereby committed to q, or perhaps *entitled* to r, whether or not you've uttered something that means q or r. Updating your score in these ways is what counts, in Brandom's system, as making inferences. When I attribute commitment to q on the basis of commitment to p, it's because I'm endorsing the inference from p to q.

[24](Frege 1884/1980, p. VI)

Thus the model of deontic scorekeeping is a kind of coarse-grained theory of reasoning. Although it is first introduced as an account of what we're doing when we interpret each others' language behavior, it is at the same time an account of the sorts of inferences we make from the claims we encounter. But the explanatory apparatus of the scorekeeping model is couched entirely in terms of the *attribution* of deontic statuses (commitment and entitlement), which is itself understood as a special sort of *attitude*. So reasoning turns out to be a matter of the adoption of attitudes, and this is where psychologism gets its grip.

That's because for Brandom, while the meaningfulness of all talking and thinking comes from our keeping score on one another, the special role of *logical* vocabulary is to express the attitudes that constitute the scorekeeping practice itself.[25] This account has been explained more fully over the last two chapters, but to take one example: the conditional has the expressive role of making explicit the sorts of inferential moves we scorekeepers can make. In a language with a conditional operator, I am able not only to update your score inferentially, but also to say how I'm updating it, in the form of a claim which can be further endorsed and disputed.[26]

Thus on Brandom's conception of logic, the role of logic is to express the attitudes constitutive of reasoning. Though Brandom never refers to them in this way, these attitudes are naturally construed as psychological states—they are the scorekeeping counterpart of the traditional propositional attitudes. But the model of deontic scorekeeping is both an account of *interpretation*, and also attempts to say what we are *doing* when we interpret. This is a nonstandard approach in the philosophy of mind, since it accounts for what it is to have a belief, or to make an inference, in terms of what it is to *interpret*

[25]Hence Brandom's slogan that "logic is the organ of semantic self-consciousness."

[26]This benefit becomes especially clear where negation is involved: if you argue from p to q, I might want to dispute p, q, or the inference itself. This last I'd do most readily by saying "Not: if p then q," but without the conditional this can't be done.

someone as having a belief, or as having made an inference.[27] And Brandom understands individual acts of interpretation as themselves constituted by other interpretations—it's interpretations all the way down, so to speak. This sounds like a troublesome regress, and it would be if Brandom were after an account of belief (or whatever) as a special ontological kind, but he isn't; he's after an account of what we're *doing* when we interpret each other as having beliefs (etc.). The basic idea that leads to regress is that interpreting someone in this way itself depends on being recognizable as having done so; acts of interpretation can take place only within a whole practice of interpretations.

Implicit in the idea of *recognition* is the fact that Brandom's conception of the mental is a normative one. The operative question is not "What's a person's internal configuration when they believe that *p*," but rather "What *counts* as believing that *p*," which on Brandom's interpretive approach amounts to "When is a person *correctly interpreted* as believing that *p*?" The standards of correctness are then deferred to further interpretations. This does not mean that internal configurations—brain states or activation patterns or sensorimotor primings or what have you—are irrelevant; on the contrary, they are pieces of the physical infrastructure required for exhibiting the behavior which gets correctly interpreted as believing or asserting. But being in such configurations is not only insufficient for having the belief that *p*; it is the wrong category of property entirely. To use an analogy of Brandom's (2002, p. 361), believing that *p* is like being legally capable of signing a contract. There is no special ontological difference between a person's signature at 17 years of age and that same signature at 18—it's still just some ink on paper—but rather a normative difference in the significance of that signature. Once a person is of legal age, their signature *counts* as legally binding, and hence acquires all sorts of consequences in the form of how it is correct to treat the signatory (in particular, about what they have committed themselves to by signing).

[27] One other conspicuous outlier in this regard is Dennett (1971, 1987, 1991), who understands belief in terms of adopting the intentional stance.

All this means that having the attitudes required for scorekeeping is not merely a matter of internal configuration, and hence is quite unlike the sorts of psychological properties appealed to by classical empiricist philosophy of mind, 20th-Century intentional realism, and psychology and the other sciences of mind after the cognitive revolution.Rather, like beliefs, the scorekeeping attitudes are normative statuses which depend for their existence on the interpretations of other scorekeepers. Indeed, the very status of *being a scorekeeper* is a normative status which depends on being (correctly) interpreted.

So the kind of psychologism Brandom's view exemplifies is interestingly different from the classical kind. Whether the differences are enough to avoid the pitfalls of classical psychologism is the subject of the next section.

4.3 classic objections to psychologism

In this section I consider several important objections to psychologism, most of which were raised by Frege and then again later by Husserl. After raising each objection I offer a response on behalf of Brandom, showing how the expressive conception of logic makes for a viable form of psychologism.

4.3.1 the shakiness of psychology

Both Frege and Husserl are concerned that, while logic and mathematics are paradigmatically secure sciences, psychology is somehow vague or shaky, and hence cannot be the foundation for logic, nor can logic be a special part of psychology.

> [I]f everyone had the right to understand by this name [i.e. 'one'] whatever he pleased, then the same proposition about one ["the number one is a thing"] would mean different things for different people,—such propositions would

> have no common content.[28]
>
> When STRICKER, for instance, calls our ideas of numbers motor phenomena and makes them dependent on muscular sensations, no mathematician can recognize his numbers in such stuff or knows what on earth to make such a proposition. An arithmetic founded on muscular sensations would certainly turn out sensational enough, but also every bit as vague as its foundation. No, sensations are absolutely no concern of arithmetic. No more are mental pictures, formed from the amalgamated traces of earlier sense-impressions. All these phases of consciousness are characteristically fluctuating and indefinite, in strong contrast to the definiteness and fixity of the concepts and objects of mathematics.[29]
>
> [O]nly vague rules could be based on vague theoretical foundations. If psychological laws lack exactness, the same must be true of the prescriptions of logic. It cannot be doubted that many of these prescriptions are infected with empirical vaguenesses. But precisely the laws which are pointedly called 'logical' ... are of absolute exactness. Every interpretation that would base them on empirical indefinitenesses, make them depend for their validity on vague 'circumstances', would fundamentally alter their true sense. Plainly they are genuine laws, and not 'merely empirical', i.e. approximate, laws.[30]

But security in this context can mean two different things, as these quotations illustrate. Logic and mathematics[31] are *epistemically* secure in the sense that our justifications for logical or mathematical statements come by way of proof, which provides *certainty*, at least relative to a choice of axioms. Indeed, Frege's official rationale for his new logic—expressed in all three of his logicist treatises—is to eliminate all "gaps" in mathematical proof, and to return its standards of rigor to those exemplified by Euclid.[32] Psychology, on the other hand, has to content itself with inductive generalizations which hold only for the most part. This was especially true on Frege's empiricist conception of psychology,

[28](Frege 1884/1980, p. I)

[29](Frege 1884/1908, pp. V–VI). Footnote excised following 'STRICKER'.

[30](Husserl 1900/2001, pp. 46–47)

[31]In the present context, I am going to switch pretty freely between talking of logic and talking of mathematics, since for Frege they are inextricably linked—not only because of his logicism, but because he sees mathematical proof as held to the standards set by logic. See the introductions to (Frege 1879/1967) and (Frege 1884/1980). Similarly, Husserl holds pure arithmetic and pure logic to be on a par, philosophically speaking, and even calls pure arithmetic logic's "mature sister discipline" (Husserl 1900/2001, p. 109).

[32]See Frege (1879/1967, pp. 1–2), (1884/1980, p. IX), (1893/1964, p. 2).

where the inventory of mental items was limited to sensations and Ideas. A logic founded on psychological notions thus could not underwrite the certainty required of proof—that's the epistemic concern.

The deeper concern about psychology's shakiness is the *semantic* one. In mathematics, for example, each of us needs to mean the same thing by '1', so that we can be talking about the same thing when we argue over whether, say, 1 should be considered prime. Indeed, the meanings of mathematical terms must be stable not only interpersonally but intrapersonally, since even a single person's understanding of a proof requires that a term mean the same thing from beginning to end. Psychologism, on Frege's conception of psychological entities, would seem to fail on both counts: there could be no interpersonally stable meanings, since each person's ideas are theirs alone, and there could be no intrapersonally stable meanings, since the stream of ideas is constantly in flux.

On Brandom's expressivism, however, the special attitudes which provide the meanings of logical terms are not like empiricism's Ideas. Their significance is not "fluctuating" or "indefinite," but rather well-defined against the background of scorekeeping, and this definition provides both epistemic and semantic stability. To take the familiar example, a conditional formed with 'if' always expresses endorsement of the inference from antecedent to consequent. Because this attitude of endorsement is a fundamental component of scorekeeping, it is something all reasoners can take up (that is, it is interpersonally available) as well as something a single reasoner can do repeatedly (that is, it is intrapersonally available). This semantic stability in turn supports an account of deductive inference which is quite close to the standard Tarskian account: an inference is logically valid if uniform substitutions of its nonlogical vocabulary invariably yield good inferences. *Modus ponens*, for example, satisfies this condition on account of the expressive role of 'if'. So Brandom's expressivism can secure something like the

certainty typically demanded of logical inference, on account of the logical terms having well-defined roles in scorekeeping, i.e. reasoning. Hence the sort of psychologism exhibited by Brandom's view does not entail the sort of shakiness that worried Frege and Husserl.[33]

4.3.2 normative and descriptive logical laws

Both Frege and Husserl thought that psychologism confused the normative and descriptive aspects of logic. Psychologistic philosophers think that because logic offers laws for thought, it is a descriptive science of thinking (Frege 1893/1964, p. 12). Or they think that because logic offers prescriptions, it is merely a technology—and since it is a technology concerned with psychological matters such as judging, conceiving, and inferring, its theoretical basis must lie in psychology (Husserl 1900/2001, §18 and §19). But according to Frege and Husserl, the same mistake underlies both views, namely, a failure to see that logic's normative force comes from its status as a descriptive science of a proprietary, non-psychological subject-matter. Here, for example, is Frege:

> Any law asserting what is, can be conceived as prescribing that one ought to think in conformity with it, and is thus in that sense a law of thought. This holds for laws of geometry and physics no less than for laws of logic.[34]
>
> I understand by 'laws of logic' not psychological laws of takings-to-be-true, but laws of truth. . . . If being true is thus independent of being acknoledged

[33]There is a possible concern here about *vagueness*. Frege, who construed concepts as mathematical functions which map their arguments to a truth-value, thought concepts needed to have sharp boundaries, at least for mathematical purposes. But vague concepts like *bald* don't have sharp boundaries, and hence don't map every possible argument to a truth-value. (If 'a' denotes someone only partially bald, then plausibly '$bald(a)$' does not map either to True or to False.) Is logical vocabulary vague in this way, on Brandom's view? The answer is certainly 'no' for the logical constants of formal languages, but these are expressions given exact specifications in the construction of artificial languages. Logical vocabulary as it occurs in natural languages is another matter (recall §3.5.1 above). Ordinary 'if' certainly wears the appearance of vagueness, in the sense that it could be indeterminate, for some choices of P and Q, whether 'if P then Q' is true or false. But even if this appearance is correct, it shouldn't worry Frege, since what ordinary 'if' expresses—the endorsement of an inference—is precisifiable, where such precisification captures a feature of existing linguistic practice but eliminates vagueness.

[34](Frege 1893/1964, p. 12)

> by somebody or other, the the laws of truth are not psychological laws: they are boundary stones set in an eternal foundation, which our thought can overflow, but never displace. It is because of this that they have authority for our thought if it would attain to truth.[35]

And similarly Husserl:

> We must first put an end to a distorted notion which both parties share, by pointing out that logical laws, taken in and for themselves, are not normative propositions at all in the sense of prescriptions, i.e. propositions which tell us, as part of their *content*, how one *should* judge. ... Anyone who judges that every *A* is also *B*, and that a certain *S* is *A*, ought also to judge that this *S* is *B*. Everyone sees, however, that this proposition is not the original proposition of logic, but one that has been derived from it by bringing in the thought of normativity.[36]

> What we have said here in regard to pure arithmetic [namely, that it is not concerned with psychological acts like counting and adding, but rather with "ideal singulars"—the numbers] carries over at all points to *pure* logic. In the latter case, too, we accept as obvious the fact that logical concepts have a psychological origin, but we deny the psychologistic conclusion to which this seems to lead. ... We deny that the theoretical discipline of pure logic, in the independent separateness proper to it, has any concern with mental facts, or with laws that might be styled 'psychological'. We saw that the laws of pure logic, e.g. the primitive 'laws of thought', or the syllogistic formulae, totally lose their basic sense, if one tries to interpret them as psychological. It is therefore clear from the start *that the concepts which constitute these and similar laws have no empirical range*. ... [T]hey must be notions truly *generic, whose range is exclusively one of ideal singulars, genuine species*.[37]

Although the negative arguments from Frege and Husserl against grounding logic in psychology are correct as far as they go, their conclusion that logic must be a descriptive science with its own "peculiar realm of truth" is too hasty.[38] If we can articulate a conception of logic which accounts for its normative status in a psychologistic fashion, without succumbing to absurd consequences, then we have a viable psychologism. I

[35](*Ibid.*, p.13)

[36](Husserl 1900/2001, p. 101)

[37](*Ibid.*, pp.110–111)

[38]Husserl's claim that logic's truths belong to its own "peculiar realm" is made in §20 of (1900/2001, p. 45).

will address some of the alleged absurd consequences, such as Husserl's charge that interpreting logical laws psychologistically destroys their "basic sense," in later sections of this chapter. In the rest of this section, I aim to show that Brandom's view lacks the problematic results for logic's normativity which are alleged to follow from psychologism. To do so I will begin by examining Frege's remarks, since it is Frege who says more about psychology, and who provides the clue for a viable psychologism.

For in the first place, Frege does *not* claim that psychology is not normative. In his lengthy argument against psychologism in the *Grundgesetze*, he claims that psychology can be normative in two ways (Frege 1893/1964, pp. 12–13). First: like any other descriptive science, psychology can be conjoined to the general norm that one ought to judge in accordance with the facts. Then, since psychology tells us facts about the mind, it will provide the norms for how we ought to think about these psychological matters. Second: psychology can also provide norms for thought

> in the sense that they give an average, like statements about 'how it is that good digestion occurs in man', or 'how one speaks grammatically', or 'how one dresses fashionably'.[39]

Frege's meaning here is obscure, but one way of reading him is as claiming that, in general, statements of typical or average human behaviors can serve as rules of thumb. If the average person takes a walk after a meal, say, then that fact could be useful to someone trying to achieve good digestion, although it certainly doesn't guarantee the effect.[40] Likewise, if the average person comes to believe instances of "All *A* are *B*" after observing just two *A*s which are *B*, then that fact could be useful to someone trying to regulate their inductive inferences, although it certainly doesn't guarantee that all such

[39]Ibid.

[40]This is not the only way to read Frege on this point, and it's not obviously the best one (matters are made worse by his choice of examples—following the average might well produce ungrammatical speech, and will almost certainly produce unfashionable dress). But other candidate readings (for example, that Frege thinks an average could be *constitutive* of correct behavior) seem equally problematic, and none of them will give psychology a stronger role in determining norms of thought, so far as I can tell.

inferences will be correct.

Neither kind of norm, however, can do the job that logical norms are supposed to do: to "prescribe universally the way in which one ought to think if one is to think at all" (*ibid.*, p.12). The first kind of norm tells us how we ought to think only when the subject-matter is psychological; hence it is too restricted in scope to serve as a law of thought in the logical sense. The second kind of norm applies to thinking in general, but is only a rule of thumb, and hence is too weak to serve as a law of thought. The psychologistic fallacy here, then, is to think that psychology can provide anything more than these two types of norm.

Now it is sometimes thought that Frege held psychologism to be a mistake because psychology is descriptive, and logic must be normative for thought. But we should see that for Frege, psychology is not disqualified from the foundations of logic *merely* in virtue of its status as a descriptive science. Indeed, recall that Frege himself thinks that logical laws are in the first place descriptive, as shown by his analogy with geometry and physics. The real problem with psychology as a grounding for logic is that it fails to provide the norms for thought of the right scope and strength. But if there is a form of psychologism which relies on a conception of psychology different from Frege's, then it could by Frege's own lights provide norms for thought as such.

One way to do this is to follow Frege himself in giving logic a *descriptive* subject-matter with wide enough scope that the norm 'judge in accordance with the facts' will provide us with normative laws for thought as such. Frege's way of pursuing this line was to posit the "third realm" of Thoughts, the facts about which we can ascertain by analysis into concept and object (Frege 1918/1956, p. 302). Once we know the facts about which Thoughts entail which others, then the norm 'judge truly' yields logical laws.

The other way to update psychologism is to take seriously Frege's suggestion that psychology provides norms "in the sense that [it gives] an average," and make the case

that when we properly understand this kind of norm, we'll see that the laws of thought Frege seeks could be nothing else.

Brandom's conception of logic incorporates elements of both of these strategies. In the model of deontic scorekeeping, Brandom has offered an account of what it is to think which is descriptive of thought as such, and logic articulates the possible inferential relationships that can hold between thoughts. But deontic scorekeeping also makes good on Frege's analogy to grammatical speech, and thus essentially involves norms in the second way.

On the first point, consider the theoretical role the deontic scorekeeping model is supposed to play. It is a model of what we're doing when interpreting each others' language behavior, but it is also an account of what we must be able to interpret others *themselves* as doing in their social transactions, such that they are language-users at all. To put the point in a way Brandom frames it late in *Making It Explicit*, to see other beings as trafficking in conceptual content (and not merely derivatively so, in the way that, say, books do), we need to be able to take up the "explicit discursive scorekeeping stance" toward them (1994, pp. 639–643). This means being able to interpret others as keeping deontic score on themselves and one another, which means taking them to adopt the deontic attitudes of commitment and entitlement. So those deontic attitudes, and the socially articulated combinations derivable from them, form a necessary condition on the very possibility of thought and speech; anything which doesn't adopt these attitudes doesn't count as a thinker. The model of deontic scorekeeping is, we might say, transcendental psychology.

As a consequence, the elements of the model of deontic scorekeeping are common to all thought as such. Moreover, they pertain not to the content but to the pragmatic *form* of thinking, as discussed in §5.1 of the previous chapter. So right away we should see that, since the expressive conception of logic is based on the form of what is common

to thought as such, and not on empirical details of what individual thinkers think or how they reason, it is a type of psychologism different from what Frege had in mind. Indeed, the structure of thought which deontic scorekeeping makes possible—not only claims standing in inferential relationships, but distinct subsentential components like singular terms, predicates, and relations—corresponds closely to the structure Frege himself discerns in sentences and thoughts.

But note, too, that the psychological facts of deontic scorekeeping are normatively laden. When a scorekeeper is committed to some claim, that fact carries normative import for what else they ought to do and believe. This normative import isn't incidental to the commitment, either, as e.g. the normative import of pain is (pain is often bad and to be avoided, but not when it is a necessary part of achieving something valuable, as in dentistry or athletic training). Rather, the commitment is itself individuated by its normative connections to other deontic attitudes, namely those inferentially related to it. So the kind of psychological fact at issue in Brandom's expressive conception of logic is in the first place already normatively significant. What logical vocabulary expresses is just these normative connections between the possible contents of thought, and so the sort of psychologism yielded by the expressive conception has a fair claim to "prescribe universally the way in which one ought to think if one is to think at all."[41]

One way to summarize the conception of psychology on which this psychologism rests is that it is concerned with *normative structural psychological facts*. And this makes Frege's analogy to "how one speaks grammatically" quite apt. For the linguist who attempts to determine a language's grammar is in a similar position vis-à-vis her subjects to the deontic scorekeeper and his interlocutors. No matter what sort of methodology the linguist uses, normative facts are what she aims to adduce in the construction of her theory. For example, generative grammarians in Chomsky's tradition test their subjects'

[41] Exactly what logic's normative status *is* is the subject of the next chapter.

(often their own) judgments of syntactical *correctness* for sentences of a language; anthropological linguists like Boas asked subjects to produce what they themselves took to be *correct* utterances under some circumstances; and even linguists who mine corpora must rely on the underlying data being purged of errors and organized by, say, parts of speech, which themselves rely on judgments of correctness by those tagging the corpus. And all researchers in these traditions decide which data are worth examining by first deciding who *counts* as a competent speaker or writer of the language under study, itself a normative judgment.

So the attempt to discern "how one speaks grammatically" from the empirical study of language turns out to be, at least in part, the study of normative linguistic facts. And a linguist who produces a grammar for a language is making explicit the underlying principles which explain why competent speakers produce the linguistic performances they do. By analogy, logic on Brandom's expressive account makes explicit the underlying principles which explain why competent scorekeepers—that is, thinkers and speakers of any language—count as trafficking in conceptual content at all. The linguist might produce "laws" of a language L along the lines of "Sequences of such-and-such syntactic form, or transformations thereof, are correct sentences of L." And likewise the Brandomian logician might produce logical laws such as "Inferring Q from P, when Q is incompatible with P, is an incorrect scorekeeping move."

Thus although Frege was right to think that psychologism, in its reliance on psychology, might produce norms along the lines of how one speaks grammatically, he was wrong to think that this kind of norm is the wrong kind for logic.

4.3.3 reference to psychological objects

In §23 of the *Logical Investigations*, Husserl argues that psychologism has the absurd consequence that logical terms refer to psychological objects:

> [I]f the laws of logic have their epistemological source in psychological matter of fact, if, e.g., as our opponents generally say, they are normative transformations of such facts, they must themselves be psychological in content, both by being laws for mental states, and also by presupposing or implying the existence of such states. This is palpably false. No logical law implies a 'matter of fact', not even the existence of presentations or judgments or other phenomena of knowledge.[42]

Although directed at mathematics rather than logic, Frege has a similar complaint when he says, more colorfully,

> What, then, are we to say of those who, instead of advancing this work where it is not yet completed, despise it, and betake themselves to the nursery, or bury themselves in the remotest conceivable periods of human evolution, there to discover, like JOHN STUART MILL, some gingerbread or pebble arithmetic! It remains only to ascribe to the flavour of bread some special meaning for the concept of number.[43]

Grounding logic in psychology, or more generally in anything empirical, would seem to entail that logic is somehow about that empirical stuff. But on Brandom's conception of logic, logical terms aren't referring terms at all. They express—but do not refer to—scorekeeping attitudes, so the basic version of this objection to psychologism doesn't affect the expressive conception.

In order to understand how, on Brandom's view, logical terms can express attitudes without referring to them, we need to distinguish semantic theory from what Brandom would call "philosophical semantics," or what is sometimes called *metasemantic* or fundamental semantic theory.Semantic theory is about what the meanings of particular expressions are; metasemantic theory is about how expressions get meanings in the first place.

This distinction might be made clearer by way of another analogy to metaethical expressivism. In metaethics, expressivism is properly a *metasemantic* view according to

[42](Husserl 1900/2001, p. 51)
[43](Frege 1884/1980, p. VII)

which certain mental states—typically attitudes of approval and disapproval—explain *why* moral terms like 'wrong' mean what they do. So conceived, expressivism is compatible with 'wrong' having any particular semantic value, e.g. the utility-non-maximizing function or the set of all rights violations. That is, expressivists are entitled to be utilitarians, deontologists, or whatever at the level of normative moral theory, since nothing about expressivism as a metasemantic view settles what makes particular acts wrong. Construed as a metasemantic view, expressivism competes with things like divine command theory, which explains what 'wrong' means in terms of God's commands. But divine command theory also settles nothing about which particular acts are wrong; God could command all sorts of things, and we need additional premises to get any information about what is in fact wrong.

But expressivism has sometimes been construed as a *semantic* view, on which the meaning of a moral term is itself the mental state expressed—'murder is wrong' has as its interpretant disapproval of murder, so that when I exclaim that murder is wrong, what I'm doing with my utterance is *not* in fact claiming something about murder, but merely expressing something about myself. What's wrong, on this construal, is what I disapprove of. Thus on this first-order semantic construal, expressivism competes with first-order normative ethical theories, which assign such meanings to 'wrong' as the set of acts which fail to maximize utility.[44]

It is tempting, and perhaps commonplace, to think of expressivism as a first-order semantic view. But it is going to be crucial in what follows to distinguish the semantic from the metasemantic, and in particular to explore the viability of the metasemantic view. This is because it is the metasemantic view which affords the possibility of a propositional, truth-apt semantics.

[44]Williams (1972, pp.16–18) is one who construes expressivism as a semantic view, and objects to it as a version of noncognitivism.

First-order semantic expressivism is typically conjoined to (or even considered a form of) noncognitivism, so that the mental states expressed are noncognitive things like attitudes of approval. Since the semantic interpretants are just these expressed mental states, it is a short step to conclude that expressive construals of discourse make sentences "neither true nor false." So first-order semantic expressivism seems to close off the possibility of construing a discourse as trafficking in truth, or as expressing propositions. If the vocabulary to be given an expressivist treatment is *logical* vocabulary, we could be in big trouble—since any claims whatsoever can be compounded using logical terms, an expressivist logic would seem to entail that *no* discourse is truth-apt except that conducted via logically atomic sentences. This is not, perhaps, a result that no one would endorse, but it is surely a radical one.

But there are further reasons to be wary of expressivism as a first-order semantic view. For one thing, it is in essentials the same view as Ayer's emotivism, and is thus susceptible to the same problems which leveled that view, such as the Frege-Geach problem. For another, it exhibits a kind of theoretical myopia akin to phenomenalism about the external world, or instrumentalism about science: it treats the stuff we use as *evidence* for the phenomenon in question as exhaustive of its nature. In the case of expressivism in ethics, this is to take the attitudes associated with moral judgment to constitute the meaning of moral claims. And just as we need not be phenomenalists (sensory input can be evidence for the real external world), nor instrumentalists (observed entities can provide evidence for real theoretical entities), we need not be first-order semantic expressivists—the attitudes associated with normative claims can *explain* what normative language means without thereby *constituting* that meaning.

What presents an immediate puzzle in interpreting Brandom's view is the fact that he says things like "conditionals express endorsement of an inference," or "negation expresses minimal incompatibility," which make it sound as if he's advancing an

expressive account in its first-order semantic form. But that isn't right.

The key to understanding Brandom's view is to ask how it's possible for him to claim inspiration from both Gentzen and Sellars. From Gentzen (1935), Brandom takes the idea that the meaning of a term is its inferential role, and in particular its introduction and elimination rules. From Sellars (1953), Brandom takes the idea that the distinctive meaning of logical terms includes the "conveying" of an underlying state of mind or attitude. These two semantic ideas, inferential and expressive, are not obviously compatible.

The way to synthesize them is that, regarded broadly enough, inferential role is what fixes the meaning of logical terms.[45] It's at this point that Brandom's recruitment of Dummett becomes helpful. Instead of thinking of inferential role specifically in terms of introduction and elimination rules, Dummett thinks generally in terms of appropriate circumstances for, and consequences of, the use of a term (Dummett 1973a). What in Gentzen's framework has the form of a premise from which a conclusion may be drawn, has in Dummett's framework the more general form of an inferential circumstance which licenses the use of a statement. (Likewise for consequences.) For logical vocabulary, the circumstances in question are certain scorekeeping situations; employing the conditional 'if p then q', for example, is appropriate in circumstances where one endorses the inference from p to q. And the inferential consequences, of course, are things like the permission (or obligation) to endorse q when you also hold that p. This structure preserves Gentzen-style inferential role semantics, but also allows logical terms to convey underlying commitments—'if p then q' doesn't outright *say* that the speaker endorses the corresponding inference, but because it is appropriately made under circumstances of such endorsement, it can be said to *convey* that information.

[45]Indeed, "strong inferentialism" is the name Brandom gives his view, which is distinguished from narrower forms of inferentialism exactly by its ability to countenance non-propositional circumstances and consequences of inference (1994, pp. 131–32).

This is the sense in which logical terms express basic scorekeeping attitudes—not as a first-order semantic view, in which the semantic values of logical terms are themselves those attitudes, but rather in the sense that these attitudes determine the conditions for appropriately using logical terms. Thus the expressive conception does *not* entail that logical terms refer to psychological objects or facts; logical terms don't refer at all, and even the attitudes they express don't constitute the meanings of the terms. (And, further, there is no danger of being forced to offer psychological interpretations of logical laws, à la Mill.)

4.3.4 truth and taking-to-be-true

Both Frege and Husserl objected that psychologism, because it treats the subject-matter of logic as the actual thoughts had by concrete thinkers, would entail that logic consists of statements about what people merely *take* to be true. But logic is about what is actually true, or about the thinker-independent laws of truth, hence psychologism is false. Frege is more forthright than Husserl on this point:

> Of course [the laws of logic will be variable and restricted]—if logic has to do with something's being taken to be true, rather than with its being true! Thus Herr B. Erdmann in the first volume of his *Logik* . . . equates truth with 'general validity', and bases this in turn upon 'general agreement among the subjects who judge'. Thus in the end truth is reduced to individuals' taking something to be true. All I have to say to this is: being true is different from being taken to be true, whether by one or many or everybody, and in no case is to be reduced to it. There is no contradiction in something's being true which everybody takes to be false. I understand by 'laws of logic' not psychological laws of takings-to-be-true, but laws of truth.[46]

A notable feature of this argument is that it is, if not exactly question-begging, one that gets no grip on the opponent. As Husserl acknowledges, there is a dialectical impasse here: "One cannot persuade the subjectivist any more than one can the open sceptic, a

[46](Frege 1893/1964, p. 13). Footnote excised following '*Logik*'.

man simply lacking the ability to see that laws such as the law of contradiction have their roots in the mere meaning of truth . . . " (Husserl 1900/2001, p.78). However, to the extent that we are not arrant "subjectivists" who insist that there is no truth over and above what each person takes to be true, there is dialectical pressure on us to reject psychologism, or articulate a form of it which isn't vulnerable to Frege's and Husserl's argument.

Is the expressive conception vulnerable to this argument? I think the answer is no, but showing that this is the case is difficult, particularly since Brandom's systematic philosophy could fairly be said to reduce truth to what scorekeepers take to be true, though that's putting things glibly. In order to see why the expressive conception nonetheless dodges this particular objection, I will first need to say a bit about how Brandom construes truth, and then consider broader issues of objectivity, correctness, and conceptual content. But the basic idea is as follows: In Brandom's systematic theoretical philosophy, there is a bootstrapping from what scorekeepers *take* to be true to a fine-grained discursive practice that enables an objective responsibility, within the practice, to the world. Scorekeepers are thus able to formulate claims (including claims which make use of logical vocabulary), the truth of which are independent of what any particular scorekeeper takes to be true. This is so even though it's nonetheless correct to say, of Brandom's picture, that it is only the taking-true which makes objective truth possible.

an expressive account of truth

In Brandom's systematic theoretical philosophy, traditional semantic vocabulary is a species of expressive vocabulary. This includes the predicate 'is true', along with 'refers' and 'denotes', and their philosophical cognates 'applies to', 'is satisfied by', 'falls under', and so on. The mechanism by which these terms do their expressive jobs is anaphora—the usually unremarkable linguistic phenomenon by which speakers can link the significances of two token expressions, as in

The world was moving and she was right there with it

where the pronoun ‘it’ has the significance of ‘the world’. In this example, ‘the world’ is the anaphoric antecedent, and ‘it’ is the anaphoric dependent, because the significance of ‘it’ is dependent on the significance of ‘the world’. This antecedent–dependent structure is common to anaphoric constructions generally, such as the verb and pro-verb in

She had blue skin, And so did he

where ‘so did’ is dependent on ‘had blue skin’. A striking philosophical proposal, due to Grover, Camp, and Belnap (1975), and taken on by Brandom, is that the predicate ‘is true’ is an anaphoric dependent—in particular, a prosentence-forming operator, which generates sentences the significance of which depends on other sentences. For a simple example, take this dialogue:

ABEL: It’s hot today.
MABEL: That’s true.

What Mabel has done by employing the truth predicate, along with demonstrative ‘that’, is produced a sentence which has the significance of ‘It’s hot today’, though without uttering any of those words. Her assertion is instead anaphorically dependent on Abel’s.

This prosentential theory of truth, according to which the truth-predicate enables the expression of whatever content is possessed by an anaphoric antecedent, entails that truth is not a property, and the truth-predicate doesn’t refer to a property. We can of course say that a sentence is true, but in doing so we aren’t ascribing a property to it—we’re endorsing its content. So Brandom’s view can’t slide down the slippery slope Frege attributes to Erdmann, at the bottom of which truth turns out to be general agreement. Truth isn’t a property at all, so there’s no danger that it will turn out to be general agreement, or any other property.

truth and objectivity

Neither Frege nor Husserl would likely be impressed by this reply, since—even though it secures the narrow result that assertions of '*P* is true' don't *mean* 'I take *P* to be true', and there's a general reason why truth isn't reducible to any property of being-taken-to-be-true—the prosentential theory of truth as it appears in Brandom's philosophical system accounts for the meaning of the truth predicate in terms of its expressive function, and not, conspicuously, in terms of any sort of discourse-independent standard which our speech and thought aims to meet.

It seems to me that what's really worrying Frege and Husserl isn't truth per se, but rather objectivity. Isn't it the case that there are mind-independent facts of the matter, which we can get right or wrong? As Frege puts it, "there is no contradiction in something's being true which everybody takes to be false." Any theoretical account of language which entails otherwise runs contrary to deeply held convictions about the objectivity and discourse-independence of the world.

The real threat to objectivity on Brandom's view, given the expressive role of the truth predicate, would seem to be not the danger of conflating truth with what *everyone* takes to be true, but rather with what each individual scorekeeper takes to be true. For when I claim that *P* is true, I am endorsing via anaphoric mechanisms the claim that *P*; and when I endorse *P* then, by the expressive role of 'is true', I'm thereby committed to '*P* is true'. It would then seem that, since both conditionals hold, their biconditional equivalent '*P* is true if and only if I claim that *P*' holds also. And now truth is just whatever I claim.

What is needed is a way to show that neither conditional in fact holds, and the way to do it is to provide counterexamples. Brandom does this himself near the end of *Making It Explicit*,[47] labeling the two threatening conditionals the *No First-Person*

[47](1994, pp. 604–607); a similar argument is offered in (Brandom 2000), pp. 198–204.

Ignorance Condition:

$$(P)[P \rightarrow (\text{I claim that } P)]$$

and the *No First-Person Error Condition*:

$$(P)[(\text{I claim that } P) \rightarrow P]$$

A counterexample to either conditional will take the form of a claim which is compatible with the antecedent but incompatible with the consequent (as we might otherwise put it, something which makes the antecedent true and the conclusion false). For the No First-Person Ignorance Condition, such a counterexample would be 'I don't claim that *P*', which is incompatible with 'I claim that *P*' but not with *P* itself. For example, take *P* = the Moorean claim 'It's raining'. This is the antecedent of the conditional, and it's compatible with 'I don't claim that it's raining', since obviously it could be raining and yet I don't claim that it is. But in this case the consequent of the conditional, 'I claim that it's raining', is incompatible with 'I don't claim that it's raining', and so we have a counterexample to the conditional.[48]

For the No First-Person Error Condition, a counterexample is provided by any *Q* which is incompatible with *P*, in which case it's (by hypothesis) incompatible with the consequent, but compatible with the antecedent. Suppose, for example, that I'm talking to a friend about planning a picnic for Saturday, but have forgotten that yesterday I checked the forecast to discover that it will rain all weekend, and we agree to hold the picnic. Let *P* = 'I will enjoy a picnic on Saturday', and let *Q* = 'It will rain on Saturday'. Then I am committed to *P* and to *Q*, and to 'I claim that *P*', since I just made the plan with my friend. My commitment to *Q*, however, undermines my entitlement to *P*, though I haven't realized it yet. So we have a counterexample to the No First-Person Error Condition:

[48]Of course, *I* can't be entitled to both *P* and 'I don't claim that *P*, but the contents themselves are compatible, as shown by the fact that someone else could consistently ascribe to me commitment to both.

something compatible with the antecedent but incompatible with the consequent.[49]

What these counterexamples show is that the structure of deontic scorekeeping, including its social articulation, makes it possible for us to express conceptual contents which show that in general, the content of *P* is not equivalent to the content of 'I claim that *P*'. Hence there is no contradiction, on Brandom's account, in supposing that *P* is true even though I believe it to be false, or vice versa. This goes for *P* = some law of logic as well as for more ordinary, humdrum choices.

Admittedly, this result doesn't rely at all on the claim that logic tells us the "laws of truth," as Frege would have it; and given the expressive conception it doesn't make much sense to call logic the laws of truth anyway. So perhaps Frege would still not be satisfied. Regardless, if Frege's real concern is for preserving the objectivity of our discourse, and not something to do peculiarly with truth, then he *ought* to be satisfied, for the counterexamples above show that there is in-principle reason why truth doesn't collapse into taking-true.

4.3.5 species relativism and universality

In the *Logical Investigations*, Husserl is concerned that there is further threat posed by psychologism, when we consider that it seems to make logic relative to the human species as a whole:

> If, however, instead of such a subject [the individual judging subject], we make some contingent *species* of judging beings the pivot of our relations, we achieve a new form of relativism. Man as *such* is then the measure of all human truth. Every judgement whose roots are to be found in what is *specific* to man, in the constitutive laws of man as species – is a true judgement, for us human beings. ... It is best to employ the term 'relativism', and to

[49] If this example sounds implausible, in that it's hard to imagine simultaneous commitment to this particular *P* and *Q*, substitute your favorite pair of incompatible claims, commitment to which is easy because their incompatibility isn't obvious. (Perhaps *P* = 'I freely choose my actions' and *Q* = 'Every event is determined by a prior event'.)

> distinguish *individual* from *specific* relativism. The restriction of the latter to the human species, stamps it as *anthropologism*.[50]
>
> Specific relativism makes the assertion: Anything is true for a given species of judging beings that, by their constitution and laws of thought, must count as true. This doctrine is absurd.[51]

Frege raises a similar concern, more colorfully:

> But what if beings were even found whose laws of thought flatly contradicted ours and therefore frequently led to contrary results even in practice? The psychological logician could only acknowledge the fact and say simply: those laws hold for them, these laws hold for us. I should say: we have here a hitherto unknown type of madness.[52]

Is the expressive conception a form of anthropologism? Clearly not, since it explains logic in terms of the practice of deontic scorekeeping, and it's quite possible that nonhuman creatures could engage in scorekeeping (and perhaps some in fact do). If such creatures count as scorekeepers, then logic applies to them as much as it applies to us.

Still, the general charge of relativism is warranted, given that on the expressive conception logic is relative to the practice. However, this doesn't have the dire consequences predicted by Frege and Husserl, as we can see by noting once again that on Brandom's view deontic scorekeeping is the *sine qua non* of sapience—any beings capable of conceptually contentful thought must be interpretable as keeping score, and any beings not so interpretable don't count as sapient. In a situation reminiscent of that described by Davidson (1973), it is quite impossible to encounter Frege's "logical aliens." We simply could not ascribe contentful thought and speech to those aliens and at the same time ascertain that their logic flatly contradicts our own. For logic makes explicit the pragmatic structure of linguistic behavior, and if the aliens' logic contradicted our own, their linguistic behavior would thereby be unintelligible to us. (How, for example, could

[50](Husserl 1900/2001, pp. 77–78)
[51]*Ibid.*, p.79
[52](Frege 1893/1964, p. 14)

we see such beings as thinking at all if they never treated at least some commitments as incompatible?)

Of course, this leaves open the possibility that we encounter beings who are interpretable as scorekeepers, and who endorse a logic which doesn't flatly *contradict* ours, but which differs in some important respects. This is exactly the situation which obtains between actual logicians today, some of whom endorse, e.g., *ex falso quodlibet*, and some of whom don't—each is partially alien to the other. But these partial aliens take themselves, at least some of the time, to have a genuine disagreement about the *correct* logical principles. So the fact that the expressive conception of logic permits the existence of partial logical aliens is, in the first place, a good thing, since there are actual such aliens, and we wouldn't want a theory to pronounce them impossible. And in the second place, Frege would not obviously shrink from this possibility. For one thing, he himself was an advocate of change in logic—revolutionary change—from what had been received as the obviously correct logical system. For another, and again as Frege's own case shows, it is entirely consistent to hold that partial logical aliens are possible and also that there is only one correct logic, which applies to all thinking beings.[53]

Thus the sort of relativism entailed by the expressive conception of logic really amounts to taking logic to be relative to thinking beings as such, and so logical laws apply to every thinking being. It does follow that logic won't apply to plants, planets, or planarians; but that doesn't seem so absurd.

[53]Brandom himself endorses a sort of pluralism, as we've seen, since according to him any logic which does the job of making explicit *some variety or other* of inferential goodness is a correct logic, and there are indefinitely many such logics which fit the bill. But even this doesn't warrant Frege's concern, since each variety of logical goodness expressed by a particular logical system—bare material goodness, relevant goodness, necessary truth preservation, and so on—will apply to all scorekeepers as such. That is, though Brandom's pluralism relaxes the uniqueness of correct logic, it does not narrow the *scope* of those subject to the norms of each possible logic.

4.4 naturalism in a foundational theory of logic

Brandom evades the traditional critique of psychologism in large part due to the character of the underlying psychological theory being appealed to: Frege and Husserl were occupied with prevailing empiricist conceptions of psychology, but Brandom is working with a kind of transcendental psychology. One might well wonder, then, whether this is ultimately a naturalistic account at all. (Indeed, Hanna (2006) appeals to a sort of transcendental psychological structure in his theory, and pretty much for that reason denies that his theory counts as a form of psychologism.) What's naturalistic about Brandom's story?

The principal reason is, I think, that nothing is appealed to in Brandom's account which isn't a transaction with the ordinary, natural world—either its humdrum nonsapient inhabitants, or the other scorekeepers in our community. We take attitudes toward ourselves and each other, and we update our attitudes as we converse and get along in the world. But nowhere in the explanation of the significance of logical vocabulary do we need to invoke any abstract objects, mathematical operators, rational intuitions, or graspings of a third realm. This is entirely in line with the kindred strategy of expressivism in metaethics, which aims to explain a discourse by appealing to speakers' *attitudes* rather than, in the first place, any objects and properties the discourse appears to be about.

To be sure, this is not a stringent naturalism. No facts about logic are being reduced to facts about physics, say, and neither are logical claims being translated into the vocabulary of some natural science. But a relaxed naturalism is still admirable, since it domesticates the possibly puzzling realm of logic, showing it to be intelligible in terms of familiar persons and practices. This epistemic goal is what I believe to be the laudable aim of the naturalist impetus, and I believe that, in this case, the goal is satisfied.

Chapter 5

The Normativity of Logic

In this chapter I explore the ways in which logic on the expressive conception can be said to be normative. Because the expressive conception in some ways stands athwart today's dominant conception of logic, but in line with a much older conception, I begin with a short historical section before moving to the expressive conception in particular.

5.1 logic's historical arc

This section recounts the history of logic across two distinctions. First is the act/content distinction: is logic primarily concerned with acts (of judging, thinking, inferring), or is it primarily concerned with the contents of such acts (propositions, statements, sentences)? Second is the norm/fact distinction: is logic something that provides, of itself, norms for reasoning, or does logic state facts about relations between contents, which may or may not particularly matter for reasoning?

5.1.1 logic as concerned with acts

For the better part of its history, logic has been conceived of as a discipline concerned with a special kind of *act*—inferring, reasoning, or judging. We might well begin with Aristotle, since his logical treatises inaugurate logic as a dedicated field of inquiry. In the *Prior Analytics*, where he develops the theory of the syllogism, Aristotle tells us that the subject being studied is *demonstration*, which is a species of deduction. Deduction is then defined as follows:

> A deduction [*sullogismos*] is a discourse [*logos*] in which, certain things being stated, something other than what is stated follows of necessity from their being so.[1]

Translators use a variety of terms to translate 'logos', but always with a term for some sort of *activity*: 'speech' or 'argument', for example.[2] The translation of 'sullogismos' by 'deduction' is commonplace, and reflects the common thread between Aristotle's conception of the syllogism and our own conception of deduction: the notion that some things follow from others with necessity.

What Aristotle tells us about deduction in the remainder of the *Prior Analytics* largely concerns the various forms that deduction may take (the various figures of the syllogism), and in the *Posterior Analytics*—in the standard arrangement of Aristotle's corpus, the very next treatise—he concentrates on the specific conditions a deduction must meet to qualify as demonstration. The nature of deduction itself gets no further discussion, but in his elucidation of the syllogistic forms, Aristotle speaks in many places of what may or may not be *inferred*, or what conclusions may be *drawn*, indicating that even when treating deduction at the abstract, schematic level, there is still an activity of some kind to which logic pertains.[3]

[1] Aristotle, *Prior Analytics* i.2, 24^b18–20, trans. Jenkinson. See also *Topics* i.1, 100^a25–26 (Barnes 1984).

[2] Smith (1989) uses 'speech', 'argument' is the choice of Striker (2009).

[3] See in particular Book ii of *APr*, 53^{ab}, 57^b, 63^a, 64^b (Barnes 1984). It is quite possible that the activities Aristotle is concerned with are *explanation* (in the case of the demonstrative syllogism) and

Just as with the Aristotelian structure of logic (discussed in Chapter 2), the notion that logic concerns acts persists into the early modern period. The Port-Royal *Logic* makes its focus on acts quite clear:

> Logic is the art of conducting reason well in knowing things, as much to instruct ourselves about them as to instruct others. . . . [T]his art does not consist in finding the means to perform these operations [conceiving, judging, reasoning, and ordering], since nature alone furnishes them in giving us reason, but in reflecting on what nature makes us do, which serves three purposes. The first is to assure us that we are using reason well, since thinking about the rule makes us pay new attention to it. The second is to reveal and explain more easily the errors or defects that can occur in mental operations . . . . The third purpose is to make us better acquainted with the nature of the mind by reflecting on its actions . . . .[4]

Arnauld and Nicole conceive of logic's subject matter as mental acts, and its purpose as the understanding and maintenance of such acts. Their debt to Aristotle is clear not only in their discussion of reasoning as syllogistic, but in their taxonomy of mental acts: conceiving, judging, reasoning, and ordering correspond to the received organization of Aristotle's logical works—the *Categories* deals with simple ideas of objects and properties, *On Interpretation* with propositions and judgment, the *Analytics* with deduction, and the *Topics* with the ordering of syllogisms in dialectical reasoning.

A little more than a hundred years later, Kant wrote, in a passage now famous, that

> . . . since the time of Aristotle [logic] has not had to go a single step backwards, unless we count the abolition of a few dispensable subtleties or the more distinct determination of its presentation, which improvements belong more to the elegance than to the security of that science. What is further remarkable about logic is that until now it has also been unable to take a single step forward, and therefore seems to all appearance to be finished and complete (B viii).[5]

---

disputation (in the case of the dialectical syllogism). In any case, this is all quite consistent with my reading of Aristotle above (in chapter 2); there is nowhere Aristotle identifies a specific mental activity such as inferring, which corresponds to the syllogism as judgment corresponds to the proposition, and thought-of-simples corresponds to nouns and verbs.

[4](Arnauld and Nicole 1683/1996)

[5](Kant 1787/1999, p. 106)

Although in its details the logic Kant endorsed was more than trivially different from Aristotle's (for example, Kant includes the use of the Stoics' compound propositions such as conditionals and disjunctions), the conception of logic as concerned with mental acts remains intact:

> The sphere of logic is quite precisely delimited; its sole concern is to give an exhaustive exposition and a strict proof of the formal rules of all thought, whether it be *a priori* or empirical, whatever be its origin or its object, and whatever hindrances, accidental or natural, it may encounter in our minds.[6]

It is only in the 19th century that philosophers began to conceive of logic as concerned not with mental acts, but rather with the *contents* of such acts, thought of as items independent of particular thinkers. But even so revolutionary a figure as Frege, whose logical innovations turned the discipline on its ear, remained wedded in his early work to the conception of logic as concerned with acts. His *Begriffsschrift* of 1879 included, as part of its new notation, signs indicating judgment—a small vertical stroke appended to the front of a formula—and inference—a long horizontal line written between formulas in a proof. And his definition of the conditional, which otherwise looks like our familiar truth-functional definition, consists of a statement of the three cases in which a conditional is "to be affirmed" and the one in which it is "to be denied" (Frege 1879/1967 pp. 10–20).

5.1.2 logic as concerned with contents

During the mid-19th century, a rival conception of logic was gestating. Bolzano's 1837 *Theory of Science* holds logic to properly concern propositions "in themselves" [*Sätze an sich*], or "objective propositions." These are the *contents* or meanings of particular mental or linguistic expressions, considered in abstraction from particular thinkers, and shareable across times, utterances, and languages:

[6] ibid.

> One will gather what I mean by proposition as soon as I remark that I do not call a proposition in itself or an objective proposition that which the grammarians call a proposition, namely, the linguistic expression, but rather simply the meaning of this expression, which must be exactly one of the two, true or false; and that accordingly I attribute [concrete] existence to the grasping of a proposition, to thought propositions as well as to the judgments made in the mind of a thinking being (existence, namely, in the mind of the one who thinks this proposition and who makes the judgment); but the mere proposition in itself (or the objective proposition) I count among the kinds of things that do not have any existence whatsoever, and never can attain existence.[7]

At the turn of the 20th century, Frege and Husserl came to share Bolzano's position. Frege's later essays even share Bolzano's locution of "grasping" abstract, mind-independent meanings, although Frege prefers the term 'thought' [*Gedanke*] to 'proposition' [*Satz*] for the content of a sentential expression.[8] The focus on contents persists through the 20th century—even Quine, a stalwart opponent of meanings construed as abstract objects, introduces his logic textbook by saying that "Logic, like any science, has as its business the pursuit of truth. What are true are certain *statements*; and the pursuit of truth is the endeavor to sort out the true statements from the others, which are false."[9]

5.1.3 norms and facts

Against the background of the act/content distinction, and the distinction between conceptions of logic that it funds, a second distinction is lurking. This is the distinction between facts and norms.

Logic has long been seen as specially concerned with norms for judging, inferring, reasoning, or mental activity generally.[10] This normative role comes out especially clearly

[7](Bolzano 2004, pp. 40–41)

[8]See e.g. (Frege 1892/1948), (Frege 1918/1956), and (Frege 1979).

[9](Quine 1950/1982, p.1); emphasis is mine.

[10]I will hereafter use the tidier phrase 'norms for thought' to indicate this sort of norm, allowing that the

in the quotation from Arnauld and Nicole's *Logic* above; logic is "the art of conducting reason well." But this conception of logic as normative can be traced (unsurprisingly) all the way back to Aristotle, for whom 'syllogism' is a success term: every syllogism is a form of correct reasoning; there are no "invalid syllogisms." And the conception can likewise be traced forward through Kant, for whom pure general logic is "a canon of the understanding and reason" (A53/B77), and is analogous to the moral law in that it sets out the merely necessary laws for the use of reason (A55/B79).[11] And Mill's introduction to his *System of Logic* is shot through with the idea that logic provides norms for thought. For example:

> Logic is the common judge and arbiter of all particular investigations. It does not undertake to find evidence, but to determine whether it has been found. Logic neither observes, nor invents, nor discovers; but judges. It is no part of the business of logic to inform the surgeon what appearances are found to accompany a violent death. This he must learn from his own experience and observation, or from that of others, his predecessors in his peculiar pursuit. But logic sits in judgment on the sufficiency of that observation and experience to justify his rules, and on the sufficiency of his rules to justify his conduct. It does not give him proofs, but teaches him what makes them proofs, and how he is to judge of them. It does not teach that any particular fact proves any other, but points out to what conditions all facts must conform, in order that they may prove other facts.[12]

For Mill, logic is clearly a canon for the inferences made in any particular inquiry. And although he does, as we saw in the previous chapter, hold part of logic's ambit to be the theoretical analysis and description of the mental process of reasoning, he also endorses

---

activity which is the proper target of logical norms might vary depending on which philosopher's view is under consideration.

[11](Kant 1787/1999). For Kant it is not only pure general logic which is normative, but transcendental logic and applied logic as well. Applied logic is the theoretical analogue of the "doctrine of virtue proper," which gives advice on how to behave given our natural limitations of attention, unreflective desires, etc. Transcendental logic is concerned with the objective validity of pure concepts, and "has to do merely with the laws of the understanding and reason, but solely insofar as they are related to objects *a priori* and not, as in the case of general logic, to empirical as well as pure cognitions of reason without distinction" (A57/B81).

[12](Mill 1843, pp. 21–22)

Whately's claim that the other part of logic is an Art: "the rule, grounded in that analysis, for conducting the process correctly" (*ibid.*, p. 18).

Although the act/content and fact/norm distinctions are in principle orthogonal, yielding four possible conceptions of logic, two pairings are particularly natural: logic as concerned with *norms for acts* of judgment and inference, and logic as concerned with *facts about contents* of such acts.

It is not hard to see why the conception of logic as concerned with acts would be bound up with the idea that logic is normative. In the first place, for any particular human activity we care to investigate, we tend to be interested in doing it well, or knowing how it ought to be done. For the activity of reasoning, philosophers in particular should be especially interested in the norms of correctness, since reasoning and its assessment are (arguably) the philosopher's stock-in-trade. If, on the other hand, we think of logic as concerned with acts but essentially descriptive, we wind up with a kind of psychologism which construes logic as a branch of empirical psychology, assimilated to the study of reasoning, judgment and decision-making, and perhaps developmental psychology. So, given that we think of logic as concerned in the first place with acts of thought, it is more natural to think of it as a normative discipline than a descriptive one.

On the other hand, the conception of logic as concerned with the *content* of thought makes logic primarily concerned with facts, not norms. On this conception, logic is in the business of stating relations between propositions (especially implication, but also equivalence, independence, etc.) and properties of propositions (logical truth, contingency, logical falsehood, etc.). Statements of such relations and properties—" '$P \wedge \neg P$' is a contradiction," etc.—are, on the face of it, just like any other descriptive claims which apply predicates to objects, although in this case the objects are abstract contents rather than concrete things. Logic now looks like a science, or a branch of mathematics, rather than a normative enterprise like ethics or medicine.

Conceiving of logic as a fact-stating enterprise is not by itself troublesome; there are many fact-stating enterprises and we like them just fine. The trouble arises when we consider that logic was, for the greater part of its history, essentially in the business of providing norms for thought, and that the conviction that it ought to provide such norms has never really left us, even as the content conception eclipsed the act conception. Introductory logic textbooks still tell students that what distinguishes logic from psychology is that logic concerns *correct* reasoning, as opposed to psychology's interest in how people *actually do* reason under various conditions.[13] And the textbooks don't say this out of mere deference to tradition; as John MacFarlane puts it,

[13]To take a few examples:

> Logic is the study of correct reasoning. It is not a study of how this reasoning originates, or what its effects are in persuading people; it is rather a study of what it is that makes some reasoning "correct" as opposed to "incorrect". (Parsons 2010, p.5)

> Logic is concerned with arguments, good and bad. With the docile and the reasonable, arguments are sometimes useful in settling disputes. With the reasonable, this utility attaches only to good arguments. It is the logician's business to serve the reasonable. Therefore, in the realm of arguments, it is the logician who distinguishes good from bad. (Kalish & Montague 1964, p. 1)

> **Logic** may be defined as the organized body of knowledge, or science, that evaluates arguments. . . . The aim of logic is to develop a system of methods and principles that we may use as criteria for evaluating the arguments of others and as guides in constructing arguments of our own. (Hurley 2014, p.1)

> Logic is about reasoning – about going from premises to a conclusion. . . . **Logic** can be defined as *the analysis and appraisal of arguments*. When you do logic, you try to clarify reasoning and separate good from bad reasoning. (Gensler 2010, p. 1)

Of course, some psychologists bring normative concerns to their inquiry. Kahneman and Tversky believe subjects to be mistaken when they deviate from what statistical principles entail; Wason believes subjects to be mistaken when they fail to respond in accordance with *modus tollens*. But the norms against which such normative judgments are made do not come from psychology; instead they come from probability theory and propositional logic. A parallel point can be made about judgments of veridicality in psychophysical and perceptual experiments.

> Logic is often said to provide norms for thought or reasoning. Indeed, this idea is central to the way in which logic has been demarcated as a discipline, and without it, *it is hard to see how we would distinguish logic from the disciplines that crowd it on all sides*: psychology, metaphysics, mathematics, and semantics. (Try saying how logic differs from geometry without mentioning thought or reasoning, and try saying how logic differs from psychology without mentioning norms.)[14]

We need to be able to explain the way in which logic is normative for thought, but the content conception of logic makes this difficult in a way that it wasn't on the act conception. For it's not obvious how facts about abstracta (and their properties and relations) should matter for individual human beings' reasoning episodes, much less generate norms by which such episodes are to be assessed. Logic says "$P \wedge Q \vdash P$"—what does this mean for how we should reason? Let's call this question—"given the content conception of logic, how do logical facts relate to norms for thought?"—the *regulative question.*

At this point, two options suggest themselves, each backed by a recent philosopher's contention:

The Harman option. In *Change in View*, Gilbert Harman (1984) argued that logic is a fact-stating enterprise like any other science, and that accordingly logic has no special normative import for thought. Suppose, for example, that logic tells us that $A, B \models C$. Does this mean that anyone who believes A and B ought to believe C? No; sometimes it's better to give up belief in either A or B rather than accept C, for example when you know C to be false. But *logic* will not tell you which of these things to do; it tells you only that A and B jointly entail C. Likewise, logic tells us that '$P \wedge \neg P$' is inconsistent. Does this mean you ought not to believe contradictory propositions? No; sometimes you don't know which one to reject, or simply don't have time to worry about it. Life is short,

[14](MacFarlane 2004, p. 1). My emphasis.

and not every inconsistency in one's belief set can be ferreted out. Knowing only the logical fact that two propositions are contradictory doesn't by itself tell you what to do.

The MacFarlane option. MacFarlane (2004) argues that we should formulate and assess "bridge principles" that connect logical facts to norms for thought. Such principles would be analogous to familiar principles from normative ethics, such as the utilitarian claim that the act which maximizes aggregate happiness is morally right. The idea is to connect putatively nonnormative facts (facts about happiness, or about logical consequence) to normative directives (how we ought to conduct ourselves, or ought to reason).[15] MacFarlane proposes a method for systematically generating candidate bridge principles, and expresses confidence that we can assess them in a less biased way than we might assess competing accounts of validity itself, such as the accounts offered by classical, intuitionist, and relevant logics.[16] For example, the following bridge principles strike MacFarlane as plausible:

> (`Wo-`) If $A, B \models C$, then you ought to see to it that if you believe A and believe B, you do not disbelieve C.
>
> (`Wr+`) If $A, B \models C$, then you have reason to see to it that if you believe A and believe B, you believe C.

But

> (`Cp+`) If $A, B \models C$, then if you believe A and you believe B, then you may believe C.

is implausible, for Harman-style reasons: there are situations in which you ought not believe C, even if you already believe both A and B and $A, B \models C$; for example, when you believe C to be false or absurd.

[15] These bridge principles should not be confused with what Hempel (1966) calls bridge principles, the point of which is to bridge observational and theoretical claims. However, MacFarlane's bridge principles are in a way analogous to Hempel's in that they specify how we ought to take statements in one domain (observational claims or logical consequences) to constrain statements in another domain (theoretical claims or norms for thought), given that the former underdetermine the latter.

[16] In a footnote, MacFarlane says "I think there is a lot that is wrong or confused in this paper, but also some things that are right." So it's not clear whether he should officially be said to endorse its content.

On the Harman option, then, we resolve the regulative question by simply denying that logical facts have any special import for norms for thought. On the MacFarlane option, we produce and then assess bridge principles by asking whether they entail plausible norms for reasoning. But the Harman option is tantamount to eliminating logic as a form of inquiry specially concerned with reasoning well, and the MacFarlane option is, at least on its face, simply a test of particular answers to the logical fact/norm question against our intuitions about how we should reason, not a general answer to the regulative question. It could, of course, be construed as the beginning of inquiry into a more systematic and general answer—and I would welcome such inquiry.[17] But in its absence, we should try to do better.

A Frege option? It might seem that Frege already answered this difficulty for us, in claiming that logic is normative because it is true. Recall his analogy with physics and geometry, and his metaphor of the boundary stones:

> Any law asserting what is, can be conceived as prescribing that one ought to think in conformity with it, and is thus in that sense a law of thought. This holds for laws of geometry and physics no less than for laws of logic. The latter have a special title to the name "laws of thought" only if we mean to assert that they are the most general laws, which prescribe universally the way in which one ought to think if one is to think at all.[18]

> . . . [T]he laws of truth are not psychological laws: they are boundary stones set in an eternal foundation, which our thought can overflow, but never displace. It is because of this that they have authority for our thought if it would attain to truth.[19]

According to Frege, logic states what's true (or, in his Platonic/Aristotelian turn of phrase, "what is"), and hence we ought to "think in conformity with it." Moreover, since logic is

[17] At present, MacFarlane appears to have tabled his work on this topic, but has the paper in the "work in progress" section of his website.

[18] (Frege 1893/1964, p. 12)

[19] *Ibid.*, p. 13. My ellipsis omits the antecedent of a conditional—"If being true is thus independent of being acknowledged by somebody or other, then . . . "—but of course Frege endorses that antecedent.

about the most general sorts of facts, which are relevant to any thinking whatsoever, we ought to think in conformity with it whenever we think, or at least insofar as we want to think truly.[20]

Although this seems like a straightforward enough account of logic's normativity, in fact it runs up against the same difficulty Harman and MacFarlane are responding to, as can be seen when we ask what it means, exactly, that we "ought to think in conformity with" logic. Both Harman and MacFarlane agree with Frege that logic tells us truths; the difficulty is to say what import these truths have for our reasoning practices. Harman, taking into account the various ways in which we are bounded rational agents, with a variety of cognitive limits and non-cognitive goals, concludes that logic is on a par with any other fact-stating discipline: yes, it tells us true things (in particular, which propositions imply which others, which are contradictory, and so on), but there is *no general account* of the way these facts matter for reasoning, any more than there is for the facts delivered by other sciences. Frege seems partly in agreement, since he holds that logic is no more or less normative than physics or geometry.

However, Frege also seems to hold that there is some general way in which logic constrains our reasoning, since we ought to think in conformity with it. He doesn't say anything specific about what this amounts to, but presumably it's something like the rules that we ought not hold contradictory beliefs, or that we ought to believe the implications of our beliefs. But if this is right, then we're now in MacFarlane's territory, looking for the right bridge principles to tell us, given some logical truths, how we ought to reason.

So Frege's position doesn't help account for logic's normativity in any detailed way, and in fact it's positioned somewhere uneasily between Harman's and MacFarlane's, insisting that logic has no distinctive normative import but also that it has normative

[20]That is, Frege thinks the norms of logic are *regulative* for all thought irrespective of its subject matter, not *constitutive* of thought; thinking which doesn't conform to the laws of logic doesn't thereby fail to count as thought in the first place. On this point, see (MacFarlane 2002).

consequences for our thinking and reasoning. If we want to understand the normative import of logic, we must side with Harman or MacFarlane, or find a genuine third way. I believe there is such a third way, but it requires abandoning the conception of logic as concerned primarily with contents, upon which the regulative question is premised.

5.2 Ryle and Sellars on norms implicit in logic

In the early 1950s, both Sellars and Ryle suggested a way of thinking about logic on which the claims of logic are already implicitly normative in some way. Their insights are incorporated into and developed by Brandom's expressive conception, as we'll see in the section following this one.

5.2.1 Ryle

" 'If', 'So', and 'Because' " (Ryle 1950) argues for a number of novel claims. The general strategy of the paper is to examine hypothetical statements (what we'd now more readily call conditionals—those made with 'if') by way of their relationships to inferences, explanations (statements using 'because'), and arguments (made by using 'so'). Ryle thinks it's clear that an argument "p, so q" is not a statement, but that in order to be valid, the associated hypothetical "if p then q" must be true; if the hypothetical is false, then q does not follow from p. But what does it mean to say that the argument requires the truth of the hypothetical?

Ryle's answer is that hypothetical statements are *inference licenses*, and that arguments are applications of such licenses, in much the same way that taking a trip by train is an application of a train ticket. Just as the ticket must be valid or legitimate for it to be used correctly, a hypothetical statement must be true if its associated argument is to be good. And like the ticket, a hypothetical must have "indents" or placeholders which

specify what it may be applied to. Train tickets specify passengers, stations, and times; hypotheticals specify statements in the antecedent and consequent positions. Thus the inference-license view explains why hypotheticals, but not arguments, may have open sentences as constituents:[21] hypotheticals are to be applied to arguments, but arguments are not themselves applied to anything; they are made.

Ryle thinks this view has two main consequences for the philosophy of logic. First, since hypothetical statements are licenses for making arguments and inferences, using hypotheticals is "more sophisticated than the activities of wielding and following arguments" (*ibid.*, p. 243). Second, because hypotheticals are "sophistications" on inferences, they do not report any facts or describe any states of affairs.

Indeed, Ryle thinks that hypotheticals are not even compound statements, properly speaking. Just as the open places in a train ticket are specifications for persons and stations—but not themselves persons or stations—the open places in a hypothetical statement are not themselves statements, even in a hypothetical such as "If today is Monday, then tomorrow is Tuesday," which appears to have statements for its constituents. As evidence for this, consider subjunctive formulations of conditionals, which have as constituents things like 'today were Monday'. Likewise, negating a hypothetical requires, according to Ryle, bringing out its implicit modal content, with a similar result: "Today could be Monday and yet tomorrow not Tuesday;" here, too, the constituents are not statements.

So what is going on when someone asserts a hypothetical statement? After all, many hypotheticals are not worded subjunctively, and it's an entrenched convention of logicians that hypotheticals are symbolized as 'if p then q', where the italic letters are propositional variables. Ryle's answer is that hypotheticals are like the demonstration of a skill by miming the movements which would, under normal circumstances, constitute

[21] Compare "If anyone is late, he'll miss the continental breakfast" to "Anyone is late, so he'll miss the continental breakfast."

performance of the skill. But in the case of argumentation, the skill being demonstrated is a verbal one, so that its demonstration requires verbalization. In this respect the assertion of a hypothetical statement is like an acting coach demonstrating how to perform; the coach needs to utter lines to do so, but such utterances are not assertions.

" 'If', 'So', and 'Because' " thus articulates three main claims which are taken up in Brandom's development of the expressive conception of logic:

- Hypothetical statements are not for reporting matters of fact, but act as inference licenses, such that when one "has" the antecedent one may make the inference to the consequent. Hence a hypothetical statement is a kind of permission.
- Hypothetical statements belong to "post-inferential levels" of discourse, and can be employed by a speaker only once the ability to make inferences and arguments is already in place.
- Hypotheticals have implicit modal content, as shown by the fact that we negate "if p then q" with "it's possible that p but not q."

These claims are not adopted as-is, however; part of Brandom's advance over Ryle is his adoption of 'expression' rather than 'application' as the key role for conditionals, as well as a more detailed account of the role of inference and its expression. But before moving to Brandom's view of hypotheticals and norms we must turn to Sellars.

5.2.2 Sellars

Recall from Chapter 3 that in "Inference and Meaning" Sellars argues, among other things, that the function performed by subjunctive conditionals is indispensable to language, that there is material correctness of inference which isn't reducible to formal correctness, and consequently that what Carnap calls 'P-consequence' cannot be traded in for 'L-consequence'.

In "Inference and Meaning" there is, however, another striking claim of Sellars's which bears on the issue of logic's normativity; this is his claim that "the language of modalities is interpreted as a 'transposed' language of norms" (Sellars 1953, p. 332). This is a claim about the occurrence of modal terms in formal languages like Carnap's, and to understand it we need to understand how Sellars sees the metalanguage in Carnap's system as shirking the job required of a proper metalanguage for logical languages.

Carnap's formal language, like formal logics generally, purports to say which object-language inferences are permissible by defining a set of inference rules in a metalanguage (Carnap's term is 'syntax language').[22] But according to Sellars, what passes for rules of inference—statements about syntax in the metalanguage—are not genuinely rulish:

> . . . [A] syntactical metalanguage cannot permit the formulation of syntactical rules, unless (1) it contains a term for the activity of asserting, and (2) it contains an expression having the force of "ought". To the extent that a so-called "syntactical metalanguage" falls short of these requirements, it is an abstraction from a syntactical metalanguage proper. It is undoubtedly convenient to study calculi by means of such truncated metalanguages as mention only the structural inter-relationships of the sign-designs of these calculi, but it is essential for our purposes to stress that these truncated metalanguages become capable of formulating rules only when supplemented by the equipment mentioned above.[23]

Carnap's metalanguage specifies the *circumstances* in which one expression may be derived from another, but it cannot *say* that the expressions may be so derived.

Now Carnap makes the further claim that statements about logical necessity are related to statements about derivability in the following way: 'If $(\phi a \wedge \psi a)$ then it is logically necessary that ϕa' conveys the same information as ' 'ϕa' is an L-consequence of '$(\phi a \wedge \psi a)$' '.[24] But if, as Sellars has argued, statements about L-consequence are

[22](Carnap 1937, p. 4)
[23](Sellars 1953, p. 331)
[24](Carnap 1937 p. 303; recounted by Sellars on 1953, p. 332)

doing duty for what *may* be derived, then claims of logical modality convey the same information as what are, properly speaking, claims of permissibility.

How, then, are these logical claims related to normative claims? Sellars quotes Carnap on "the material mode of speech":

> The material mode of speech is a transposed mode of speech. In using it, in order to say something about a word (or a sentence) we say instead something parallel about the object designated by the word (or the fact described by the sentence . . .) . . .[25]

When Sellars claims that "the language of modalities is interpreted as a 'transposed' language of norms," the quotation marks around 'transposed' suggest that Sellars is playing on Carnap's use of that term. If that's right, then the reason Sellars claims that the language of modalities is a transposed language of norms is that, in formal languages like Carnap's, when we want to talk about norms using the material mode, we instead say something parallel about modality. For example, without transposition, we can say "You may infer 'Falafel is an animal' from 'Falafel is a cat'." Transposing to the material mode, we instead say that " 'Falafel is an animal' *follows necessarily* from 'Falafel is a cat'."[26]

What explains the fact that this transposition is possible? Why is it modal language in particular (and not, say, gustatory language) which functions as material-mode expression of norms? For Sellars, the answer is in the relation of *conveyance*. The idea is that linguistic performances can not only assert, but also convey, information or content. When my roommate asks whether we have beer and I respond "We're out of beer," I've asserted the proposition *that Per and Adam are out of beer*, but also conveyed

[25]Quoted on (Sellars 1953, p. 331); quotation is from (Carnap 1937 p. 309).

[26]This reading of Sellars thus disagrees with that provided by Brandom in (Brandom 2008) on two counts. First, Brandom thinks that Sellars never "manages to say clearly just what sort of 'transposition' he has in mind" (p. 100), while I think it's pretty clear, given Sellars's quotation of Carnap, that Sellars is using Carnap's sense of 'transposition'. Second, Brandom's reconstruction of the transposition claim itself is that normative vocabulary is sufficient to specify the practices (inferences) which are then sufficient to deploy modal vocabulary (p. 101). On my reading, it's the modal vocabulary which says something about the norms.

the information *that Adam believes that Per and Adam are out of beer*. In general, a speaker's assertion that *p* asserts (but doesn't convey) the proposition that *p*, but it also conveys (and doesn't assert) the information that the speaker believes that *p*. Conveyance in this technical sense can be conceived of as a relation between an utterance and a state of mind.

Sellars claims, then, that although logical-modal and normative vocabulary don't assert the same thing, they convey the same thing. And the thing conveyed is "the existence of a rule-governed mode of behavior" (Sellars 1953, p. 333). Since what is conveyed is a state of mind, we must think of this mode of behavior itself as a state of mind, such as a general pro-attitude toward rule-conforming behavior, or a disposition or policy to so behave. For example, my assertions of "*Q* follows necessarily from *P*" and "*Q* may be inferred from *P*" both convey my pro-attitude toward concluding *Q* from *P*, or my disposition to do so.

Here, then, are three claims Sellars makes in "Inference and Meaning" regarding the normative status of logic, each of which is taken up in Brandom's later work:

- Subjunctive conditionals express inference licenses—in the language of Sellars's essay, they are sentences of the object language which express what in the metalanguage is said by P-rules (if the inference licensed is material) and L-rules (if it is formal).

- The modal language of logic is a transposed language of norms, meaning that insofar as logic is not merely a set of syntactic relations among sentences, its traffic in modal notions is ancillary to its purpose as a canon for thought (or assertion, as Sellars prefers).

- Corresponding logical-modal claims and normative claims, although they assert different things, convey the same thing, namely a speaker's state of mind regarding

inferences.

With these claims (and some argument for them) on the table, it is time to look at their inheritance in the expressive conception of logic.

5.3 acts and norms: the expressive conception

Consider again MacFarlane's search for bridge principles: it's exactly because logical consequence, conceived of as a factual relation between propositions, appears to be normatively inert that the need for a bridge principle arises—logical consequences sure *seem* like they ought to have normative import, and yet all we see are facts about sentences. (Likewise for other logical properties and relations.) So a bridge principle is required, which will tell us the rational norms associated with logical facts.

But it looks like Sellars and MacFarlane are responding to the same situation in two different ways: MacFarlane looks for bridge principles; Sellars claims that the language of logical consequence has a hidden normative dimension.

These two responses aren't necessarily in conflict: suppose an expressive conception of logic has it (as Sellars does) that logical vocabulary is implicitly normative. Then we can see the search for a bridge principle as the search for a specification—an explicitation—of that normative aspect. But what the expressive account would also offer us is an explanation for *why* that bridge principle is true (or, alternately, why logical vocabulary has that aspect).

In the rest of this section I articulate the normative character of logic on Brandom's expressive conception.

5.3.1 conveyance as a constraint on bridge principles

Take one of MacFarlane's bridge principles,

> (`Wr+`) If $A, B \models C$, then you have reason to see to it that if you believe A and believe B, you believe C.

which is a conditional in which the antecedent is a statement of entailment and the consequent is a normative directive for reasoning. That structure is the general form a bridge principle will take. Suppose that this particular principle is correct. If we follow Sellars, then the antecedent and consequent convey the same state of mind. That is, the claims

> $A, B \models C$

and

> you have reason to see to it that if you believe A and believe B, you believe C,

although they have different *asserted* contents, convey the same thing. Again, since Sellars says no more about what's conveyed other than a "rule-governed mode of behavior," we must speculate about what, exactly, this is; but it's not implausible to think of it as a disposition to infer C from A and B together.

Irrespective of exactly what sort of state of mind is conveyed, we now have a formal Sellarsian constraint on bridge principles: a correct bridge principle will be such that its antecedent and consequent convey the same state of mind. Already this is a more informative thesis than anything MacFarlane tells us about bridge principles, and it is even a weak sort of explanation: it tells us something about *why* any particular correct bridge principle obtains, and not merely *that* it does. Furthermore, there is also something expressivist in spirit about this weak explanation, since we are appealing to a state of mind in order to explain a kind of semantic fact. This is so even without having imputed to Sellars a fully worked-out expressivist metasemantics for either logical or normative vocabulary.

However, these advantages of following Sellars on conveyance carry a *prima facie* cost: how do we explain the different asserted content of logical and normative claims? That is, if the two sorts of claims convey the same state of mind, why don't they simply mean the same thing when asserted? The difficulty is made more pointed when we look back at what Carnap—who inspired the conveyance idea in the first place—had to say about the material mode as a "transposed" mode of speech:

> The origin of a transposed mode of speech can sometimes be explained psychologically by the fact that the conception of the substituted object *b* is for some reason more vivid and striking, stronger in feeling-tone, than the conception of the original object *a*. This is the case with the material mode of speech. The image of a word (for instance, of the word 'house') is often much less vivid and lively than that of the object which the word designates (in the example, that of the house). Further, the fact, which is perhaps a consequence of the psychological fact just mentioned, that the approach and method of syntax have hitherto not been sufficiently known, and that, in consequence, the majority of the necessary syntactical terms have not been a part of ordinary language, may have contributed to the origin of the material mode of speech. ... People are not accustomed to direct their attention to the sentence instead of the fact; and it is apparently much more difficult to do so.[27]

Carnap appears to think that if it only weren't so difficult, or if we'd developed his "method of syntax" sooner, or if houses weren't so much more vivid and lively than 'house', we could really just dispense with the material mode of speech entirely, and do all the work we need in the formal mode of speech. So what, if we follow Sellars, is the significance of logical vocabulary distinct from the corresponding normative vocabulary, if the former is just the material-mode correlate of the latter?

5.3.2 logic and norms on Brandom's expressive conception

We can see Brandom's deontic scorekeeping model as a way of upholding Sellars's contention that logical claims and normative claims convey the same thing, while at

[27](Carnap 1937, p. 309)

the same time offering an explanation of the distinctive content associated with assertions of each. Although Brandom does not himself distinguish these, it turns out that there are several senses in which logic can be involved in linguistic performances, and several senses in which such performances can "be normative." So it will be best to approach the explanation in stages, starting with the least sophisticated sort of vocabulary and then building in more sophistication at each stage.

First stage: logically atomic sentences. Recall that the pragmatic model of deontic scorekeeping, in terms of which all linguistic activity is to be explained, construes conceptual content in terms of its potential to change the deontic score, which is a constellation of normative statuses associated with each interlocutor. So on Brandom's view, *all* linguistic activity is normative at least in the sense that it makes a difference to how scorekeepers are appropriately treated. Every assertion, for example, is something like the signing of a contract, or participation in a bar mitzvah—it changes what others may expect of you, and what you may expect of them. This true even when what's asserted contains no logical or normative vocabulary, such as the philosophical chestnut 'The cat is on the mat'.

Sellars would say that my assertion of *P* conveys my belief that *P*; someone hearing me assert it would be licensed to believe *that Adam believes that P*, even though I haven't asserted it. In Brandom's idiom, my assertion of *P expresses* my *commitment* to *P*, since the assertion makes explicit what was otherwise only implicit in what I do. (Someone keeping score on me would be licensed to attribute *commitment to P* to me, and thus to undertake commitment to *Adam is committed to P*.) What is expressed is thus something normative, since *qua* commitment it carries consequences for what other commitments I ought to have, which others I ought not have, what I may be challenged on, etc.

These two ways of thinking about the normative significance of language—in terms of scorekeeping potential, and in terms of what underlying attitude is expressed—will be common to the next three stages. Of course, they are really just two sides of the same coin; it is because of its scorekeeping potential that a linguistic performance expresses what it does. But it's helpful to distinguish them, since they correspond to what Sellars distinguishes as asserted content and conveyed content, and so we can see Brandom as taking over this Sellarsian strategy.

Second stage: logically compound sentences. If even logically atomic sentences are normative in some sense, then how, if at all, does the addition of logical vocabulary matter, normatively speaking?

Take the conditional as a first case. One way to understand the significance of conditional statements on Brandom's view is in terms of their contribution to the deontic score. Like all assertions, conditionals have normative significance in the sense that they make *some* difference to the deontic score, but the distinctive scorekeeping contribution of 'if *P* then *Q*' is to establish a scorekeeping context in which, given commitment to *P*, it's appropriate to undertake commitment to *Q*.

But as with the case of logically atomic assertions, the scorekeeping significance of conditionals enables us to see them as expressing an underlying attitude. What is expressed (or as Sellars would have it, conveyed) by asserting a conditional is that the speaker endorses the inference from antecedent to consequent. This is on the one hand nothing fancier than how asserting *P* expresses the speaker's commitment to *P* (and hence might seem to have no more normative significance than the assertion that *P*), but on the other hand what's expressed isn't merely a matter of what the speaker takes-true (and hence how they think one ought to believe), but a matter of how the speaker thinks one ought to *reason*—namely, one ought to infer *Q* from *P*.

If that's right, then the distinctive normative character of logical vocabulary, something not shared with logically atomic empirical claims, is *not* some new grade of normativity, but rather that the object of the expression is distinctive—it's an attitude regarding the appropriateness of reasoning rather than mere belief.

This is as it should be, since it matches nicely with Brandom's way of demarcating logical vocabulary in the first place. It is natural, when first reading Brandom's work, to get the mistaken impression that what separates logical from nonlogical vocabulary is that the former, but not the latter, is expressive. But in fact all vocabulary is expressive, and the demarcation of logical vocabulary concerns *what* is expressed: the underlying structure of scorekeeping proprieties in virtue of which linguistic performances have conceptual content in the first place. Likewise, when we turn to consider normativity, it's not that logically compound claims are somehow normative but atomic claims aren't; instead, logically compound claims express proprieties of reasoning, whereas atomic claims express proprieties of belief.

Third stage: valid arguments. When we consider the making of arguments, rather than mere assertions, a new normative question arises. By 'valid argument' I mean any argument made in an object-language that instantiates a logically valid form, for example:

If today is Monday, then tomorrow is Tuesday
Today is Monday

Tomorrow is Tuesday

We can ask three questions about such arguments. First, what is the scorekeeping significance of making them; second, what is expressed by making them; and third, given that validity is a good-making feature of arguments, how should we understand this goodness? The first two questions are the same that we've discussed in the first two

stages; the third is new. I'll take them in turn.

What is the scorekeeping significance of making a valid argument? Nothing other than the combined scorekeeping significances of the individual claims making it up. An assertion of '*P*, if *P* then *Q*, so *Q*' simply combines whatever significances *P* and *Q* have, along with the significance of the conditional. This is extremely close to the scorekeeping significance of the associated (and invalid) material inference '*P*, so *Q*'. A second party who takes me to have made the material inference attributes to me not only commitment to *P* and to *Q*, but endorsement of the inference itself. If I in addition assert the conditional 'If *P* then *Q*', and thereby make a valid argument, the only new scorekeeping significance added by my assertion is that second parties may make claims about what I've explicitly avowed, and not merely what I've implicitly endorsed by making a material inference.

So when we turn to ask what is *expressed* by making a valid argument, the answer is again nothing more than what is expressed by the individual claims—to take the same example, commitments to *P* and to *Q*, and endorsement of the inference from *P* to *Q*. The fact that the argument is valid adds nothing to what is expressed by making it. (Predicating validity of an argument, however, is different. That's the next stage of sophistication.)

What, then, of the third question about valid arguments—how should we understand their goodness? Recall that on Brandom's account, an argument is valid just in case all substitutions of nonlogical for nonlogical vocabulary yield materially good substitution-instances. If an argument is valid, then, all its substitution-instances are materially good, including the argument as it stands (since it is one such substitution-instance). This means that insofar as we're concerned with the goodness of valid arguments in the sense of their *rational authority*, they are good in the way that any materially good inference is. However, this isn't to say there's nothing good about valid arguments over

and above their material goodness. To quote again a passage I've quoted once already:

> On this view [the expressive conception], the philosophical significance of logic is not that it enables those who master the use of logical locutions to *prove* a special class of claims—that is, to entitle themselves to a class of commitments in a formally privileged fashion. The significance of logical vocabulary lies rather in what it lets those who master it *say*—the special class of claims it enables them to express.[28]

In chapter 3 I quoted this passage to illustrate Brandom's conception of the discipline or enterprise of logic, which focuses on the content-displaying role of logical vocabulary. Here I'm quoting it to emphasize that the expressive benefit of logical vocabulary occasions a corresponding virtue in valid arguments. The fact that, in a valid argument, the inference being endorsed is made explicit in the form of a claim among the premises—as Frege would put it, that "everything necessary for a correct inference is expressed in full"[29]—is a *social* virtue of the argument.

Take, for example, the fact that philosophical articles are chock full of valid arguments, but the validity of the arguments doesn't generally—maybe ever—change a reader's views. Philosophers don't just see that an argument is valid and then accept its conclusion. Far more common is to acknowledge the argument's validity but dispute the truth of one or more premises, and it is putting arguments into deductively valid form which helps readers see exactly what's at issue, and what may be disputed. (This is just a special case of a point made in §3.3.1.) So there is a normative quality peculiar to valid arguments, and it's a social virtue rather than one of rational authority.

Fourth stage: predications of logical properties and relations. By predications of logical properties and relations, I have in mind metalinguistic attributions of properties like validity, contradiction, and equivalence to groups of object-language sentences, or forms of such. For example, " 'Today is Monday, if today is Monday then tomorrow

[28](Brandom 1994, p. xix)
[29](Frege 1879/1967, p. 12)

is Tuesday, so tomorrow is Tuesday' is valid," "*Barbara* is valid," or " 'P' and '$\neg P$' are contradictory." It is constructions like these, finally, which bring us to the question about the normativity of logic raised by Harman and MacFarlane, for each of them is concerned about statements like '$A, B \models C$', which in this case is a metalanguage statement predicating the entailment relation of $\{A, B\}$ and C, or the property of validity to the argument 'A, B, so C'.[30]

Since these are assertions, we can ask the same two questions about them that we asked in the first two stages: what is the scorekeeping significance, and what is expressed, by a predication of a logical property or relation? Validity is the most interesting case, or at least the paradigm case of concern to philosophers, so I'll take it up first.

What is the scorekeeping significance of a predication of validity? Given Brandom's account of what validity is, to assert that an argument is valid is to undertake a *quantificational* commitment to the effect that all of the argument's substitution-instances are materially good, which is in turn commitment to every claim 'α is materially good' for substitution-instances α.[31] So to understand the scorekeeping significance of predicating validity, we need to understand the scorekeeping significance of predicating material goodness. What is it to say that an inference is materially good? Just that it's rationally appropriate to infer its conclusion from its premises, that it's a legitimate move in the game of giving and asking for reasons. 'Appropriate' can be made precise in many ways—including, of course, that one is committed or entitled to make the inference, but getting as fine-grained as we can manage using specialized logical terms (more on which in a moment)—but predicating material goodness of an inference without qualification does not distinguish between these precisifications.

[30] Although officially variables like 'A', 'B', and 'C' should be distinguished from object-language sentences like 'Today is Monday', this distinction doesn't matter for the discussion here, which really concerns the metalanguage predicates and their significance. (For the same reason I am going to dispense with Quine's corner-quotes when mentioning variable schemata.)

[31] "Attributing commitment to a claim of the form $(x)Px$ is attributing commitment to all claims of the form Pa" (Brandom 1994, p. 434).

On this account, anyone calling an argument valid thereby undertakes commitments to each member of a class of arguments being materially good, even if they lack the concept of material goodness. This might seem like a bad or even fatal consequence of the account, but it isn't. That's because it's an account of what they have committed themselves to by making an assertion of the form 'α is valid', and a general feature of Brandom's scorekeeping semantics is that in making an assertion one is always committing oneself to more than one would acknowledge, and second parties are always in position to draw out and articulate those unacknowledged commitments. In this case, Brandom is the second party, offering a perspective from which validity—the thing philosophers and logicians have been theorizing since at least Aristotle—is a social-inferential phenomenon, the correct analysis of which is a quantificational commitment to the material goodness of an argument's substitution-instances. This is not incompatible with the fact that few (if any) other philosophers would themselves analyze validity in these terms, or endorse explicit formulations of such. We can say that although such philosophers do not believe, *de dicto*, that a valid argument is one with all and only materially good substitution-instances, they nonetheless believe, *de re*, of material goodness that all and only a valid argument's substitution-instances are such.

What should we make, then, of other theoretical accounts of validity? I believe the answer is that, from Brandom's perspective, those other accounts are attempts to single out various sorts of material goodness one might be concerned with. Take, for an easy example, classical propositional validity in terms of truth-tables. This is an analysis of validity on which the kind of goodness being ascribed to all substitution instances is failing to have true premises and a false conclusion. But it's a mistake to think that this somehow captures validity *simpliciter*! That's because an argument's failing to have true premises and false conclusion hardly exhausts material goodness. Indeed, complaints from e.g. relevant logicians can be seen as pointing out exactly this. (*Their* mistake is

then to think that their own analysis of validity captures validity *simpliciter* any more than does the classical account.)

One thing to be said in favor of this way of thinking about validity, and about different theoretical accounts of it, is that it explains what might otherwise look like a peculiarity of the history of logic. As Timothy Smiley characterizes the matter:

> The idea of one proposition's following from others – of their implying it – is central to argument. It is, however, an idea that comes with a history attached to it, and those who blithely appeal to an 'intuitive' or 'pre-theoretic' idea of consequence are likely to have got hold of just one strand in a string of diverse theories. [Here I introduce] the main alternatives – to call them rivals would be too strong, since it suggests that they are necessarily in competition with another. Simply put, consequence may be conceived as a relation that is or is not modal in character, and is or is not formal. Thus for Aristotle, consequence is both necessary and formal; for Chrysippus it is necessary but not formal; for Bolzano and Tarski it is formal but not necessary; and for Philo and Russell it is neither necessary nor formal. Conceptions of consequence that are neither necessary nor formal are also needed if justice is to be done to deduction in science, the law and daily life. Cutting across all these other differences there is a perennial controversy about relevance. Does implication always require a full-blooded connection between premises and conclusion, or may it hold simply because of some property of either separately, for example, because the premises are impossible?[32]

A Brandomian explanation for the "string of diverse theories" of consequence would be that material-inferential goodness is the most generic sort of consequence, and that each strand in the string isolates a species of that genus. 'Validity' usually indicates a species of consequence which is necessary and formal, but this species admits of further divisions, such as the classical, intuitionist, and relevant accounts. And just as Smiley says, these accounts are not *rivals* in the sense that none of them is the correct analysis of validity to the exclusion of the others.

Since Brandom does not himself discuss the scorekeeping significance of validity, nor situate it against historical treatments, I have had to reconstruct this position somewhat

[32](Smiley 1999)

from what he does say about logical vocabulary and his definition of validity. However, what I've said here comports well with what Brandom has said elsewhere about different formal logical systems. Quoting again:

> On [my] account, the question of correctness [i.e., which is the right logic?] lapses. *No reasonably well-behaved logic is any more correct than any other* (though some – such as classical logic – have other distinctive virtues such as being able to specify in their own terms the inferential roles of their own vocabulary). The right question is rather which aspects of inferential role do the various kinds of vocabulary serve to make explicit. Thus the classical two-valued conditional lets us say that an inference is good in the sense that it does not have true premises and a false conclusion. (Admittedly, this is a pretty weak endorsement of an inference, but it is still an important good-making property of inferences. Those that do not have it are bad.) The intuitionistic conditional lets us make explicit assessments of an inference as good in the sense that there is a recipe for turning a proof of its premises into a proof of its conclusion. C. I. Lewis's hook of strict implication lets us claim that an inference is good in the sense that it is impossible for its premises to be true without its conclusion being true. And so on. I take it that there is no definite totality of dimensions along which we might want to assess the goodness of inferences, and so no definite totality of possible logical vocabulary.[33]

Although Brandom is here talking about different formally defined *conditionals*, and hence about object-language expressions, the point stands when we go up a level to talk about predications of validity for arguments given in those object languages. All such kinds of validity stand as species to Brandom's generic validity, defined in terms of generic material goodness.

Given that this is the correct way to understand the scorekeeping significance of predications of validity (whether generic or specific), what is thereby *expressed* about the speaker's attitudes? Here it's worth remembering that 'valid' is a philosophers' word, and is used (of arguments) only by those who have been trained to think of language and linguistic performances as things to think and talk *about*, not merely things to use and do. As Ryle puts it:

[33](Brandom 2010, p. 353). Emphasis is mine.

> Such locutions [as '*P* entails *Q*'] are used (roughly) not by the players in the field but by the spectators, critics, and selectors in the grandstand. They belong to the talk of logicians, cross-examiners, and reviewers.[34]

So we should expect that, whatever it is that's expressed by a predication of validity, it's something more sophisticated than what's expressed by the "players in the field" who are *using* logical vocabulary to make valid arguments. But it shouldn't be wholly unrelated to what's expressed by making such arguments, either.

Indeed, I think it's possible to find two aspects of what's expressed by predicating validity of an argument, one of which is shared with what's expressed by making the argument itself, and one of which is novel. What's shared is the endorsement of the material inference from premises to conclusion; this follows from the fact that predicating validity entails commitment to the material goodness of the argument. What's novel is the indefinitely large class of commitments to the material goodness of the argument's substitution-instances, commitments which someone who merely makes the argument need not undertake. (If I make the argument above regarding Monday and Tuesday, I need not commit myself to any other argument of the same form, even though all of them are valid.) There is also a *lack* of commitment to any of the premises or conclusion, which is a further difference from what's expressed by making the argument.

5.3.3 Brandomian bridge principles

We are now in a position to discuss how Brandom's expressive conception yields the sort of bridge principle which MacFarlane is looking for, and which Harman thinks isn't forthcoming. Following the Sellarsian strategy from §3.1, we look to what underlying state of mind is conveyed—or, in Brandom's terms, what attitude of the speaker is expressed—by predication of logical properties and relations, particularly validity.

[34](Ryle 1950, p. 249)

We should start by considering the validity of concrete arguments expressed without variables, such as (once again)

If today is Monday, then tomorrow is Tuesday
Today is Monday

Tomorrow is Tuesday

If the story told in the four stages above is correct, then someone who asserts " 'If today is Monday, then tomorrow is Tuesday; Today is Monday; So tomorrow is Tuesday' is valid" expresses, in the first place, the material goodness of the argument as it stands. And since material goodness is just generic appropriateness of inference, we have the bridge principle

> If 'If today is Monday, then tomorrow is Tuesday; Today is Monday; So tomorrow is Tuesday' is valid, then if you're committed to 'If today is Monday, then tomorrow is Tuesday' and 'Today is Monday', it is appropriate to undertake commitment to 'Tomorrow is Tuesday'.

This isn't very interesting, but it *does* do what MacFarlane asks of a bridge principle: it's a conditional in which the antecedent states a fact about a logical property or relation, and the consequent states a normative directive for reasoning. But the predication of validity doesn't merely express endorsement of the argument as it stands; it also expresses the quantificational commitment to the material goodness of all substitution-instances. Because predications of validity express this whole class of normative attitudes, we can formulate a more general version using variable schemata:

> If 'if P then Q, P; so Q' is valid, then if you're committed to 'if P then Q' and to P, it is appropriate to undertake commitment to Q

But of course the particular logical forms of the premises and conclusion don't really concern us; if we want a bridge principle to express the normative import of validity as

such, we can just introduce new variables to stand for whatever those statements happen to be:

> If 'A, B; so C' is valid, then if you're committed to A and to B, it is appropriate for you to undertake commitment to C.

or, in more ordinary terms:

> If 'A, B; so C' is valid, then if you believe both that A and that B, it is appropriate for you to come to believe that C.

This is the normative import of validity on Brandom's expressive conception of logic, due to the underlying scorekeeping attitudes expressed by predications of validity. It's nothing fancy, but it's a bridge principle, and, more importantly, we have an explanation for why the bridge principle obtains which isn't simply an intuition about how reasoning ought to go.

5.3.4 Harman's challenge

So what about Harman's objection that a bridge principle like this is untenable, given that in some cases it's not appropriate to come to believe C—most saliently, in those cases where it's better to give up belief in A or B? Although Harman is surely right that there are such cases, I don't think they constitute counterexamples to the bridge principle. Consider a parallel sort of objection to Harman's:

> In some cases one ought not believe that C, even though 'A, B, so C' is valid and one believes that A and that B, because the effects of believing that C would be morally or prudentially bad. (Perhaps coming to believe that C would make it easier to be a jerk to one's friends, or to take foolish risks with one's life.)

Here I think we'd all be unmoved by the objection, since the existence of moral or prudential (or other broadly practical) reasons to avoid belief in C doesn't nullify the existence of any *theoretical* or *epistemic* reasons in play. To the extent that a bridge

principle is supposed to articulate such theoretical reasons, it is not made false by pointing out possible practical reasons in conflict with those theoretical reasons.

Similarly, even if we confine ourselves to considering theoretical reasons only, Harman's point that in some cases we ought to give up belief in *A* or *B* rather than come to believe *C*, and hence have reason not to believe *C*, doesn't in such cases nullify the opposing reason to believe *C*, namely that there is a valid argument from premises we believe to the conclusion that *C*. Once again, an analogy from metaethics will help. In a short unpublished paper, Mark Schroeder (2007) points out that we often talk as if there is *no* reason to do something when in fact there is reason to do it, but the reason is very weak, or is obviously outweighed by countervailing reasons. His example is the claim that you have reason to eat your car, which seems clearly false until it's pointed out that your car contains (at least) the recommended daily intake amount of iron. Of course, all things considered you have most reason not to eat your car, but this doesn't nullify the (weak, totally swamped, but real) reason to eat it.

I think we should say the same thing about validity and its normative import. It's *always* a reason to believe *C* that there's a valid argument for it from premises you believe—or as Brandom would put it, it's always appropriate to undertake commitment to *C*—even when, in light of your other beliefs, *C* is absurd. All things considered, you ought to give up belief in one of the premises, but this doesn't nullify the validity-based reason to believe *C*.

Chapter 6

Conclusion

There is still more to say, of course, about the expressive conception of logic—one issue that looms particularly large in my view is the pluralism about logic which follows from Brandom's conception, and how to situate it relative to the more *au courant* pluralism of Beall and Restall (2006), and the ensuing discussion of their proposal. The expressive demarcation of the logical constants is another idea which deserves dedicated attention, and applying Brandom's ideas to the question of the relative priority of proof theory versus model theory is yet another. This is not to mention the questions Brandom's work raises for the psychology and cognitive science of logic—for example the question of what non-logical abilities young children are able to deploy as we train them into the use of terms like 'if'—or questions about the pedagogy of logic, as we teach students to use formal logical systems.

Although I have not even begun to address such questions, in this work I have taken some successful first steps into the elaboration of the expressive conception, and shown how it represents a promising new way of thinking about some old logical and philosophical issues. I hope, too, that at least some of what I've produced here can serve as a helpful companion to anyone studying Brandom's work for the first time, whether or

not they're particularly interested in logic. Other than a reasonably clear exposition of Brandom's views, I take my main contributions in this work to be a thorough investigation of the notion of material inference, a novel expressive interpretation of conjunction and disjunction, a portrait of a psychologistic conception of logic which isn't dead in the water, and a novel account of what validity amounts to within Brandom's conception of logic. This is certainly not the last word on these topics, but I hope it is a good first few.

References

Anscombe, G. E. M. (1958). Modern moral philosophy. *Philosophy 33*, 1–19.

Arnauld, A., & Nicole, P. (1996). *Logic or the Art of Thinking*. (J. V. Buroker, Trans.). Cambridge: Cambridge University Press. (Original work published 1683)

Ayer, A. J. (1952). *Language, truth and logic*. New York: Dover Publications.

Barnes, J. (Ed.). (1984). *The complete works of Aristotle* (Vols. 1–2). Princeton, NJ: Princeton University Press.

Beall, J. C., & Restall, G. (2006). *Logical Pluralism*. Oxford: Clarendon Press.

Belnap, N. (1962). Tonk, plonk, and plink. *Analysis 22*, 130–134.

Benacerraf, P. (1965). What numbers could not be. *Philosophical Review 74*, 47–73.

Benacerraf, P. (1973). Mathematical truth. *Journal of Philospohy 70*, 661–679.

Blackburn, S. (1984). *Spreading the word: Groundings in the philosophy of language*. Oxford: Oxford University Press.

Blackburn, S. (1998). *Ruling passions: A theory of practical reasoning*. Oxford: Oxford University Press.

Bolzano, B. (2004). *On the mathematical method and correspondence with Exner* (P. Rusnock & R. George, Trans.). Amsterdam: Rodopi.

Brandom, R. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Cambridge, MA: Harvard University Press.

Brandom, R. (2000). *Articulating reasons: An introduction to inferentialism*. Cambridge, MA: Harvard University Press.

Brandom, R. (2002). *Tales of the mighty dead: Historical essays in the metaphysics of intentionality*. Cambridge, MA: Harvard University Press.

Brandom, R. (2008). *Between saying and doing: Towards an analytic pragmatism*. Oxford: Oxford University Press.

Brandom, R. (2010a). Reply to Allan Gibbard's "Thought, norms, and discursive practice". In B. Weiss & J. Wanderer (Eds.), *Reading Brandom: On Making It Explicit* (pp. 297–300). New York, NY: Routledge.

Brandom, R. (2010b). Reply to Bernhard Weiss's "What is logic?". In B. Weiss & J. Wanderer (Eds.), *Reading Brandom: On Making It Explicit* (pp. 353–356). New York, NY: Routledge.

Brandom, R. (2010c). Reply to Michael Kremer's "Representation or inference". In B. Weiss & J. Wanderer (Eds.), *Reading Brandom: On Making It Explicit* (pp. 347–352). New York, NY: Routledge.

Brentano, F. (1911). *Von der klassifikation der psychischen phänomene, (Psychologie vom empirischen standpunkt, vol. 2)*. Leipzig: Duncker und Humblot.

Carnap, R. (1937). *Logical syntax of language*. Reprinted 2000, 2001, London: Routledge.

Carroll, L. (1895). What the tortoise said to Achilles. *Mind 4*, 278–280.

Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *The Journal of Philosophy 78*, 67–90.

Churchland, P. M. (1989). A neurocomputational perspective. Cambridge, MA: MIT Press.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated?. *Behavioral and Brain Sciences, 4*(3), 317-331.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 163–228). Oxford: Oxford University Press.

Davidson, D. (1973). On the very idea of a conceptual scheme. *Proceedings and Addresses of the American Philosophical Association 47*, 5–20.

Dennett, D. C. (1971). Intentional systems. *Journal of Philosophy 68*, 87–106.

Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Dennett, D. C. (1991). Real patterns. *Journal of Philosophy 87*, 27–51.

Dretske, F. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT/Bradford Press.

Dummett, M. (1959). Truth. *Proceedings of the Aristotelian Society, New Series, 59*, 141–162.

Dummett, M. (1973a). *Frege: Philosophy of language*. New York: Harper & Row.

Dummett, M. (1973b). The philosophical basis of intuitionistic logic. In *Truth and Other Enigmas* (1978) (pp. 215–247). Cambridge, MA: Harvard University Press.

Enoch, D. (2011). *Taking morality seriously*. Oxford: Oxford University Press.

Etchemendy, J. (1983). The doctrine of logic as form. *Linguistics and Philosophy 6*, 319–334.

Etchemendy, J. (2008). Reflections on consequence. In D. Patterson (Ed.), *New Essays on Tarski and Philosophy* (pp. 263–299). Oxford: Oxford University Press.

Feferman, S. (1999). Logic, logics, and logicism. *Notre Dame Journal of Formal Logic 40*, 31–54.

Field, H. (1972). Tarski's theory of truth. *The Journal of Philosophy 69*, 347–375.

Field, H. (1980). *Science without numbers*. Princeton, NJ: Princeton University Press.

Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT/Bradford Press.

Fodor, J., & Lepore, E. (1992). *Holism: A shopper's guide.* Oxford: Blackwell.

Fodor, J., & Lepore, E. (2001). Brandom's burdens. *Philosophy and Phenomenological Research 63*, 465–481.

Fodor, J., & Lepore, E. (2007). Brandom beleaguered. *Philosophy and Phenomenological Research 74*, 677–691.

Frege, G. (1948). Sense and Reference. *Philosophical Review 57*, 209–230. (Original work published 1892)

Frege, G. (1956). The thought: A logical inquiry. *Mind 65*, 289–311. (Original work published 1918)

Frege, G. (1964). The basic laws of arithmetic: Exposition of the system. (M. Furth, Trans.) Berkeley/Los Angeles, CA: University of California Press. (Original work published 1893)

Frege, G. (1967). Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought. (S. Bauer-Mengelberg, Trans.) In J. van Heijenoort (Ed.), *From*

Frege to G odel: A Source Book in Mathematical Logic, 1879–1931. Cambridge, MA: Harvard University Press. (Original work published 1879)

Frege, G. (1979). Logic (Unpublished essay). (P. Long and R. White, Trans.). In H. Hermes, F. Kambartel, & F. Kaulbach (eds.), *Posthumous Writings* (pp. 126–151). Oxford: Basil Blackwell.

Frege, G. (1980). The foundations of arithmetic. (J. L. Austin, Trans.) Evanston, IL: Northwestern University Press. (Original work published 1884)

Gensler, H. J. (2010). *Introduction to logic* (2nd ed.). New York, NY: Routledge.

Gentzen, G. (1935). Untersuchungen über das logische schliessen (Investigations into logical deduction). *Mathematisches Zeitschrift 39*, 176–210, 405–431.

Gibbard, A. (1992). *Wise choices, apt feelings: A theory of normative judgment*. Cambridge, MA: Harvard University Press.

Gibbard, A. (2003). *Thinking how to live*. Cambridge, MA: Harvard University Press.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European review of social psychology*, 2(1), 83-115.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: a reply to Kahneman and Tversky. *Psychological Review 103*, 592–596.

Gómez-Torrente, M. (2002). The problem of logical constants. *Bulletin of Symbolic Logic 8*, 1–37.

Grover, D., Camp, J., & Belnap, N. (1975). A prosentential theory of truth. *Philosophical Studies 27*, 73–125.

Hacking, I. (1979). What is logic? *Journal of Philosophy 76*, 285–319.

Hale, B., & Wright, C. (2010). Assertibilist truth and objective content: Still inexplicit. In B. Weiss & J. Wanderer (Eds.), *Reading Brandom: On Making It Explicit* (pp. 276–294). New York, NY: Routledge.

Hanna, R. (2006). *Rationality and logic*. Cambridge, MA: MIT Press.

Hanson, W. H. (1997). The concept of logical consequence. *The Philosophical Review 106*, 365–409.

Hare, R. M. (1952). *The language of morals*. Oxford: Clarendon Press.

Harman, G. (1984). *Change in view: Principles of reasoning*. Cambridge, MA: MIT Press.

Hempel, C. G. (1966). *Philosophy of natural science*. Englewood Cliffs, NJ: Prentice-Hall.

Hume, D. (1999). *An enquiry concerning human understanding*. T. L. Beauchamp (Ed.). Oxford/New York: Oxford University Press. (Original work published 1748)

Hurley, P. (2014). *A concise introduction to logic* (12th ed.). Boston, MA: Cengage Learning.

Husserl E. (2001). *Logical investigations*. (J. N. Findlay, Trans.) New York, NY: Routledge. (Original work published 1900)

Jacquette, D. (1997). Psychologism the philosophical shibboleth, *Philosophy and Rhetoric 30*, 312–331.

Kalderon, M. (2005). *Moral Fictionalism*. Oxford: Clarendon Press.

Kalish, D., & Montague, R. (1964). *Logic: Techniques of formal reasoning*. New York, NY: Harcourt, Brace & World, Inc.

Kant, I. (1999). *Critique of pure reason*. P. Guyer & A. W. Wood (Eds., Trans.). Cambridge: Cambridge University Press. (Original work published 1787)

Kapitan, T. (1982). On the concept of material consequence. *History and Philosophy of Logic 3*, 193–211.

Keeler, M. (2004). Using Brandom's framework to do Peirce's normative science: Pragmatism as the game of harmonizing assertions? In K. E. Wolff, H. D. Pfeiffer, & H. S. Delugach (Eds.), *Lecture Notes in Computer Science 3127* (pp. 242–260). Berlin/Heidelberg: Springer-Verlag.

Kibble, R. (2006). Reasoning about propositional commitments in dialogue. *Research on Language and Computation 4*, 179–202.

Kim, J. (2003). The American origins of philosophical naturalism. *Journal of Philosophical Research 28* (Supplement), 83–98.

Kusch M. (1995). *Psychologism: A case study in the sociology of philosophical knowledge*. London: Routledge.

Kusch, M. (2014). Psychologism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Spring 2014 Edition)*. Retrieved from http://plato.stanford.edu/archives/spr2014/entries/psychologism/

Lance, M. (2001). The logical structure of linguistic commitment III: Brandomian scorekeeping and incompatibility. *Journal of Philosophical Logic 30*, 439–464.

Lance, M. N., & Kremer, P. (1994). The logical structure of linguistic commitment I: Four systems of non-relevant commitment entailment. *Journal of Philosophical Logic 23*, 369–400.

Lance, M. N., & Kremer, P. (1994). The logical structure of linguistic commitment II: Systems of relevant commitment entailment. *Journal of Philosophical Logic 25*, 425–449.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic 8*, 339–359.

Linsky, B., & Zalta, E. N. (1995). Naturalized platonism versus platonized naturalism. *The Journal of Philosophy 92*, 525–555.

MacFarlane, J. (2000). *What does it mean to say that logic is formal?* (Unpublished doctoral dissertation). University of Pittsburgh, Pittsburgh, PA.

MacFarlane, J. (2002). Frege, Kant, and the logic in logicism. *Philosophical Review 111*, 25–65.

MacFarlane, J. (2004). In what sense (if any) is logic normative for thought? (Unpublished paper). Delivered at the American Philosophical Association Central Division meeting, 2004. Retrieved from http://johnmacfarlane.net/normativity_of_logic.pdf

MacFarlane, J. (2008). Brandom's demarcation of logic. *Philosophical Topics 36*, 55–62.

MacFarlane, J. (2009). Logical constants. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*. Retrieved from http://plato.stanford.edu/archives/fall2009/entries/logical-constants/

Mackie, J. L. (1977). *Ethics: Inventing right and wrong*. New York, NY: Penguin.

Maddy, P. (2007). *Second philosophy: A naturalistic method*. Oxford: Oxford University Press.

Martin, E. P., & Meyer, R. K. (1982). Solution to the P – W problem. *The Journal of Symbolic Logic 47*, 869–887.

McDowell, J. (1997). Brandom on representation and inference. *Philosophy and Phenomenological Research 57*, 157–162.

McGee, V. (1996). Logical operations. *Journal of Philosophical Logic 25*, 567–580.

Mill, J. S. (1843). *A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence, and the methods of scientific investigation* (8th ed.). New York, NY: Harper & Brothers.

Mill, J. S. (1865). *An examination of Sir William Hamilton's philosophy, and of the principal philosophical questions discussed in his writings* (2nd ed.). London: Longmans, Green, and Co.

Millikan, R. (1984). *Language, thought, and other biological categories*. Cambridge, MA: MIT Press.

Millikan, R. (1989). Biosemantics. *Journal of Philosophy 86*, 281–97.

Oesterle, J. (Trans.). (1962). *Aristotle on interpretation: Commentary by St. Thomas and Cajetan. Mediaeval philosophical texts in translation 11*. Milwaukee: Marquette University Press.

Papineau, D. (1987). *Reality and representation*. Oxford: Basil Blackwell.

Parsons, T. (2010). Introduction. In *An introduction to symbolic logic* (Introduction). Retrieved from http://www.philosophy.ucla.edu/people/faculty/tparsons/Logic%20Text/Text0.pdf

Peacocke, C. (1976). What is a logical constant? *Journal of Philosophy 73*, 221–240.

Pelletier, F. J., Elio R., & Hanson, P. (2008). Is logic all in our heads? From naturalism to psychologism. *Studia Logica 88*, 3–66.

Price, H. (2011). *Naturalism without mirrors*. Oxford: Oxford University Press.

Priest, G. (2008). *An introduction to non-classical logic: From if to is*. Cambridge: Cambridge University Press.

Prior, A. N. (1960). The runabout inference-ticket. *Analysis 21*, 38–39.

Prior, A. N. (1964). Conjunction and contonktion revisited. *Analysis 24*, 191–195.

Quine, W. V. (1951). Two dogmas of empiricism. *Philosophical Review 60*, 20–43.

Quine, W. V. (1982). *Methods of Logic*. Cambridge, MA: Harvard University Press. (1st ed. 1950)

Read, S. (1988). *Relevant logic*. Oxford: Blackwell.

Rosen, G. (1990). Modal fictionalism. *Mind 99*, 327–354.

Rosen, G. (1997). Who makes the rules around here? *Philosophy and Phenomenological Research 57*, 163–171.

Rosen, G. (2001). Brandom on modality, normativity, and intentionality. *Philosophy and Phenomenological Research 63*, 611–623.

Rosenberg, A. (2011). *The atheist's guide to reality: Enjoying life without illusions*. New York, NY: W. W. Norton & Company.

Rosenberg, J. F. (1997) Brandom's *Making It Explicit*: A first encounter. *Philosophy and Phenomenological Research 57* 179–187.

Routley, R., Meyer, R. K., Plumwood, V., & Brady, R. (1983) *Relevant logic and its rivals* (Vol. 1). Atascardero, CA: Ridgeview.

Ryle, G. (1949). *The concept of mind*. London: Hutchinson.

Ryle, G. (1950). 'If', 'so', and 'because'. In M. Black (Ed.), *Philosophical Analysis*. Reprinted in *Collected Papers* (Vol. 2) (pp. 234–249). London: Hutchinson.

Ryle, G. (1954). *Dilemmas: The Tarner lectures 1953*. Cambridge: Cambridge University Press.

Schroeder, M. (2007). The negative reason existential fallacy. (Unpublished paper). Retrieved from http://www-bcf.usc.edu/ maschroe/research/Schroeder_Negative_Reason_Existential_Fallacy.pdf

Schroeder, M. (2008). *Being for*. Oxford: Oxford University Press.

Sellars, W. (1949). Language, rules and behavior. In S. Hook (Ed.), *John Dewey: Philosopher of Science and Freedom*, (pp. 289–315). New York: Dial Press. Reprinted in J. F. Sicha (Ed.), *Pure Pragmatics and Possible Worlds — The Early Essays of Wilfrid Sellars*, (pp. 129–155). Atascadero, CA: Ridgeview Publishing Co (1980).

Sellars, W. (1953). Inference and meaning. *Mind 62*, 313–338.

Sellars, W. (1956). Empiricism and the philosophy of mind. In H. Feigl & M. Scriven (Eds.), *Minnesota Studies in the Philosophy of Science* (Vol. 1), (pp. 253–329). Minneapolis, MN: University of Minnesota Press.

Shapiro, S. (1991). *Foundations without foundationalism: A case for second-order logic*. Oxford: Oxford University Press.

Shafer-Landau, R. (2003). *Moral realism: A defence*. Oxford: Oxford University Press.

Sher, G. (1991). *The bounds of logic: A generalized viewpoint*. Cambridge, MA: MIT Press.

Sher, G. (2013). The foundational problem of logic. *The Bulletin of Symbolic Logic 19*, 145–198.

Smiley, T. (1999). Consequence, conceptions of. In E. Craig & E. Craig, (Eds.), *Concise Routledge encyclopedia of philosophy* (p. 169). London: Routledge.

Smith, R. (1989). *Aristotle's Prior Analytics*. (R. Smith, Trans.). Indianapolis: Hackett.

Stelzner, W. (2005). Psychologism, universality and the use of logic. In J. Faye, P. Needham, U. Scheffler, & M. Urchs (Eds.), *Nature's Principles* (pp. 269–288). Berlin: Springer.

Stevenson, C. (1944). *Ethics and language*. New Haven, CT: Yale University Press.

Strawson, P. F. (1959). *Individuals: An essay in descriptive metaphysics*. New York, NY: Routledge.

Striker, G. (2009). *Aristotle, Prior Analytics I*. G. Striker, Trans.). Oxford: Clarendon Press.

Swinburne, R. (1977). *The coherence of theism*. Oxford: Oxford University Press.

Tarski, A. (1933). The concept of truth in formalized languages. In J. Corcoran (Ed.), *Logic, Semantics, Metamathematics* (2nd ed.). (1983). Indianapolis, IN: Hackett.

Tarski, A. (1936). On the concept of logical consequence. In J. Corcoran (Ed.), *Logic, Semantics, Metamathematics* (2nd ed.). (1983). Indianapolis, IN: Hackett.

Tarski, A. (1986). What are logical notions? *History and Philosophy of Logic 7*, 143–154.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, New Series, 185*, 1124–1131.

Unger, P. (2006). *All the power in the world*. Oxford: Oxford University Press.

Warmbrod, K. (1999). Logical constants. *Mind 108*, 503–538.

Wason, P. C. (1966). Reasoning. In B.M. Foss (Ed.), *New horizons in psychology* (Vol. I). Harmondsworth: Penguin.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology 20* 273–281.

Weiss, B., & Wanderer, J. (Eds.). (2010). *Reading Brandom: On making it explicit*. New York, NY: Routledge.

Whiting, D. J. (2008). Conservatives and racists: Inferential role semantics and pejoratives. *Philosophia 36*, 375–388.

Williams, B. (1972). *Morality: An introduction to ethics*. New York, NY: Harper & Row.