

---

# **Making It Explicit**

## **Reasoning, Representing, and Discursive Commitment**

---

---

**Robert B. Brandom**

Harvard University Press  
Cambridge, Massachusetts  
London, England

---

To Wilfrid Sellars and Richard Rorty

without whom most of it would not even be *implicit*

Copyright © 1994 by the President and Fellows of Harvard College

All rights reserved

Printed in the United States of America

Page xxiv constitutes an extension of the copyright page.

Fourth printing, 2001

First Harvard University Press paperback edition, 1998

*Library of Congress Cataloging-in-Publication Data*

Brandom, Robert.

*Making it explicit : reasoning, representing, and discursive commitment /*  
Robert B. Brandom.

p. cm.

Includes bibliographical references and index.

ISBN 0-674-54319-X (cloth)

ISBN 0-674-54330-0 (pbk.)

1. Language and languages—Philosophy. 2. Semantics (Philosophy)  
3. Pragmatics. 4. Representation (Philosophy) I. Title.

P106.B694 1994

121'.68—dc20

93-50631

---

. . . both a new world,  
and the old made explicit . . .

We shall not cease from exploration  
And the end of all our exploring  
Will be to arrive where we started  
And know the place for the first time.

T. S. ELIOT, "Four Quartets"



---

# Contents

---

<i>Preface</i>	xii
----------------	-----

## PART ONE

<b>1. Toward a Normative Pragmatics</b>	<b>3</b>
I. Introduction	3
II. From Intentional State to Normative Status	7
III. From Norms Explicit in Rules to Norms Implicit in Practices	18
IV. From Normative Status to Normative Attitude	30
V. From Assessment to the Social Institution of Norms	46
VI. From Intentional Interpretation to Original Intentionality	55
Appendix: Wittgenstein's Use of <i>Regel</i>	64
<b>2. Toward an Inferential Semantics</b>	<b>67</b>
I. Content and Representation	67
II. The Priority of the Propositional	79
III. Conceptual Classification and Inference	85
IV. Material Inference, Conceptual Content, and Expression	94
V. Circumstances and Consequences of Application	116
VI. Conclusion	132

<b>3. Linguistic Practice and Discursive Commitment</b>	<b>141</b>
I. Intentional States and Linguistic Practices	141
II. Deontic Status and Deontic Attitudes	157
III. Asserting and Inferring	167
IV. Scorekeeping: Pragmatic Significance and Semantic Content	180
<b>4. Perception and Action: The Conferral of Empirical and Practical Conceptual Content</b>	<b>199</b>
I. Assertions as Knowledge Claims	199
II. Reliability	206
III. Observation Reports and Noninferential Authority	213
IV. Rational Agency	229
V. Practical Reasoning: Inferences from Doxastic to Practical Commitments	243
VI. Intentions	253
<b>PART TWO</b>	
<b>5. The Expressive Role of Traditional Semantic Vocabulary: 'True' and 'Refers'</b>	<b>275</b>
I. From Inference to Truth, Reference, and Representation	275
II. Truth in Classical Pragmatism	285
III. From Pragmatism to Prosentences	299
IV. Reference and Anaphorically Indirect Descriptions	305
V. The Function of Traditional Semantic Vocabulary Is Expressive, Not Explanatory	322
<b>6. Substitution: What Are Singular Terms, and Why Are There Any?</b>	<b>334</b>
I. Multivalued Logic and Material Inference	334
II. Substitution, Sentential Embedding, and Semantic Roles	346
III. Subsentential Expressions	360
IV. What Are Singular Terms?	367
V. Why Are There Singular Terms?	376
VI. Objections and Replies	384
VII. Conclusion	399
<i>Appendix I: From Substitutional Derivation of Categories to Functional Derivation of Categories</i>	404
<i>Appendix II: Sentence Use Conferring the Status of Singular Terms on Subsentential Expressions—An Application</i>	409
<b>7. Anaphora: The Structure of Token Repeatables</b>	<b>413</b>
I. Frege's <i>Grundlagen</i> Account of Picking Out Objects	413
II. Definite Descriptions and Existential Commitments	432
III. Substitution, Token Recurrence, and Anaphora	449
IV. Deixis and Anaphora	459

V. Interpersonal Anaphora and Communication <i>Appendix: Other Kinds of Anaphora—Paychecks, Donkeys, and Quantificational Antecedents</i>	473 490
<b>8. Ascribing Propositional Attitudes: The Social Route from Reasoning to Representing</b>	<b>495</b>
I. Representation and <i>De Re</i> Ascription of Propositionally Contentful Commitments	495
II. Interpretation, Communication, and <i>De Re</i> Ascriptions	508
III. <i>De Re</i> Ascriptions and the Intentional Explanation of Action	520
IV. From Implicit Attribution to Explicit Ascription	529
V. Epistemically Strong <i>De Re</i> Attitudes: Indexicals, Quasi-Indexicals, and Proper Names	547
VI. The Social-Perspectival Character of Conceptual Contents and the Objectivity of Conceptual Norms <i>Appendix: The Construction and Recursive Interpretation of Iterated Ascriptions That Mix <i>De Dicto</i> and <i>De Re</i> Content Specifications</i>	584 608
<b>9. Conclusion</b>	<b>614</b>
I. Two Concepts of Concepts	614
II. Norms and Practices	623
III. We Have Met the Norms, and They Are Ours	639
<i>Abbreviations</i>	653
<i>Notes</i>	655
<i>Index</i>	717



---

## Preface

---

We work in the dark—we do what we can—we give what we have. Our doubt is our passion and our passion is our task. The rest is the madness of art.

HENRY JAMES, "The Middle Years"

---

That old philosopher Fred Allen used to say he could not understand why someone would spend years writing a novel, when for a few dollars you could buy one practically anywhere. A similar remark might be made about contributions to that peculiar genre of creative nonfiction writing to which philosophical works such as this one belong. This book is an investigation into the nature of *language*: of the social practices that distinguish us as rational, indeed logical, concept-mongering creatures—knowers and agents. This is of course a topic that has been much explored by philosophers, both the mighty dead and the ablest contemporary thinkers. Surrounded as we are by the riches they have bequeathed, it is hard to avoid asking why one should bother reading—let alone writing—yet another such work. This question may seem all the more urgent inasmuch as it is acknowledged (indeed, some pains are taken to show) that the basic building blocks out of which this account is constructed—its motivating insights, commitments, and strategies—are not novel or original.

Still, though the ways of thinking and talking about thinking and talking presented here arise naturally out of a reading of the philosophical tradition (above all Kant, Frege, and Wittgenstein) and of its development by more recent thinkers, both that tradition and its significance for us today are by current standards seen decidedly on a slant. As a result, the story told in these pages comes at familiar things from an unfamiliar direction. Its promise lies

in the sort of added depth and dimension that only binocular vision affords; that is the point of laying a substantially different conceptual perspective alongside our more accustomed line of sight. In keeping with this understanding of the sort of payoff that can be hoped for, the body of the work aims to set criteria of adequacy for a theory of discursive practice, motivate the approach adopted, work the model out in detail, and apply it. The idea is to show what kind of understanding and explanatory power one gets from talking this way, rather than to argue that one is somehow rationally *obliged* to talk this way.

Of course I take it that the claims made in what follows are true; I endorse those assertions; they express my commitments. One of the central tenets of the account of linguistic practice put forward here is that the characteristic *authority* on which the role of assertions in communication depends is intelligible only against the background of a correlative *responsibility* to vindicate one's entitlement to the commitments such speech acts express. It is possible to secure entitlement to the commitments (assertional, inferential, and referential) implicit in an idiom without gainsaying the possibility of entitlement to a different one. But even such a modest justificatory project is of interest only to someone who both understands the commitments in question and has some reason to *want* to become entitled to talk in ways that presuppose them.

Both of these ends are served by starting the story with some historical lessons. Accordingly, Chapters 1 and 2 form an entrance hall to the rest of the edifice, one whose main architectural features are made more noticeable by the judicious placement of ancestor portraits. The same figure appears on many walls, almost always recognizable, but often portrayed from unusual vantage points (from behind, from above) or highlighting something other than the familiar face. In particular, the portrait of Frege will seem to some to be like one of those odd photographs of a reclining figure taken with the lens so close that the subject's left foot assumes gigantic proportions, dwarfing the rest of the individual, whose head and torso dwindle to the dimensions of insignificant appendages. Nonetheless, the tradition that is retrospectively constituted by the unusual emphases and filiations to be found here is meant to be coherent and compelling in its own terms. It is not just whatever rewriting of the history of philosophy happens to be needed to make the waning years of the twentieth century safe for the views I put forward. Rather, those views have the shape they do because of this reading of how we got to where we are.

One of the overarching methodological commitments that orients this project is to explain the *meanings* of linguistic expressions in terms of their *use*—an endorsement of one dimension of Wittgenstein's pragmatism. For although he drove home the importance of such an approach, other features of his thought—in particular his theoretical quietism—have discouraged his admirers from attempting to work out the details of a theory of meaning or,

for that matter, of use. One result has been a substantial disjunction between *semantic* theorizing (about the sorts of contents expressed by various locutions), on the one hand, and *pragmatic* theorizing (about the linguistic practices in which those locutions are employed), on the other. The explanatory strategy pursued here is to begin with an account of social practices, identify the particular structure they must exhibit in order to qualify as specifically *linguistic* practices, and then consider what different sorts of semantic contents those practices can confer on states, performances, and expressions caught up in them in suitable ways. The result is a new kind of conceptual-role semantics. It is at once firmly rooted in actual practices of producing and consuming speech acts, and sufficiently finely articulated to make clear how those practices are capable of conferring the rich variety of kinds of content that philosophers of language have revealed and reveled in.

Claims about the relations between meaning and use have a clear sense only in the context of a specification of the vocabulary in which that use is described or ascribed. At one extreme, use clearly determines meaning in the strongest possible sense if admissible specifications of the use can include such phrases as "using the word 'not' to express negation" or "using the term 'Julius Caesar' to refer to Julius Caesar." At an opposite extreme, if admissible specifications of use are restricted to descriptions of the movements of particles expressed in the vocabulary of physics, not only will the use, so described, fail to settle *what* is meant or expressed by various noises or inscriptions, it will fail to settle even *that* anything is meant or expressed by them. The specification of use employed here is neither so generous as to permit semantic or intentional vocabulary nor so parsimonious as to insist on purely naturalistic vocabulary.

Instead, it makes essential use of *normative* vocabulary. The practices that confer propositional and other sorts of conceptual content implicitly contain norms concerning how it is *correct* to use expressions, under what circumstances it is *appropriate* to perform various speech acts, and what the *appropriate* consequences of such performances are. Chapter 1 introduces and motivates this normative pragmatics, which is rooted in considerations advanced by Kant, Frege, and Wittgenstein. No attempt is made to eliminate, in favor of nonnormative or naturalistic vocabulary, the normative vocabulary employed in specifying the practices that are the use of a language. Interpreting states, performances, and expressions as semantically or intentionally contentful is understood as attributing to their occurrence an ineliminably normative pragmatic significance.

Though this normative dimension of linguistic practice is taken to be ineliminable, it is not treated as primitive or inexplicable. It is rendered less mysterious in two ways. First, linguistic norms are understood as *instituted* by social-practical activity. The pragmatic significances of different sorts of speech acts are rendered theoretically in terms of how those performances

affect the *commitments* (and *entitlements* to those commitments) acknowledged or otherwise acquired by those whose performances they are. The norms implicit in linguistic practice are accordingly presented in a specifically *deontic* form. But these deontic statuses are understood in turn as a form of *social* status, instituted by the practical attitudes of those who attribute and acknowledge such statuses.

The natural world does not come with commitments and entitlements in it; they are products of human activity. In particular, they are creatures of the *attitudes* of taking, treating, or responding to someone in practice as committed or entitled (for instance, to various further performances). Mastering this sort of norm-instituting social practice is a kind of practical know-how—a matter of keeping deontic *score* by keeping track of one's own and others' commitments and entitlements to those commitments, and altering that score in systematic ways based on the performances each practitioner produces. The norms that govern the use of linguistic expressions are implicit in these deontic scorekeeping practices.

The second way norms are rendered less mysterious is by explaining exactly what is expressed by normative vocabulary. Beginning with basic deontic scorekeeping attitudes and the practices that govern them, an account is offered of how locutions must be used in order to express explicitly the very normative notions—*is committed*, *is permitted*, *ought*, and so on—that are appealed to in laying out the normative pragmatics. This is an explication of explicitly normative conceptual contents in terms of implicitly normative practices, rather than a reduction of normative terms to nonnormative ones. It illuminates the normative dimension of discursive practice in line with the methodological principle that implicit structures are often best understood by looking at how they can be made explicit.

The first step in the project is accordingly the elaboration of a pragmatics (a theory of the use of language) that is couched in terms of practical scorekeeping *attitudes* of attributing and acknowledging deontic *statuses* of commitment and entitlement. The pragmatic significance of performances—eventually, speech acts such as assertions—is then understood to consist in the difference those performances make to the commitments and entitlements attributed by various scorekeepers. The next step is to say what structure such a set of social practices must have in order to qualify as specifically *discursive* practice. This is a matter of moving from pragmatics to semantics. The defining characteristic of discursive practice is the production and consumption of specifically *propositional* contents. It is argued in Chapter 2 that propositional contentfulness should be understood in terms of *inferential* articulation; propositions are what can serve as premises and conclusions of inferences, that is, can serve as and stand in need of *reasons*. Chapter 3 describes (in deontic scorekeeping terms) a model of social practices of giving and asking for reasons—specifically *linguistic* discursive prac-

tices, which suffice to confer propositional contents on states, attitudes, performances, and expressions that play suitable roles in those practices.

This account of the conferral of semantic content by inferentially articulated, social scorekeeping practice is further generalized in two different directions. First, it is shown how this model can be applied in order to understand not only linguistic meanings but intentional contents generally. The propositional contentfulness of *beliefs*, no less than of *claims*, should be understood in terms of their role in reasoning of various kinds. The inferentially articulated commitments expressed by assertional speech acts are *doxastic* commitments. Much of the theoretical work done by the concept of belief can be done instead by appeal to this sort of deontic status, and to the practical scorekeeping attitudes of acknowledging or undertaking such commitments. A social, linguistic account of *intentionality* is accordingly elaborated in Chapter 3.

It is extended in Chapter 4 to incorporate treatments of *perception* and *action* and of the contribution those phenomena make to the empirical and practical dimensions of the propositional contents of the states, acts, and attitudes involved in them. It is not denied that it makes sense to talk about nonlinguistic creatures as having intentional states, but it is claimed that our understanding of such talk is parasitic on our understanding of the sort of full-blooded linguistic intentionality characteristic of states and attitudes that only beings who engage in discursive social practices can have. This story amounts, then, to an account of the relations of mindedness—in the sense of sapience rather than mere sentience—to behavior. As in the parallel case of meaning and use, the clarification of these relations must begin with a determination of what vocabulary it is admissible to use in specifying the relevant behavior; again there is a spectrum of possibilities, from allowing intentional vocabulary with semantic locutions, such as “acting as if one believed that snow is white,” ranging down to restrictions to physicalistic or other naturalistic vocabulary, such as “one’s left wrist rotating twenty degrees.” The *via media* pursued here eschews intentional or semantic specifications of behavior but permits normative and therefore social specifications of what is in fact linguistic behavior.

Where the first sort of generalization involves moving from consideration of *language* to consideration of *mind*, from talking to thinking and believing, the second involves moving from an account of the practices that constitute treating something as *propositionally* contentful to the practices that constitute treating something as *conceptually* contentful in a broader sense. In Chapter 6 the notion of *substitution* and substitutional inferences is used to show how expressions such as singular terms and predicates, which cannot directly play the inferential role of premise or conclusion in an argument, nonetheless can play an *indirectly* inferential role in virtue of their systematic contributions to the directly inferential roles of sentences in which they

occur. In Chapter 7 the notion of *anaphora* (whose paradigm is the relation between a pronoun and its antecedent) and anaphoric inheritance of substitutional commitment is used to show how even unrepeatable expressions such as demonstrative tokenings play substitution-inferential roles and hence express conceptual contents. The result is a kind of conceptual-role semantics that is distinguished first by the nature of the functional system with respect to which such roles are individuated and attributed: what is appealed to is role in the implicitly normative linguistic social practices of a community, rather than the behavioral economy of a single individual. It is also different from familiar ways of using the notion of conceptual role in conceiving of the conceptual in terms of specifically *inferential* articulation, and in its elaboration of the fundamental *substitutional* and *anaphoric* substructures of that inferential articulation.

This semantic explanatory strategy, which takes *inference* as its basic concept, contrasts with the one that has been dominant since the Enlightenment, which takes *representation* as its basic concept. The inferentialist approach is by no means without precedent—though it has been largely a minority platform. Indeed, the distinction canonically drawn between Continental rationalists such as Spinoza and Leibniz, on the one hand, and British empiricists such as Locke and Hume, is for many purposes more perspicuously rendered as a distinction between those endorsing an inferentialist order of explanation and those endorsing a representationalist order of explanation. The elements of the contemporary inferentialist program are extracted (in Chapter 2) from Frege of the *Begriffsschrift*, Sellars, and some of Dummett's writings.

The complementary theoretical semantic strategies of representationalism and inferentialism are bound by the same pair of general explanatory obligations: to explicate the concept treated as primitive, and to offer an account of other semantic concepts in terms of that primitive. The representationalist tradition has developed good answers to the second sort of concern, primarily by employing a variety of set-theoretic methods to show how proprieties of inference can be determined by representational properties of the claims that serve as their premises and conclusions. The explanatory challenge to that tradition lies rather in the first sort of demand, in saying what it is for something to have representational content, and in what the grasp or uptake of that content by speakers and thinkers consists. As the inferentialist program is pursued here, the proprieties of inference that serve as semantic primitives are explicated in the pragmatics; they are implicit in the practices of giving and asking for reasons. The major explanatory challenge for inferentialists is rather to explain the representational dimension of semantic content—to construe *referential* relations in terms of *inferential* ones.

The second part of the book responds to this challenge. Chapter 5 explains the expressive role of traditional representational semantic vocabulary. An account is offered there of the use of the sort of expression of which 'true'

and ‘refers’ are paradigmatic. Following the lead of the prosentential approach to truth, the key semantic concept employed in that unified account is *anaphora*. Chapter 7 then explains anaphoric relations in terms of the substitution-inferential structure of discursive scorekeeping elaborated in Chapter 6. Chapter 6 also offers an account in those terms of what it is for claims—which are understood in the first instance (in Chapter 3) as what can serve as premises and conclusions of inferences—to be and be understood to be about *objects*, and to characterize them as having properties and standing in relations.

The primary treatment of the representational dimension of conceptual content is reserved for Chapter 8, however. There the representational properties of semantic contents are explained as consequences of the essentially *social* character of inferential practice. Words such as the ‘of’ that expresses intentional directedness, and ‘about’ and ‘represents’ in their philosophically significant uses, have the expressive role they do—making representational relations explicit—in virtue of the way they figure in *de re* ascriptions of propositional attitudes. These are the tropes used to say explicitly what someone is thinking *about*, what a belief *represents*, what a claim is true *of*. Chapter 8 offers a discursive scorekeeping account of the practices that constitute using locutions to express such *de re* ascriptions, and hence of how expressions must be used in order to mean ‘of’, ‘about’, or ‘represents’. This account of what is expressed by the fundamental explicitly representational locutions makes possible an explanation of the *objectivity* of concepts. It takes the form of a specification of the particular sort of inferential structure social scorekeeping practices must have in order to institute objective norms, according to which the correctness of an application of a concept answers to the facts about the object to which it is applied, in such a way that anyone (indeed *everyone*) in the linguistic community may be wrong about it.

In summary, in the theoretical place usually occupied by the notion of *intentional states*, the pragmatics presented here elaborates a conception of *normative statuses*; in the place usually occupied by the notion of *intentional interpretation*, it puts *deontic scorekeeping*—that is, the social practices of attributing and acknowledging commitments and entitlements, which implicitly institute those statuses. The theoretical work typically done by semantic assessments according to correctness of *representation* and satisfaction of *truth* conditions is done by assessments of proprieties of *inference*. Semantic articulation is attributed and acknowledged by keeping score not only of directly inferential commitments, which relate sentential (that is, claimable or believable) contents, but also of indirectly inferential *substitutional* and *anaphoric* commitments, which relate the subsentential contents of expressions of other grammatical categories.

The pragmatics and semantics maintain particularly intimate relations throughout. The aim is always to show how some bit of vocabulary must be used—the significance its utterance must have in various circumstances, the

practical scorekeeping attitudes its usage must elicit and be elicited by—in order for it to express a certain kind of semantic content: to be being taken or treated in practice by the linguistic community *as* a conditional, a singular term, a bit of normative vocabulary, a propositional-attitude-ascribing locution, and so on. A fundamental methodological criterion of adequacy of the account is that the theorist not attach semantic contents to expressions by *stipulation*; it must always be shown how such contents can be *conferred* on expressions by the scorekeeping activities the theorist attributes to the linguistic practitioners themselves. That is, the aim is to present conditions on an interpretation of a community as discursive scorekeepers that are sufficient (though perhaps not necessary) to ensure that interpreting the community as engaged in *those* implicitly normative practices *is* interpreting them as taking or treating their speech acts as expressing the sorts of semantic contents in question.

The obligation to say what it is about the use of locutions in virtue of which they express various sorts of content dictates that the master concept articulating the relation between the pragmatic and semantic portions of the theory is that of *expression*. To express something is to make it *explicit*. What is explicit in the fundamental sense has a *propositional* content—the content of a claim, judgment, or belief (claimable, judgeable, believable contents). That is, making something explicit is *saying* it: putting it into a form in which it can be given as a reason, and reasons demanded for it. Putting something forward in the explicit form of a claim is the basic move in the game of giving and asking for reasons.

The relation of expression between what is *implicit* in what practitioners *do* and what is *explicit* in what they *say* structures the story told here at two different levels. At the basic level, the question is how the capacity to entertain principles, and so to know *that* something is the case, arises out of the capacity to engage in practices—to know *how* to do something in the sense of being *able* to do it. What must practitioners be able to *do* in order to be able thereby to *say that* things are thus and so—that is, to express something explicitly? The explanatory force of a response to this question can be judged by the constraints that are acknowledged on the vocabulary in which those practical capacities are specified; normative vocabulary is employed here, but intentional vocabulary (which would permit at the outset the ascription of propositionally contentful states, attitudes, and performances) is not. The first level of the account of expression accordingly consists in explaining—making theoretically explicit—the implicit structure of linguistic practices in virtue of which they count as making anything explicit at all.

The second level of the account of expression consists in working out a theory of the expressive role distinctive of *logical* vocabulary. The claim is that 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. Conditionals serve as a paradigm illustrating this expressive role. According to the inferential approach to semantics and the deontic scorekeeping approach to pragmatics, practitioners confer determinate propositional contents on states and expressions in part by their scorekeeping practice of treating the acknowledgment of one doxastic commitment (typically through assertional utterance of a sentence) as having the pragmatic significance of an undertaking of further commitments that are related to the original commitment as its inferential consequences. At the basic level, treating the claim expressed by one sentence as an inferential consequence of the claim expressed by another sentence is something practitioners can *do*, and it is because such practical attitudes can be implicit in the way they respond to each other's performances that their sentences come to mean what they do. With the introduction of conditional locutions linking sentences, however, comes the expressive power to *say explicitly that* one claim is a consequence of another. The expressive role distinctive of conditionals is making *implicit* inferential commitments *explicit* in the form of declarative sentences, the assertion of which acknowledges a propositionally contentful doxastic commitment. In a similar way, at the basic level, scorekeepers can treat the claims expressed by two sentences as incompatible—namely by treating commitment to one as in practice precluding entitlement to the other. The introduction of a locution with the expressive power of negation makes it possible to express such implicit practical scorekeeping attitudes explicitly—by *saying* that two claims are incompatible (one entails the negation of the other). Identity and quantificational expressions are analyzed on this model as making explicit the substitutional relations characteristic of singular terms and predicates respectively, and further locutions are considered that play a corresponding expressive role in making anaphoric relations explicit.

So an *expressive* theory of logic is presented here. On this view, 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. Logical vocabulary endows practitioners with the expressive power to make explicit as the contents of claims just those implicit features of linguistic practice that confer semantic contents on their utterances in the first place. Logic is the organ of semantic self-consciousness. It brings out into the light of day the practical attitudes that determine the conceptual contents members of a linguistic community are able to express—putting them in the form of explicit claims, which can be debated, for which reasons can be given and alternatives proposed and assessed. The formation of concepts—by means of which practitioners can come to be aware of anything at all—comes itself to

be something of which those who can deploy logical vocabulary can be aware. Since plans can be addressed to, and intentional practical influence exercised over, just those features of things of which agents can become explicitly aware by the application of concepts, the formation of concepts itself becomes in this way for the first time an object of conscious deliberation and control.

Explaining the features of the use of *logical* vocabulary that confer its characteristic sort of semantic content is accordingly explaining how the sort of expressive power the theorist requires to explain the features of the use of *nonlogical* vocabulary that confer semantic content on it can become available to those whose linguistic practice is being theorized about. It is this fact that sets the expressive scope of the project pursued here. The aim is twofold: to make explicit deontic scorekeeping social practices that suffice to confer conceptual contents on nonlogical sentences, singular terms, and predicates in general; and to make explicit the deontic scorekeeping social practices in virtue of which vocabulary can be introduced as playing the expressive roles characteristic of a variety of particular logical locutions. How much logical vocabulary is worth reconstructing in this fashion? In this project, neither more nor less than is required to make explicit *within* the language the deontic scorekeeping social practices that suffice to confer conceptual contents on nonlogical vocabulary in general. At that point it will have been specified what practices a theorist must attribute to a community in order to be interpreting its members as engaging not just in specifically *linguistic* practices but in linguistic practices that endow them with sufficient expressive power to *say* how their practices confer conceptual content on their states, attitudes, performances, and expressions. That is, they will be able to express the theory offered here.

To make the semantic theory explicit requires logical vocabulary capable of expressing inferential, substitutional, and anaphoric relations. This vocabulary corresponds pretty well to the language of standard first-order logic, with the addition of classical semantic vocabulary. To make the pragmatic theory explicit requires logical vocabulary expressing the endorsement of norms generally, and the attribution and acknowledgment of the deontic statuses of commitment and entitlement in particular. The discussion of action in Chapter 4 includes an account of the use of vocabulary that makes norms explicit, and Chapter 8 explains how the social-practical attitudes of attributing and acknowledging deontic statuses (paradigmatically doxastic commitment) are made explicit by the use of propositional-attitude-ascribing locutions such as the regimented ‘. . . is committed to the claim that . . .’, which does duty here for ‘. . . believes that . . .’. Along this expressive dimension, the project eats its own tail, or lifts itself up by its own bootstraps—presenting an explanation of what it is to say something that is powerful enough to explain what it itself is saying.

Interpreting the members of a community as engaging in specifically

discursive practices, according to the view put forward here, is interpreting them as engaging in social practices that include treating some performances as having the pragmatic significance of *assertions*. For it is in terms of the constellation of inferentially articulated commitments and entitlements characteristic of the making (staking) of claims that the notion of specifically *propositional* contentfulness is to be understood. Since all other varieties of conceptual contentfulness derive (substitutionally) from the propositional, this is to say that the application of concepts is a linguistic affair—not in the sense that one must be talking in order to do it, but in the sense that one must be a player of the essentially linguistic game of giving and asking for reasons in order to be able to do it. There can be sets of practices that are linguistic in this sense but that do not incorporate the expressive resources provided by logical vocabulary. Indeed, the way the use of such vocabulary in making explicit what is implicit in the use of nonlogical vocabulary is specified is by showing what would be required to introduce vocabulary with that expressive function into idioms that did not already contain it. The contribution made by logical locutions to the reflective processes in virtue of which the evolution of our concepts and commitments qualifies as *rational* is so important, however, that linguistic practices that at least permit their introduction form a special class. In a weak sense, any being that engages in linguistic practices, and hence applies concepts, is a *rational* being; in the strong sense, rational beings are not only linguistic beings but, at least potentially, also *logical* beings. This is how we should understand ourselves: as beings that meet this dual expressive condition.

It turns out that there is a surprising connection between being a rational creature—in the sense that includes the possibility of using the expressive resources of specifically *logical* vocabulary to reflect on one's conceptual content-conferring linguistic practices—on the one hand, and the structure of the *facts* that make up the world one can become aware of by applying those concepts, on the other. Rational beings live in a world of propertied and related particulars. Chapter 6 presents an expressive deduction of the necessity of this structure; it shows not why there is something rather than nothing but why what there is must come in the form of *things*; it shows why judgments or beliefs—the commitments expressed by claims—must in the basic case be about particulars (paradigmatically objects) and their properties and relations.

Particular objects are what is referred to by singular terms, and the demonstration proceeds by showing that the only semantically significant sub-sentential structure that is compatible with the introduction of logical vocabulary is one that decomposes basic sentences into singular terms and predicates. This would not be a surprising result if the logical vocabulary appealed to included identity and quantificational locutions, for (it will be argued) these have precisely the expressive role of making explicit in the form of claims the substitution-inferential commitments characteristic of

singular terms and predicates. But the result presented here is much stronger: any discursive practices that permit the introduction even of *sentential* logical operators such as negation and conditionals require that any *subsentential* substitutional structure be of the term-predicate variety. Thus the investigation of the nature and limits of the explicit expression in principles of what is implicit in discursive practices yields a powerful transcendental argument—a formal answer to the question, Why are there objects? that turns on a deep relation between the expressive capacities required to think critically about the inferential connections among claims and the structures in virtue of which those claims are properly understood as characterizing objects as having properties and standing in relations.

This is a long book. Its length is a consequence of the demands made by its governing methodological aspirations: to eschew representational primitives, to show how content is related to use, and to achieve self-referential expressive completeness. The first is pursued by elaborating inferentialist and expressivist alternatives to the representationalist idiom for thinking and talking about thinking and talking that has been so well worked out over the last three centuries. The aim is not to replace that familiar idiom but to enrich it. The promised enrichment is of two sorts. First, there is the greater depth of field afforded by the stereoscopic vision made available by an alternative to familiar ways of talking about intentional phenomena. Second, there is the grounding and illumination of representational tropes secured by displaying the implicit features of discursive practice that are expressed explicitly by their use. Doing this requires that both the pragmatics and the semantics be developed in a reasonable amount of detail. The account of norm-instituting social practices must appeal to capacities that are plausibly available in primitive prelinguistic cases, and yet provide raw materials adequate for the specification of sophisticated linguistic practices, including logical ones. The account of the semantic contents conferred by those practices must encompass expressions of grammatical categories that are reasonably well understood already within the representationalist tradition—for example predicates, definite descriptions, proper names, familiar sorts of logical expressions, and whatever other kinds of locutions are required to make the processes by which content is conferred explicit within the linguistic practices being modeled.

Chapters 3 and 4 present the core theory—the model according to which a pragmatics specifying the social practices in which conceptual norms are implicit and a broadly inferential semantics are combined. It is here that sufficient conditions are put forward for the practices a community is interpreted as engaging in to count as according performances the pragmatic significance characteristic of *assertions*—and hence for those practices to count as conferring specifically *propositional* contents. Everything else in the book either leads up to the presentation of this model or elaborates and

extracts consequences from it. These chapters can be read on their own; the cost of omitting the first two chapters, which motivate the approach to pragmatics and semantics pursued here (in part by a rational reconstruction of the history of discussions of conceptual norms and contents), is that without this conceptual and historical background, one will not understand why things are done as they are here, rather than in some more familiar way. The time spent developing that motivation, however, means that one must wait a while for the actual theory to show itself.

The cost of missing Part 2 (Chapters 5 through 8) would be largely that one would then not see what the model can do, what it is good for. The most essential bit is Chapter 8, for that is where the *representational* dimension of discursive practice is explained in terms of the interaction of the *social* and the *inferential* articulation of the communication of reasons for belief and action. It is this interaction that is appealed to there also to make intelligible how *objective* norms come to apply to the essentially social statuses—paradigmatically the doxastic and practical propositionally contentful commitments that correspond to beliefs and intentions—and so underwrite such fundamental practices as assessing the *truth* of beliefs and the *success* of actions. The next most important part of the second half of the book is Chapter 6: the substitutional analysis is crucial to understanding how the inferential approach can generalize beyond sentences. And it would be truly a shame to miss the transcendental expressive argument for the existence of objects—the argument that (and why) the *only* form the world we talk and think of can take is that of a world of facts about particular objects and their properties and relations. It is worth keeping in mind Diderot's thought that one "must have gone deep into art or science to master their elements . . . The darkness of the beginnings lights up only toward the middle or the end."

The aim throughout is to present a unified vision of language and mind—one that starts with a relatively clear philosophical rationale and works it out in convincing detail, addressing a sufficiently wide range of potentially puzzling phenomena to engender confidence in its adaptability and power. It is animated by the ideal of the systematic philosophers of old: the invigorating clarifying prospect achievable by laying alongside our ordinary ways of talking and thinking an alternate idiom in which everything can be said. I am sensible, of course, of many ways in which this product falls short of that ideal. Particularly in matters of detail (but by no means there alone), a myriad of choices have had to be made at the cost of spurning attractive, perhaps ultimately superior, alternatives. The approach seldom dictates just one way of doing things. Yet the choice of which large limb to follow off the trunk of the tradition must be made on the basis of the tempting fruit to be seen on the smaller branches it supports. It can only be hoped that where upon closer inspection some of them are found wanting, the fundamental soundness of the tree is not impugned, but only the judgment of the gardener, who pruned

the better and nurtured the worse. As Johnson says in the Preface to his *Dictionary*: "A large work is difficult because it is large, even though all its parts might singly be performed with facility; where there are many things to be done, each must be allowed its share of time and labor in the proportion only which it bears to the whole; nor can it be expected that the stones which form the dome of the temple should be squared and polished like the diamond of a ring." Truth may or may not be in the whole, but understanding surely is.

I am grateful to the copyright holders for permission to reprint the following excerpts:

From Donald Davidson, *Actions and Events* (Oxford: Oxford University Press, 1980), by permission of Oxford University Press.

From Michael Dummett, *Frege: Philosophy of Language* (Cambridge, Mass.: Harvard University Press, copyright © 1991 by Michael Dummett), reprinted by permission of Harvard University Press and Gerald Duckworth & Co. Ltd.

From T. S. Eliot, "Burnt Norton" and "Little Gidding" in *Four Quartets*, copyright © 1943 by T. S. Eliot and renewed 1971 by Esme Valerie Eliot, reprinted by permission of Harcourt Brace & Company and Faber and Faber Ltd.

From Gottlob Frege, *The Foundations of Arithmetic*, trans. J. L. Austin (Evanston, Ill.: Northwestern University Press, 1978), by permission of the publisher.

From H. Hermes, F. Kambartel, and F. Kaulbach, eds., *Gottlob Frege: Posthumous Writings* (Chicago: University of Chicago Press, 1979), by permission of the University of Chicago Press and Blackwell Publishers.

From Wilfrid Sellars, "Inference and Meaning," *Mind* 62, no. 247 (1953), by permission of Oxford University Press.

From Wallace Stevens, "On the Road Home," *Collected Poems*, copyright 1942 by Wallace Stevens and renewed 1970 by Holly Stevens. Reprinted by permission of Alfred A. Knopf, Inc., and Faber and Faber Ltd.

Some of the material presented here has appeared elsewhere, in article form. In particular, "Inference, Expression, and Induction" (*Philosophical Studies* 54 [1988]: 257–285, reprinted by permission of Kluwer Academic Publishers) is essentially incorporated as Sections III–V of Chapter 2; "Asserting" (*Nous* 17, no. 4 [November 1983]: 637–650) is an early ancestor of the core of Chapter 3; and "Pragmatism, Phenomenalism, and Truth Talk" (*Midwest Studies in Philosophy* 12, "Realism and Antirealism" [1988]: 75–93) and "Reference Explained Away" (*The Journal of Philosophy* 81, no. 9 [September 1984]: 469–492) coincide extensively with the middle sections of Chapter 5.

During the long gestation of this work I have been helped immensely by discussions of these ideas with many other philosophers, who gave generously of their time and acumen. The confusions and unclarities they have talked me through and out of are legion; those that remain belong on my account alone. It is impossible properly to acknowledge all these contributions, but some require particular mention. To begin with, I want to thank those with and to whom I most proximally say 'we': my colleagues in the Philosophy Department at the University of Pittsburgh, especially John McDowell, John Haugeland, and Nuel Belnap. The example, encouragement, and indulgence of this nearly ideal intellectual community made this book possible. I want also to express my appreciation for the exceptionally talented group of graduate students at Pitt who have shared this intellectual adventure over the years, particularly Mark Lance, Marc Lange, Danielle Macbeth, and Katarzyna Paprzycka. Finally, I am immeasurably grateful to Allan Gibbard for the gargantuan efforts he expended in mastering the ideas presented here, and for his suggestions as to how the final product might be improved.



---

## **PART ONE**

---



# Toward a Normative Pragmatics

---

An ounce of practice is worth a pound of precept.

ENGLISH PROVERB

---

## I. INTRODUCTION

### 1. *Saying 'We'*

'We' is said in many ways. We may be thee and me. We may be all that talks or all that moves, all that minds or all that matters. Since these boundaries are elastic, we have a task of demarcation: telling who or what we are, distinguishing ourselves from the other sorts of objects or organisms we find in our world. Saying who we are can contract to an empty exercise in self-congratulation—a ritual rehearsal of the endless, pitiable disabilities of clockworks, carrots, cows, and the clan across the river. Such a mean-spirited version of the demarcational enterprise is not forced on us by the way things are, however.

For what we are is made as much as found, decided as well as discovered. The sort of thing we are depends, in part, on what we take ourselves to be. One characteristic way we develop and make ourselves into what we are is by expressing, exploring, and clarifying our understanding of what we are. Arbitrary distinctions of biology, geography, culture, or preference can be and have been seized on to enforce and make intelligible the crucial distinction between us and them (or it). But philosophical thought is coeval with the impulse to understand ourselves according to a more principled, less parochial story—and so to be a more principled, less parochial sort of being.

The wider perspective enjoined by principle poses the question, Who are we? in the form: What would have to be true—not only of the quaint folk across the river, but of chimpanzees, dolphins, gaseous extraterrestrials, or digital computers (things in many ways quite different from the rest of us)—for them nonetheless to be correctly counted among us? Putting the issue this way acknowledges an expansive demarcational commitment to avoid, as far as possible, requiring the sharing of adventitious stigmata of origin or material constitution. In understanding ourselves we should look to conditions at once more abstract and more practical, which concern what we are able to *do*, rather than where we come from or what we are made of. Candidates for recognition as belonging among us should be required to share only the fundamental abilities that make possible participation in those central activities by which we (thereby) define ourselves. How should we think of these?

The most cosmopolitan approach begins with a pluralistic insight. When we ask, Who are we? or What sort of thing are we? the answers can vary without competing. Each one defines a different way of saying ‘we’; each kind of ‘we’-saying defines a different community, and we find ourselves in many communities. This thought suggests that we think of ourselves in broadest terms as the ones who say ‘we’. It points to the one great Community comprising members of all particular communities—the Community of those who say ‘we’ with and to someone, whether the members of those different particular communities recognize each other or not.

The reflexive character of the proposal that we use self-demarcation as the criterion by which we demarcate ourselves does not suffice to render it purely formal, however. It does not save us the trouble of contentful self-understanding. For until it has been specified in other terms what one must be able to do in order to count as “saying ‘we’,” demarcation by appeal to such attitudes remains an aspiration tacked to a slogan—empty, waiting for us to fill it. ‘We’-saying of the sort that might be of demarcational interest is not a matter merely of the production of certain vocables—indeed perhaps the relevant kind of attitude is not a linguistic matter at all. Nor again does it consist simply in the engendering of warm mammalian fellow-feeling. Making explicit to ourselves who we are requires a theoretical account of what it is in practice to treat another as one of us.

## 2. *Sapience*

What is it we do that is so special? The answer to be explored here—a traditional one, to be sure—is that we are distinguished by capacities that are broadly cognitive. Our transactions with other things, and with each other, in a special and characteristic sense *mean* something to us, they have a *conceptual content* for us, we *understand* them in one way rather than another. It is this demarcational strategy that underlies the classical iden-

tification of us as *reasonable* beings. Reason is as nothing to the beasts of the field. We are the ones on whom reasons are binding, who are subject to the peculiar force of the better reason.

This force is a species of *normative* force, a rational ‘ought’. Being rational is being bound or constrained by these norms, being subject to the authority of reasons. Saying ‘we’ in this sense is placing ourselves and each other in the space of reasons, by giving and asking for reasons for our attitudes and performances. Adopting this sort of practical stance is taking or treating ourselves as subjects of cognition and action; for attitudes we adopt in response to environing stimuli count as *beliefs* just insofar as they can serve as and stand in need of reasons, and the acts we perform count as *actions* just insofar as it is proper to offer and inquire after reasons for them. Our attitudes and acts exhibit an intelligible content, a content that can be grasped or understood, by being caught up in a web of reasons, by being inferentially articulated. Understanding in this favored sense is a grasp of reasons, mastery of proprieties of theoretical and practical *inference*. To identify ourselves as *rational*—as the ones who live and move and have our being in the space of reasons, and so to whom things can be intelligible—is to seize demarcationally on a capacity that might well be shared by beings quite different from us in provenance and demeanor.

Picking us out by our capacity for reason and understanding expresses a commitment to take *sapience*, rather than *sentience* as the constellation of characteristics that distinguishes us. Sentience is what we share with non-verbal animals such as cats—the capacity to be *aware* in the sense of being *awake*. Sentience, which so far as our understanding yet reaches is an exclusively biological phenomenon, is in turn to be distinguished from the mere reliable differential responsiveness we sentients share with artifacts such as thermostats and land mines. Sapience concerns understanding or intelligence, rather than irritability or arousal. One is treating something as sapient insofar as one explains its behavior by attributing to it intentional states such as belief and desire as constituting reasons for that behavior.

Another familiar route to understanding the sort of sapience being considered here for demarcational duty goes through the concept of *truth*, rather than that of *inference*. We are believers, and believing is taking-true. We are agents, and acting is making-true. To be sapient is to have states such as belief, desire, and intention, which are contentful in the sense that the question can appropriately be raised under what circumstances what is believed, desired, or intended would be *true*. Understanding such a content is grasping the conditions necessary and sufficient for its truth.

These two ways of conceiving sapience, in terms of inference and in terms of truth, have as their common explanatory target contents distinguished by their *propositional* form. What we can offer as a reason, what we can take or make true, has a propositional content: a content of the sort that we express by the use of declarative sentences and ascribe by the use of ‘that’ clauses.

Propositional contents stand in inferential relations, and they have truth conditions. One of the tasks of this work is to explain what it is to grasp specifically propositional contents, and so to explain who we are as rational or sapient beings. A central subsidiary task is accordingly to offer an account of the relation between the concepts of inference and truth, which complement one another and in some measure compete with one another for explanatory priority in addressing the issue of propositional contentfulness, and so of rationality.

### 3. *Intentionality*

The general self-understanding in view so far identifies us by our broadly cognitive capacities: We are makers and takers of reasons, seekers and speakers of truth. The propositional focus of the approach marks this understanding of intelligible contents as *discursive*. This conception, hallowed by ancient tradition, was challenged during the Enlightenment by a rival approach to cognitive contentfulness that centers on the concept of *representation*. Descartes's seminal demarcational story distinguishes us as representers—producers and consumers of representings—from a world of merely represented and representable things. The states and acts characteristic of us are in a special sense *of*, *about*, or *directed at* things. They are representings, which is to say that they have representative content. To have such a content is to be liable to assessments of *correctness* of representation, which is a special way of being *answerable* or *responsible* to what is represented.

Another task of this work is accordingly to address the question, How should the relation between representation—the master concept of Enlightenment epistemology—and the discursive concepts of reasons and truth be understood? One of the great strengths representationalist explanatory strategies have developed is the capacity to offer accounts of truth and goodness of inference. There are familiar set-theoretic routes that set out from representational primitives corresponding to subsentential linguistic expressions such as singular terms and predicates, lead to assignments of truth conditions to sentences compounded out of those expressions, and pass from there to determinations of which inferences are correct. While doubts have been raised, perhaps legitimately, about nearly every phase of this construction, no other semantic approach has been worked out so well.

Yet for all that, the primitives involved have never been well understood. Descartes notoriously fails to offer an account either of the nature of representational contents—or of what the representingness of representings consists in—or of what it is to grasp or understand such contents, that is to say, of their intelligibility to the representer. He does not tell us what makes a rabbit-idea an idea of (or purporting to be of) rabbits, or of anything at all, nor what it is for the one whose idea it is to understand or take it *as* being of or about something. That things could be represented by and to the mind (have

"objective reality" in it and for it, for the mind to be "as if of" things) is treated as a basic property, an unexplained explainer. But an adequate treatment of the representational dimension of discursive sapience should include an account both of representational *purport*, and of its *uptake*.

The topic to be investigated here, then, is intentionality in the sense of the propositional contentfulness of attitudes, not in the sense (if that should turn out to be different) of the directedness of sense. The aim is to understand ourselves as judges and agents, as concept-users who can reason both theoretically and practically. This is not to say that we should understand ourselves exclusively as sapients rather than sentients, in terms of understanding rather than awareness. 'We' is and by rights ought to be said in many ways. The point is just to register and delineate the way that is to be discussed here.

This inquiry is directed at the fanciest sort of intentionality, one that involves expressive capacities that cannot be made sense of apart from participation in linguistic practices. The aim is to offer sufficient conditions for a system of social practices to count as specifically *linguistic* practices, in the sense of defining an idiom that confers recognizably *propositional* contents on expressions, performances, and attitudes suitably caught up in those practices. Looking at this sort of high-grade intentionality accordingly risks being beastly to the beasts—not only by emphasizing sapience over sentience, comprehension over consciousness, but also by unfairly ignoring the sorts of beliefs and desires that are appropriately attributed to non- or pre-linguistic animals.

So it is a further criterion of adequacy of this explanatory enterprise that it have something to say about the lower grades of intentionality: not only as to how the lines should be drawn (corresponding to different senses of 'we'), but also as to how the advent of the favored sort of linguistic intentionality can be made less mysterious. How can linguistic abilities arise out of nonlinguistic ones? Or to ask a related question, What would sentient creatures have to be able to do in order to count as sapient as well? What is needed is to tell a story about practices that are sufficient to confer propositionally contentful intentional states on those who engage in them, without presupposing such states on the part of the practitioners. The hope is that doing so will offer guidance concerning what would be involved in diagnosing aliens as exhibiting such states, and programming computers or teaching merely sentient animals to exhibit them.

## II. FROM INTENTIONAL STATE TO NORMATIVE STATUS

### 1. Kant: *Demarcation by Norms*

The demarcational proposal being pursued picks us out as the ones capable of judgment and action. Not only do we respond differentially to environing stimuli, we respond by forming perceptual judgments. Not

only do we produce behavior, we perform actions. Various ways of talking about this fundamental distinction have been put on the table. It can be made out in terms of *truth*. In perception what we do is responsively *take-true* some propositional content that is intelligible to us. In action what we do is responsively *make-true* some propositional content that is intelligible to us.

The distinction can be made out in terms of *reasons*. The judgments that are our perceptual responses to what is going on around us differ from responses that are not propositionally contentful (and so are not in that sense intelligible) in that they can serve as reasons, as premises from which further conclusions can be drawn. Actions, which alter what is going on around us in response to propositionally contentful intentions, differ from performances that are merely behavior (and so not intelligible in terms of the propositionally contentful intentions that elicit them) in that reasons can be given for them; they can appear as the conclusions of practical inferences.

The distinction can also be made out in terms of the employment of *concepts*. To be a perceiver rather than just an irritable organism is to be disposed to respond reliably and differentially to the perceptible environment by the application of appropriate concepts. To be an agent rather than just a behaver is to be disposed to respond reliably and differentially to applications of appropriate concepts by altering the accessible environment. Intelligibility in the sense of propositional contentfulness, whether the latter is conceived in terms of truth conditions or capacity to serve as a reason, is a matter of conceptual articulation—in the case of perception and action, that the reliably elicited response and the reliably eliciting stimulus, respectively, essentially involve the use of concepts.

So sapience, discursive intentionality, is concept-mongering. What is distinctive of specifically *conceptual* activity? Contemporary thought about the use of concepts owes great debts to Kant. One of his cardinal innovations is his introduction of the idea that conceptually structured activity is distinguished by its *normative* character. His fundamental insight is that judgments and actions are to be understood to begin with in terms of the special way in which we are *responsible* for them.

Kant understands concepts as having the form of *rules*, which is to say that they specify how something *ought* (according to the rule) to be done.<sup>1</sup> The understanding, the conceptual faculty, is the faculty of grasping rules—of appreciating the distinction between correct and incorrect application they determine. What is distinctive about judgings and doings—acts that have contents that one can take or make true and for which the demand for reasons is in order—is the way they are governed by rules. They are conceptually contentful and so are subject to evaluation according to the rules that express those contents. Being in an intentional state or performing an intentional action accordingly has a normative significance. It counts as undertaking (acquiring) an obligation or commitment; the content of the commitment is determined by the rules that are the concepts in terms of which the act or

state is articulated. Thus Kant's version of the sort of demarcation criterion being considered picks us out as distinctively *normative*, or rule-governed, creatures.

## 2. From Cartesian Certainty to Kantian Necessity

This emphasis on the normative significance of attributions of intentionally contentful states marks a decisive difference between Kantian and Cartesian ways of conceiving cognition and action. For Kant the important line is not that separating the mental and the material as two matter-of-factually different kinds of stuff. It is rather that separating what is subject to certain kinds of normative assessment and what is not. For Descartes, having a mind (grasping intentional contents) is having representings: states that purport or seem to represent something. Some things in the world exhibit this sort of property; others do not. Where Descartes puts forward a *descriptive* conception of intentionality, Kant puts forward a normative, or *prescriptive*, one—what matters is being the subject not of *properties* of a certain kind but of *proprieties* of a certain kind. The key to the conceptual is to be found not by investigating a special sort of mental *substance* that must be manipulated in applying concepts but by investigating the special sort of *authority* one becomes subject to in applying concepts—the way in which conceptually articulated acts are liable to assessments of correctness and incorrectness according to the concepts they involve.

This approach contrasts sharply with Cartesian demarcations of cognition and action according to the presence of items of a certain matter-of-factual kind. The objection is not to the details of Descartes's understanding of the descriptive features required for intentionally contentful states and acts: his conception of mental events that are self-intimating cognitions or infallibly performable volitions, takings-true and makings-true that are minimal in that they cannot fail to be successful. It is, more radically, that what sets off the intentional is its liability to assessments of correctness, its being subject to norms (which are understood as codified in rules), rather than any missing feature that it could be described as having or lacking.

Descartes inaugurated a new philosophical era by conceiving of what he took to be the *ontological* distinction between the mental and the physical in *epistemological* terms: in terms of accessibility to cognition—in terms, ultimately, of certainty. Kant launched a new philosophical epoch by shifting the center of concern from *certainty* to *necessity*. Where Descartes's descriptive conception of intentionality, centering on certainty, picks out as essential our grip on the concepts employed in cognition and action, Kant's normative conception of intentionality, centering on necessity, treats their grip on us as the heart of the matter. The attempt to understand the source, nature, and significance of the norms implicit in our concepts—both those that govern the theoretical employment of concepts in inquiry and knowl-

edge and those that govern their practical employment in deliberation and action—stands at the very center of Kant's philosophical enterprise. The most urgent question for Kant is how to understand the rulishness of concepts, how to understand their authority, bindingness, or validity. It is this normative character that he calls *Notwendigkeit* (necessity).

The nature and significance of the sea change from Cartesian certainty to Kantian necessity will be misunderstood unless it is kept in mind that by 'necessary' Kant means 'in accord with a rule'. It is in this sense that he is entitled to talk about the *natural* necessity whose recognition is implicit in cognitive or theoretical activity, and the *moral* necessity whose recognition is implicit in practical activity, as species of one genus. The key concept of each is obligation by a rule. It is tempting, but misleading, to understand Kant's use of the notion of necessity anachronistically, in terms of contemporary discussions of alethic modality. It is misleading because Kant's concerns are at base normative, in the sense that the fundamental categories are those of deontic modality, of commitment and entitlement, rather than of alethic modality, of necessity and possibility as those terms are used today. Kant's commitment to the primacy of the practical consists in seeing both theoretical and practical consciousness, cognitive and conative activity, in these ultimately normative terms.

So for Kant, concepts are to be understood by the theorist in terms of the *rules* that make them explicit, rules that specify how the concepts are *properly* or *correctly* applied and otherwise employed. Kant's appreciation of this normative significance of concept use is one of the lenses through which he views his relationship to his rationalist and empiricist predecessors. With the wisdom of hindsight, Kant can see a normative strand of concern with *responsibility* as fundamental to the Enlightenment.<sup>2</sup> Thus the *Meditations* is to be read as motivated by the demand that the meditator take personal responsibility for every claim officially endorsed—be prepared to answer for it, demonstrate entitlement to that commitment by justifying it. This theme remained merely implicit in Descartes's theorizing about us (as opposed to his motivation and methodology), for his explicit theory remains naturalistic (though not, of course, physicalistic).

Leibniz insists, against the empiricists, that inferential transitions between representations ought not to be assimilated to matter-of-factual, habitually acquired causal dispositions. He understands them rather as applications of general principles that must accordingly be available prior to any knowledge of empirical matters of fact. Kant takes over from his reading of Leibniz the general idea of rules as what underwrite cognitive assessments of inferences and judgments. He understands such a priori principles, however, not as very general statements of fact (even metaphysical fact), but as rules of reasoning. They are conceived not as descriptive but as prescriptive—as (in Sellars's phrase) "fraught with ought."

This lesson dovetails neatly with the moral he draws from Hume's

thought. On Kant's reading, Hume's contribution is to see that ordinary empirical discourse involves commitments that reach beyond the sequences of representations, however regular, in which the concepts deployed in that discourse are taken to originate. Kant's Hume recognizes that cognitive experience crucially involves the application and assessment of the correctness of the application of *rules*. For Kant, Hume's inquiry after the nature of the authority for this inferential extension takes the form of a quest for the nature of the necessity, understood as normative bindingness, exhibited by the rules implicit in empirical concepts. It is under this conception that Kant can assimilate Hume's point about the distinction between saying what *happens* (describing a regularity) and saying what is causally *necessary* (prescribing a rule) to his point about the distinction between saying what *is* and saying what *ought* to be. One need not buy the metaphysics that Kant uses to ground and explain his norms, nor accept his answer to Hume, in order to appreciate the transformation of perspective made possible by his emphasis on the normativeness of the conceptual, and hence of cognition and action—the latter distinguished in the first instance as what we are responsible for.

### 3. Frege: *Justification versus Causation*

Kant's lesson is taken over as a central theme by Frege, whose campaign against psychologism relies on respecting and enforcing the distinction between the normative significance of applying concepts and the causal consequences of doing so. For Frege, it is possible to investigate in a naturalistic way acts of judging or thinking (even thinking conceived in a dualistic way), but such an investigation inevitably overlooks the normative dimension that is essential to understanding the propositional contents that are judged or thought. Sometimes this point is put in terms of *reasons*, invoking inferential relations among judgeable contents, as when he complains that psychologism "loses the distinction between the grounds that justify a conviction and the causes that actually produce it,"<sup>3</sup> or again when he argues that "the laws in accordance with which we actually draw inferences are not to be identified with the laws of correct [*richtigen*] inference; otherwise we could never draw a wrong inference."<sup>4</sup> Sometimes the point is put in terms of *truth*, as when he says, "It is not the holding something to be true that concerns us, but the laws of truth. We can also think of these as prescriptions for making judgments; we must comply with them if our judgments are not to fail of the truth."<sup>5</sup> Put either way, the point is that concern with the propositional contents that are thought or judged is inseparable from the possibility of assessments of *correctness*. Besides empirical regularities, there are also proprieties governing inferring and holding-true. Besides questions of which judgeable contents *are* held true, and under what circumstances, there is the question of which ones *ought* to be, and when. Besides the question of what consequences holding-true or making a judgment with

a certain content in fact leads to, there is the question of what those consequences ought or must rationally be. Psychology can study the matter-of-factual *properties* of contentful acts of judging and inferring, but not the semantically determined *proprieties* that govern them, the norms according to which assessments of truth and rationality are to be made.

Psychologism misunderstands the pragmatic significance of semantic contents. It cannot make intelligible the applicability of norms governing the acts that exhibit them. The force of those acts is a prescriptive rather than a descriptive affair; apart from their liability to assessments of judgments as true and inferences as correct, there is no such thing as judgment or inference. To try to analyze the conceptual contents of judgments in terms of habits or dispositions governing sequences of brain states or mentalistically conceived ideas is to settle on the wrong sort of modality, on causal necessitation rather than rational or cognitive right. Such natural processes "are no more true than false; they are simply processes, as an eddy in the water is a process. And if we are to speak of a right, it can only be the right of things to happen as they do happen. One phantasm contradicts another no more than one eddy in water contradicts another."<sup>6</sup> Contradiction, correct inference, correct judgment are all *normative* notions, not *natural* ones.

The laws of nature do not forbid the making of contradictory judgments. Such judgments are forbidden in a normative sense. It is *incorrect* to endorse incompatible contents: rationally incorrect, incorrect according to rules of reason, prescriptions governing what is *proper* in the way of inferring and judging. The 'must' of justification or good inference is not the 'must' of causal compulsion. But the possibility of expressing each in terms of rules or laws, so central to Kant's enterprise, misleads if these two different sorts of laws are not kept distinct, as they are not by psychologism and associationism. "What makes us so prone to embrace such erroneous views is that we define the task of logic as the investigation of the laws of thought, whilst understanding by this expression something on the same footing as the laws of nature . . . So if we call them laws of thought, or, better, laws of judgment, we must not forget we are concerned here with laws which, like the principles of morals or the laws of the state, prescribe how we are to act, and do not, like the laws of nature, define the actual course of events."<sup>7</sup>

Frege expresses his views about the normative character of judgeable contents, which he understands as having truth conditions, and so about the application of concepts, which he understands as functions whose values are truth-values, by talking about the nature of logic, which he understands as the study of the laws of truth.

Logic, like ethics, can also be called a normative science.<sup>8</sup>

The property 'good' has a significance for the latter analogous to that which the property 'true' has for the former. Although our actions and

endeavours are all causally conditioned and explicable in psychological terms, they do not all deserve to be called good.<sup>9</sup> Discussion of just how these remarks about the normative or prescriptive character of logic relate to a commitment to the normative or prescriptive significance of the exhibition of conceptual content by judgments must await more detailed consideration of Frege's *Begriffsschrift* theory of logical vocabulary as expressive of conceptual contents, in Chapter 2.

#### 4. Wittgenstein on the Normative Significance of Intentional Content

Frege emphasizes that concern with the contents of concepts and judgments is inseparable from concern with the possibility of the concepts being correctly or incorrectly applied, the judgments correctly or incorrectly made, whether this correctness is conceived in terms of truth or of the goodness of inference. Understanding this point requires distinguishing normative from causal modalities. Beyond enforcing this distinction, however, Frege has little to say about the nature of the norms that matter for the study of conceptual contents, and so for logic—though of course he has a great deal to say about the structure of such contents (some of which will be rehearsed in subsequent chapters). His concerns are at base semantic rather than pragmatic. In the twentieth century, the great proponent of the thesis that *intentionally* contentful states and acts have an essentially *normative* pragmatic significance is the later Wittgenstein.

The starting point of his investigations is the insight that our ordinary understanding of states and acts of meaning, understanding, intending, or believing something is an understanding of them as states and acts that *commit* or *oblige* us to act and think in various ways. To perform its traditional role, the meaning of a linguistic expression must determine how it would be *correct* to use it in various contexts. To understand or grasp such a meaning is to be able to distinguish correct from incorrect uses. The view is not restricted to meaning and understanding but extends as well to such intentionally contentful states as believing and intending. This is one way of developing and extending Kant's point that to take what we do as judging and acting is to treat it as subject to certain kinds of assessments as to its correctness: truth (corresponding to the world) and success (corresponding to the intention). A particular belief may actually relate in various ways to how things are, but its content determines how it is *appropriate* for it to be related, according to the belief—namely that the content of the taking-true should be true. A particular intention may or may not settle how one *will* act, but its content determines how it is *appropriate* to act, according to the intention—namely by making-true that content. To say this is in no way to deny that occurrences of intentional states of meaning, understanding, intending, and believing have causal significances. It is simply to point out that

understanding them as contentful involves understanding them as also having normative significances.<sup>10</sup>

The issue constantly before us in Wittgenstein's later works is how to understand these normative significances of intentional contents—the way in which they incorporate standards for assessments of correctness. Many of his most characteristic lines of thought are explorations of the inaptness of thinking of the normative 'force', which determines how it would be appropriate to act, on the model of a special kind of causal 'force'. The sense in which understanding or grasping a meaning is the source of the *correct* use is quite different from the sense in which it is the source of what one in fact goes on to do.<sup>11</sup> Enforcing the Kantian and Fregean distinction between grounds in the order of justification and grounds in the order of causation is what is behind talk of the "hardness of the logical 'must'"<sup>12</sup> and the picture of the dominion or compulsion intentional states exercise over what counts as correct performance as a machine whose "super-rigid" construction precludes any sort of malfunction. "The machine as symbolizing its action . . . We talk as if these parts could only move in this way, as if they could not do anything else. How is this—do we forget the possibility of their bending, breaking off, melting, and so on? Yes; in many cases we don't think of that at all . . . And it is quite true: the movement of the machine-as-symbol is predetermined in a different sense from that in which the movement of any given actual machine is predetermined."<sup>13</sup>

The relation between the content of an intention and the performances that would fulfill that intention does not leave any room for misfire, corresponding to the melting or bending of the parts of a mechanism, for it is already a *normative* relation. The state is to settle what *ought* to be done, what *must* be done if it is to be realized. What actually does or would happen is another matter. The images of superrigidity—of being guided by rails that one cannot fall away from—are what one gets if one assimilates normative compulsion to causal compulsion, ignoring the Kantian distinction. That is, if the normative 'must' were a kind of causal 'must', it would have to be a puzzling, superrigid sort—but the point is not to start with this sort of naturalistic prejudice.

In fact, by contrast, "The laws of inference can be said to compel us; in the same sense, that is to say, as other laws in human society."<sup>14</sup> They determine, in a sense yet to be specified, what one *ought* to do. Being compelled in this sense is entirely compatible with failing to act as one 'must'. Indeed, the physical or causal possibility of making a mistake, or doing what one is obliged, by what one means, intends, believes, and desires, not to do, is essential to the conception of such states and shows the essentially normative nature of their significance. "But I don't mean that what I do now (in grasping a sense) determines the future use causally and as a matter of experience, but that in a queer way, the use itself is already present".—But of course it is, 'in some sense'! Really the only thing wrong with what you say

is the expression ‘in a queer way’. The rest is all right; and the sentence only seems queer when one imagines a different language game for it from the one in which we actually use it”<sup>15</sup>—the different game, namely, of attributing *natural states* and properties, rather than *normative statuses* such as commitments. What is determined is not how one *will* act but how one *ought* to, given the sense or content grasped, or the rule one has endorsed. “How am I able to obey a rule?”—if this is not a question about causes, then it is about the justification for my following the rule in the way I do.”<sup>16</sup> That is, it is a question about what actions accord with the rule, are obliged or permitted by it, rather than with what my grasp of it actually makes me do.

### 5. Norms and Intentional Explanation

Although Wittgenstein often uses specifically linguistic examples, and some commentators have focused exclusively on these cases, the normative phenomena he highlights are part and parcel of intentional attribution generally, whether or not language is in the picture. *Ceteris paribus*, one who believes that it is raining, and that moving under the tree is the only way to stay dry, and who desires to stay dry, *ought* to move under the tree. The intentional states make the action appropriate. Indeed, the concept of rationality achieves its paradigmatic application in just such circumstances, as conduct warranted by the attributed intentional states is characterized as *rationally appropriate*. The qualification marks the option being reserved to deny that the conduct is, say, morally, politically, or aesthetically appropriate. (It is a further question whether this explanatory role of rationality justifies conceiving of what is rationally appropriate as reducible to what is prudentially or instrumentally appropriate.) Taking the category of rationality to be essentially involved in intentional explanation, as Dennett and Davidson for instance do, is one way of recognizing the normative dimension of intentionality.

It is important to keep this acknowledgment distinct from further theses one may then want to endorse concerning that normative dimension. If one keeps one’s eye resolutely on the causal dimension of intentional explanation, the normative aspect can be masked. For instance, Dennett conjoins his recognition of the constitutive role of rationality in intentional explanation with the claim that such explanation involves a substantive “rationality assumption,” the assumption in effect that the system will by and large in fact act as it ought rationally to act. There is nothing wrong with considering explanations of this sort as intentional explanations, but it is important to distinguish normative intentional explanation from causal intentional explanation. The former explains only what the subject of the intentional states ought or is obliged or committed (rationally) to do in virtue of its exhibition of the attributed states. The latter makes the substantive rationality assumption and goes on to explain what in fact happens. Normative intentional

explanations are more fundamental; they are presupposed and built upon by causal ones.

The same normative considerations arise if one approaches intentional states from the direction of their functional roles in mediating perception and action. Where Wittgenstein, Sellars, and Dummett, for instance, look at meaning, understanding, believing, or intending something in terms of mastery of the public proprieties governing the use of linguistic expressions, others would see intentional states as already definable by their role in accounting for the conduct of rational agents, whether linguistically adept or not. Both views are functionalist, in a broad sense. They differ over how to draw the boundaries around the functional systems within which alone something can have the significance of an intentional state. The issue of the extent to which mastery of linguistic social practice is a prerequisite for possession of intentional states of various sorts is of course an important one. But one need not have settled on one or the other of these explanatory approaches in order to appreciate that intentional states belong on the normative side of the Kantian divide. For talk of functional roles is itself already normative talk.

Specifying the functional role of some state in a system is specifying how it *ought* to behave and interact with other states. It is with reference to such a role that one makes sense of the notion of a *malfuctioning* component, something that is not behaving as it is supposed to. One may go on to offer various stories about the source of the correctnesses involved in functional roles: invoking the intentions of the designer, the purposes of the user, or the way the system must function if it is to realize the evolutionary good of survival, or even the cognitive good of accurately representing its environment. The point is that all of these are accounts of the source of the norms of proper functioning that are an integral part of functional explanations. The job of a designer's drawing of a machine is to specify how the machine is supposed to work, how it ought to work, according to the intentions of the designer. It is for this reason that "we forget the possibility of [the pieces] bending, breaking off, melting, and so on." Wittgenstein is of course concerned to understand how it is possible to understand such normative roles. But the current point is just that the roles one seeks to specify, in explaining the significance of intentional states, must, to begin with, be understood in normative terms of proper or correct functioning. Once again, this is not to deny that the fact that some component or system ought (functionally) to behave in a certain way may under many circumstances have a causal significance regarding how it will in fact behave. The issues are in principle distinct, however, and causal functional accounts presuppose normative functional ones.

The recognition that the consequences of attributing intentionally contentful states must be specified in normative terms may be summed up in the slogan, "Attributing an intentional state is attributing a normative

status." This is one of the leading ideas to be pursued in the present investigation. Intentional states and acts have contents in virtue of which they are essentially liable to evaluations of the "force of the better reason." It is this mysterious "force"—evidently the core of the social practices of giving and asking for reasons—that Greek philosophy investigated and appealed to in demarcating us from the nonrational background of items that we can think and find out about but that cannot themselves think or find out about other things. This "force of the better reason" is a normative force. It concerns what further beliefs one is *committed* to acknowledge, what one *ought* to conclude, what one is *committed* or *entitled* to say or do. Talk of what is a reason for what has to do in the first instance not with how people do or *would* act but with how they *should* act, what they should acknowledge. The sophist may not in fact respond to this "force," but even the sophist *ought* to. To understand rationality and states whose contents are articulated according to their role in reasoning, one must understand the force of such 'ought's. The relevance of reasons to the attributing and undertaking of intentional states and acts is *prima facie* reason to employ a normative metalanguage in analyzing such activity.

The normative dimension of intentional attributions is equally apparent if the propositional contents of the states and acts that are attributed, exhibited, or performed are conceived, not in terms of their accessibility to *reasons*, but in terms of there being circumstances under which they would be *true*. Assessments of truth, no less than assessments of rationality, are normative assessments. Truth and rationality are both forms of correctness. To ask whether a belief is true is to ask whether it is in some sense proper, just as to ask whether there are good reasons for it is to ask whether it is proper in a different sense. The business of truth talk is to evaluate the extent to which a state or act has fulfilled a certain kind of responsibility. This normative aspect of concern with truth can be masked by offering a descriptive, matter-of-factual account of what truth consists in. But doing so should be understood as offering a theory about this variety of semantic correctness, not as a denial that correctness is what is at issue. Thus Dummett argues that one does not understand the concept of truth when one has only a method for determining when it correctly applies to a claim or belief—a practical mastery of its circumstances of application. One must also know the *point* of applying it, must understand that truth is the *proper goal* of assertion and belief, that the language game of assertion and belief implicitly but essentially involves the injunction that one *ought* to speak and believe the truth. That is what one is supposed to be trying to do. Without an appreciation of this normative significance of application of the concept *truth*, one does not understand that concept.

Raising the question of what a belief or claim *represents* or is *about* can be understood as treating it as in a special way answerable for its *correctness* to what is represented, what it is about.<sup>17</sup> Thus the claim that semantically

or intentionally contentful states and acts have, as such, pragmatic significances that should be specified in normative terms does not depend upon what particular model (for instance, reasons, truth conditions, or representation) is employed in understanding such contents. The theoretical task of the intentional content of a state or act is to determine, in context, the normative significance of acquiring that state or performing that act: when it is appropriate or correct to do so and what the appropriate consequences of doing so are. The content is to determine proprieties of use, employment, or performance for states, acts, and expressions that exhibit or express such contents. The content must (in context) settle when it is *correct* to apply a concept in judging, believing, or claiming, and what correctly follows from such an application. Correctnesses of application are discussed under the general headings of assessments of truth or representation; correctnesses of inference are discussed under the general heading of assessments of rationality.<sup>18</sup> To pick out intentional states and acts as ones to which any of these sorts of assessments—truth, accuracy of representation, or reasonability—are in principle appropriate is to treat their normative articulation as essential to them. For this point, it does not matter which sort of assessment is treated as fundamental, whether the goodness of claiming of the sort concepts of truth try to capture, the goodness of representation that concepts of correspondence try to capture, or the goodness of reasoning of the sort concepts of rationality try to capture. All are *prima facie* normative or evaluative notions.

### III. FROM NORMS EXPLICIT IN RULES TO NORMS IMPLICIT IN PRACTICES

#### 1. *Regulism: Norms as Explicit Rules or Principles*

The first commitment being attributed to Wittgenstein, then, is to taking the significance of attributing intentional states to be normative, a matter of the difference it makes to the correctness or justification of possible performances (including the adoption of other intentional states). The second commitment he undertakes concerns how to understand the normative status-uses of correct and incorrect, justified and not justified, which this approach to intentionality concentrates on. The question of how the normative significances of intentional states are to be taken to be related to the matter-of-factual consequences of those states, which would be one way into this issue, can be put to one side for the moment. It is a question Wittgenstein is much interested in, but it ought to be seen as arising at a different point in the argument. For an account of the normative pole of the Kantian dualism need not take the form of a specification of how the normative is related to the nonnormative. Instead, Wittgenstein considers, and rejects, a particular model of correctness and incorrectness, roughly Kant's, in which what makes

a performance correct or not is its relation to some explicit rule. To understand his argument and the lesson he draws from it, it is necessary to see what this model of the normative is, and for what sort of explanatory role he claims it is unsuitable.

According to this more specific Kantian view,<sup>19</sup> norms just *are* rules of conduct. Normative assessments of performances are understood as always having the form of assessments of the extent to which those performances accord with some rule. Reference to proprieties of performance is taken as indirect reference to rules, which determine what is proper by explicitly *saying* what is proper. On this account, acts are liable to normative assessments insofar as they are governed by propositionally *explicit* prescriptions, prohibitions, and permissions. These may be conceived as rules, or alternatively as principles, laws, commands, contracts, or conventions. Each of these determines what one may or must do by *saying* what one may or must do. For a performance to be correct is, on this model, for the rules to permit or require it, for it to be in accord with principle, for the law to allow or demand it, for it to be commanded or contracted. It is because Kant is someone for whom the normative always appears in the explicit form of rules, laws, and commandments that he could see the rationalists' insistence on the essential role of principles in cognition and action as a dark appreciation of the fundamentally normative character of those faculties. It is for this reason that when Kant wants to say that we are creatures distinguished from others by the normative dimension of our conduct (both cognitive and practical), he puts this in terms of our being bound by *rules*.

On an approach according to which normative assessment of conduct—whether prospectively, in deliberation, or retrospectively, in appraisal—always begins with the question of what rule is followed in producing the performances in question, norms are likened to laws in the sense of statutes. For conduct is legally appropriate or inappropriate just insofar as it is governed by some explicit law that says it is. Assessments of legal praise and blame must at least implicitly appeal to the relation of the performance in question to some law. In this way, the model appeals to a familiar institutional context, in which the norms most in evidence clearly take the form of explicit principles, commands, and the like.

The influence of the jurisprudential analogy is evident in Kant's conception of the normative aspect of cognition and action in terms of following rules. Kant inherits the Enlightenment tradition, handed down from Grotius and Pufendorf, which first studied the normative in the form of positive and natural laws, conceived as the explicit commandments of sovereigns or superiors of one sort or another. As a result, Kant takes it for granted that it is appropriate to call a 'rule' or a 'law' whatever it is that determines the propriety or impropriety of some judgment or performance. For him, as for most philosophers before this century, explicit rules and principles are not simply one form among others that the normative might assume. Rules are

the form of the norm as such. This view, that proprieties of *practice* are always and everywhere to be conceived as expressions of the bindingness of underlying *principles*, may be called *regulism* about norms.<sup>20</sup>

According to this intellectualist, platonist conception of norms, common to Kant and Frege, to assess correctness is always to make at least implicit reference to a rule or principle that *determines* what is correct by explicitly *saying so*. In the best-known portion of his discussion of rule-following in the *Philosophical Investigations*, Wittgenstein argues that proprieties of performance that are governed by explicit rules do not form an autonomous stratum of normative statuses, one that could exist though no other did. Rather, proprieties governed by explicit rules rest on proprieties governed by practice. Norms that are *explicit* in the form of rules presuppose norms *implicit* in practices.

## 2. Wittgenstein's Regress Argument

Norms explicit as rules presuppose norms implicit in practices because a rule specifying how something is correctly done (how a word ought to be used, how a piano ought to be tuned) must be applied to particular circumstances, and applying a rule in particular circumstances is itself essentially something that can be done correctly or incorrectly. A rule, principle or command has normative significance for performances only in the context of practices determining how it is correctly applied. For any particular performance and any rule, there will be ways of applying the rule so as to forbid the performance, and ways of applying it so as to permit or require it. The rule determines proprieties of performance only when correctly applied.

If correctnesses of performance are determined by rules only against the background of correctnesses of application of the rule, how are these latter correctnesses to be understood? If the regulist understanding of all norms as rules is right, then applications of a rule should themselves be understood as correct insofar as they accord with some further rule. Only if this is so can the rule-conception play the explanatory role of being the model for understanding *all* norms. A rule for applying a rule Wittgenstein calls an "interpretation" (*Deutung*). "There is an inclination to say: every action according to the rule is an interpretation. But we ought to restrict the term 'interpretation' to the substitution of one expression of the rule for another."<sup>21</sup> The question of the autonomy of the intellectualist conception of norms, presupposed by the claim that rules are the form of the normative, is the question of whether the normative can be understood as "rules all the way down," or whether rulish proprieties depend on some more primitive sort of *practical* propriety. Wittgenstein argues that the latter is the case. Rules do not apply themselves; they determine correctnesses of performance only in the context of practices of distinguishing correct from incorrect applications of the rules. To conceive these practical proprieties of application as themselves rule-governed is to embark on a regress. Sooner or later the theorist will have to

acknowledge the existence of practical distinctions between what is appropriate and what not, admitting appropriatenesses according to practice as well as according to rules or explicit principles.

This regress argument shows that the platonist conception of norms as rules is not an autonomous one, and so does not describe the fundamental form of norm. "What does a game look like that is everywhere bounded by rules? whose rules never let a doubt creep in, but stop up all the cracks where it might?—Can't we imagine a rule determining the application of a rule, and a doubt which it removes—and so on?"<sup>22</sup> In each case the doubt is the possibility of a mistake, of going wrong, of acting incorrectly, for instance in applying a rule. The point is to be that a rule can remove such a doubt, settle what is correct to do, only insofar as it is itself correctly applied. "'But how can a rule show me what I have to do at this point? Whatever I do is, on some interpretation, in accord with the rule.'—That is not what we ought to say, but rather: any interpretation [*Deutung*] still hangs in the air along with what it interprets [*dem Gedeuteten*], and cannot give it any support. Interpretations by themselves do not determine meaning."<sup>23</sup> No sequence of interpretations can eliminate the need to apply the final rules, and this is always itself subject to normative assessment. Applied incorrectly, any interpretation misleads. The rule says how to do one thing correctly only on the assumption that one can do something else correctly, namely apply the rule.

This was our paradox: no course of action could be determined by a rule, because every course of action can be made out to accord with the rule. The answer was: if everything can be made out to accord with the rule, then it can also be made out to conflict with it. And so there would be neither accord nor conflict here. It can be seen that there is a misunderstanding here from the mere fact that in the course of our argument we give one interpretation after another; as if each one contented us for at least a moment, until we thought of yet another standing behind it. What this shows is that there is a way of grasping a rule [*eine Auffassung einer Regel*] which is not an interpretation, but which is exhibited in what we call "obeying the rule" and "going against it" in actual cases.<sup>24</sup>

Absent such a practical way of grasping norms, no sense can be made of the distinction between correct and incorrect performance—of the difference between acting according to the norm and acting against it. Norms would then be unintelligible.

### 3. Wittgenstein's Pragmatism about Norms

The conclusion of the regress argument is that there is a need for a *pragmatist* conception of norms—a notion of primitive correctnesses of performance *implicit in practice* that precede and are presupposed by their *explicit* formulation in *rules* and *principles*. "To use the word without a

justification does not mean to use it wrongfully [*zu Unrecht gebrauchen*].<sup>25</sup> There is a kind of correctness that does not depend on explicit justifications, a kind of correctness of practice.

And hence also ‘obeying a rule’ is a practice [*Praxis*.]<sup>26</sup>

—To obey a rule, to make a report, to give an order, to play a game of chess, are customs [*Gepflogenheiten*] (uses, institutions).<sup>27</sup>

The regress argument does not by itself provide such a conception of properties of practice; it just shows that without one we cannot understand how rules can codify the correctnesses that they do.

This argument shares its form with the regress Lewis Carroll invokes in “The Tortoise and Achilles”<sup>28</sup> but takes that line of thought one level deeper. That story depends on the fact that in a formal logical system, statements are inferentially inert. Even conditionals, whose expressive job it is to make inferential relations explicit as the contents of claims, license inferential transitions from premises to conclusions only in the context of rules permitting detachment. Rules are needed to give claims, even conditional claims, a normative significance for action. Rules specify how conditionals are to be used—how it would be correct to use them. It is the rules that fix the inference-licensing role of conditionals, and so their significance for what it is correct to do (infer, assert). Although particular rules can be traded in for axioms (in the form of conditional claims), one cannot in principle trade in *all* rules for axioms. So one cannot express all of the rules that govern inferences in a logical system in the form of propositionally explicit postulates within that system.

Carroll uses the regress of conditionals that results from the attempt to replace the rule of conditional detachment by explicitly postulated conditionals as an argument to show this. Wittgenstein’s regress-of-rules argument shows further that, while rules can codify the pragmatic normative significance of claims, they do so only against a background of practices permitting the distinguishing of correct from incorrect applications of those rules. Carroll’s point is that the significance of claims for what it is correct to do must somehow be secured. Logical claims, like others, must have some normative pragmatic significance. Wittgenstein’s point is then that conceiving such significances in regulist terms, as the invocation of rules of inference does, is not the whole story. Rule-based proprieties of performance depend on practice-based ones. The regulist, platonist, intellectualist conception of norms must be supplemented by that of the pragmatist.

Two commitments have now been attributed to Wittgenstein. The first is a *normative* thesis about the *pragmatics* of intentionality. The second is a *pragmatic* thesis about the *normativeness* of intentionality. In the first case, pragmatics is distinguished from semantics, as the theory of the significance of contentful states and performances from the theory of their contents. In

the second case, pragmatic theories of norms are distinguished from platonist theories, in treating as fundamental norms *implicit in practices* rather than norms *explicit in principles*. The first point enforces attention to the significance of intentional states for what it is *correct* to do. The second point is that proprieties of practice must be conceivable antecedently to their being expressly formulated into propositionally explicit governing rules or principles. For performances can be rule-governed only insofar as they are governed as well by practices of applying rules.

It is useful to approach the sort of understanding that is involved in mastering a practice, for instance a practice of applying or assessing applications of a rule, by means of Ryle's distinction between knowing *how* and knowing *that*.<sup>29</sup> Knowing how to do something is a matter of practical ability. To know how is just to be reliably able. Thus one knows how to ride a bicycle, apply a concept, draw an inference, and so on just in case one can discriminate in one's practice, in the performances one produces and assesses, between correct and incorrect ways of doing these things.

The explicit knowing-that corresponding to such implicit knowing-how is a theoretical formulation or expression of that practical ability, in a rule or principle, that *says* what is correct and what not. The intellectualist picture underwrites every bit of know-how with a bit of knowledge-that, which may be only implicit in practical discriminations. "Compare knowing and saying: how many feet high Mont Blanc is—how the word 'game' is used—how a clarinet sounds. If you are surprised that one can know something and not be able to say it, you are perhaps thinking of a case like the first. Certainly not like the third."<sup>30</sup> What Wittgenstein shows is that the intellectualist model will not do as an account of the nature of the normative as such. For when applied to the norms governing the application of rules and principles, it generates a regress, which can be halted only by acknowledging the existence of some more primitive form of norm. The regress is Wittgenstein's master argument for the appropriateness of the pragmatist, rather than the regulist-intellectualist, order of explanation.<sup>31</sup>

#### 4. Sellars against Regulism

Another thinker who, like Wittgenstein, takes his starting point from Kant's and Frege's appreciation of the normative character of intentionality (for him, coeval with language use) is Wilfrid Sellars. He takes up this theme in one of his earliest papers, published in 1947. The opening section of that paper is entitled "Behavior, Norm, and Semantic Meta-Language" and makes this point:

The psychologicistic blunder with respect to 'means' is related to another fundamental error, that, namely, of confusing between (1) language as a descriptive category for which symbols are empirical classes to which

certain events belong (and hence are symbol-events) by virtue of performing an empirical function, with (2) language as an epistemological category for which the relation of type to token is not that of empirical class to member. . .

For the moment it will help clarify the epistemological distinction between symbol-types and symbol-tokens, if we think of the former as norms or standards, and of the latter as events which satisfy them. We can therefore, for the moment at least, contrast the above two senses of ‘language’ as the descriptive and the normative respectively. Making use of this distinction, we argue that ‘meaning’ or, better, ‘designation’ is a term belonging to language about languages in the second sense. Its primary employment is therefore in connection with linguistic expressions as norms.<sup>32</sup>

Like Wittgenstein, Sellars sees that an adequate conception of these norms must move beyond the pervasive regulist tradition, which can understand them only in the form of rules.

Focusing on linguistic intentionality, Sellars in another paper examines the regulist conception as it applies to the linguistic norms in virtue of which it is possible to say anything at all. “It seems plausible to say that a language is a system of expressions, the use of which is subject to certain rules. It would seem, thus, that learning to use a language is learning to obey the rules for the use of its expressions. However, taken as it stands, this thesis is subject to an obvious and devastating objection.”<sup>33</sup> The objection is that taking ‘correct’ to mean ‘correct according to a rule’ generates a familiar sort of regress:

The refutation runs as follows: Thesis. Learning to use a language (*L*) is learning to obey the rules of *L*. But, a rule which enjoins the doing of an action (*A*) is a sentence in a language which contains an expression for *A*. Hence, a rule which enjoins the using of a linguistic expression (*E*) is a sentence in a language which contains an expression for *E*—in other words, a sentence in a metalanguage. Consequently, learning to obey the rules for *L* presupposes the ability to use the metalanguage (*ML*) in which the rules for *L* are formulated. So that, learning to use a language (*L*) presupposes having learned to use a metalanguage (*ML*). And by the same token, having learned to use *ML* presupposes having learned to use a metametalanguage (*MML*) and so on. But, this is impossible (a vicious regress). Therefore, the thesis is absurd and must be rejected.<sup>34</sup>

The metalanguage expresses rules for the proper application of concepts of the object language. But these rules, too, must be applied. So the

metametalanguage expresses rules for applying the rules of the metalanguage, and so on.

If any talk is to be possible, there must be some meta . . . metalevel at which one has an understanding of rules that does not consist in offering another interpretation of them (according to rules formulated in a metalanguage) but which consists in being able to distinguish correct applications of the rule in practice. The question is how to understand such practical normative know-how. Although he, like Wittgenstein, uses 'rule' more broadly than is here recommended, Sellars is clearly after such a notion of norms implicit in practice: "We saw that a rule, properly speaking, isn't a rule unless it lives in behavior, rule-regulated behavior, even rule-violating behavior. Linguistically we always operate within a framework of living rules. (The snake which sheds one skin lives within another.) In attempting to grasp rules as rules from without, we are trying to have our cake and eat it. To describe rules is to describe the skeletons of rules. A rule is lived, not described."<sup>35</sup>

This line of thought, common to Wittgenstein and Sellars, raises the key question of how to understand proprieties of practice, without appealing to rules, interpretations, justifications, or other explicit claims that something is appropriate. What does the practical capacity or 'know-how' to distinguish correct from incorrect performances (for instance—but this is only one example—applications of a rule) consist in? This is to ask what it is to take or treat a performance as correct-according-to-a-practice. It should also be asked, What is it for an act to be correct-according-to-a-practice? Both questions are important ones to ask: In what sense can norms (proprieties, correctnesses) be implicit in a practice? and What is it for someone to acknowledge those implicit norms as governing or being binding on a range of performers or performances?

The answers to these questions may be more intimately related to one another than at first appears. To foreshadow: On the broadly phenomenalist line about norms that will be defended here, norms are in an important sense in the eye of the beholder, so that one cannot address the question of what implicit norms are, independently of the question of what it is to acknowledge them in practice. The direction of explanation to be pursued here first offers an account of the practical attitude of *taking* something to be correct-according-to-a-practice, and then explains the status of *being* correct-according-to-a-practice by appeal to those attitudes. Filling in a story about normative attitudes as assessments of normative status, and explaining how such attitudes are related both to those statuses and to what is actually done, will count as specifying a sense of "norms implicit in practice" just insofar as the result satisfies the criteria of adequacy imposed on the notion of practice by the regress-of-rules argument.

Another central explanatory criterion of adequacy for such a conception of implicit practical normative knowing-how is that it be possible in terms

of it to understand explicit knowing-*that*. The effect is to reverse the regulist-intellectualist order of explanation. The regulist starts with a notion of norms explicit in principles and is obliged then to develop an account of what it would be for such things to be implicit in practices. The pragmatist starts rather with a notion of norms implicit in practice and is obliged then to develop an account of what it would be for such things to become propositionally explicit, as claims or rules. One of the primary tasks of this book is accordingly to offer an account of what it is to take some propriety that is implicit in a practice and make it explicit in the form of a claim, principle, or rule.

### 5. Regularism: Norms as Regularities

The regress-of-rules or regress-of-interpretations argument common to Wittgenstein and Sellars sets up criteria of adequacy for an account of contentful states that acknowledges their essentially normative significance, their characteristic relevance to assessments of the correctness of acts (including the adoption of further states). It must be possible to make sense of a notion of norms implicit in practice—which participants in the practice are bound by, and can acknowledge being bound by—without appeal to any explicit rules or capacities on the part of those participants to understand and apply such rules. Since the regress arises when the rule-following model of being bound by norms is applied to the agent, one strategy for avoiding it is to shift to a different model. Perhaps rules are relevant only as describing regularities, and not as being followed in achieving them.

Sellars (who does not endorse it) introduces such an approach this way: "Now, at first sight there is a simple and straightforward way of preserving the essential claim of the thesis while freeing it from the refutation. It consists in substituting the phrase 'learning to conform to the rules . . .' for 'learning to obey the rules . . .' where 'conforming to a rule enjoining the doing of *A* in circumstances *C*' is to be equated simply with 'doing *A* when the circumstances are *C*'—regardless of how one comes to do it . . . A person who has the habit of doing *A* in *C* would then be conforming to the above rule even though the idea that he was to do *A* in *C* never occurred to him, and even though he had no language for referring to either *A* or *C*."<sup>36</sup> What generates the regress is the demand that each practical capacity to act appropriately be analyzed as following an explicit rule that says what is appropriate, since understanding what is said by such a rule turns out to involve further practical mastery of proprieties.

If the practices in which norms are implicit are understood simply as regularities of performance, then there is nothing the practitioner need already understand. If such regularities of performance can be treated as practices governed by implicit norms, then there will be no regress or circularity in appealing to them as part of an account of knowing-*that*, of expressing

norms explicitly in rules and principles. For the only one who needs to understand how to apply correctly the rule conforming to which makes performances count as regular is the theorist who describes the regularity in terms of that rule. The norms implicit in regularities of conduct can be expressed explicitly in rules, but need not be so expressible by those in whose regular conduct they are implicit.

The view that to talk about implicit norms is just to talk about regularities—that practices should be understood just as regularities of behavior—may be called the simple regularity theory. It is clear how such a *regularist* account of the normative avoids the regress that threatens *regulist* accounts. The proposal is to identify being correct according to (norms implicit in) practice—in the sense required to avoid the regress of rules as interpretations that plagues fully platonist accounts—with conforming to (norms explicit in) a rule, where ‘conforming to a rule’ is just producing performances that are regular in that they count (for us) as correct according to it. The immediate difficulty with such a proposal is that it threatens to obliterate the contrast between treating a performance as subject to normative assessment of some sort and treating it as subject to physical laws.

For this reason simple regularity theories seem to abandon the idea that the significance of contentful states is to be conceived in normative terms. No one doubts that actions and linguistic performances are subject to laws of the latter sort and so conform to rules or are regular. The thesis of the normative significance of intentional states sought to distinguish intentional states from states whose significance is merely causal, and that distinction seems to be taken back by the simple regularity account. After all, as Kant tells us, in this sense “everything in nature, in the inanimate as well as the animate world, happens according to rules . . . All nature is actually nothing but a nexus of appearances according to rules; and there is nothing without rules.”<sup>37</sup> Everything acts regularly, according to the laws of physics. In what special sense do intentional states then involve specifically normative significances?

For a regularist account to weather this challenge, it must be able to fund a distinction between what is in fact done and what ought to be done. It must make room for the permanent possibility of mistakes, for what is done or taken to be correct nonetheless to turn out to be incorrect or inappropriate, according to some rule or practice. The importance of this possibility to the genuinely normative character of the force or significance associated with contentful states is a central and striking theme in Wittgenstein’s later works. What is correct or appropriate, what is obligatory or permitted, what one is committed or entitled to do—these are normative matters. Without the distinction between what is done and what ought to be done, this insight is lost.

The simple regularity approach is committed to identifying the distinction between *correct* and *incorrect* performance with that between *regular* and

irregular performance. A norm implicit in a practice is just a pattern exhibited by behavior. To violate that norm, to make a mistake or act incorrectly according to that norm, is to break the pattern, to act irregularly. The progress promised by such a regularity account of proprieties of practice lies in the possibility of specifying the pattern or regularity in purely descriptive terms and then allowing the relation between regular and irregular performance to stand in for the normative distinction between what is correct and what is not. Wittgenstein explicitly considers and rejects this approach. Where his master argument against regulism has the form of an appeal to the regress of interpretations, his master argument against regularism has the form of an appeal to the possibility of gerrymandering.

The problem is that any particular set of performances exhibits many regularities. These will agree on the performances that have been produced and differ in their treatment of some possible performances that have not (yet) been produced. A performance can be denominated ‘irregular’ only with respect to a specified regularity, not *tout court*. Any further performance will count as regular with respect to some of the patterns exhibited by the original set and as irregular with respect to others. For anything one might go on to do, there is some regularity with respect to which it counts as “going on in the same way,” continuing the previous pattern. Kripke has powerfully expounded the battery of arguments and examples that Wittgenstein brings to bear to establish the point in this connection.<sup>38</sup> There simply is no such thing as *the* pattern or regularity exhibited by a stretch of past behavior, which can be appealed to in judging some candidate bit of future behavior as regular or irregular, and hence, on this line, as correct or incorrect. For the simple regularist’s identification of impropriety with irregularity to get a grip, it must be supplemented with some way of picking out, as somehow *privileged*, some out of all the regularities exhibited. To say this is to say that some regularities must be picked out as the ones that *ought* to be conformed to, some patterns as the ones that *ought* to be continued. The simple regularity view offers no suggestions as to how this might be done and therefore does not solve, but merely puts off, the question of how to understand the normative distinction between what is done and what ought to be done.

One might respond to the demand that there be some way to pick out the correct regularity, from all the descriptively adequate but incompatible candidates, by shifting what one describes, from finite sets of performances to the set of performances (for instance, applications of a concept) the individual is *disposed* to produce. This set is infinite, in that any bearer of an intentional state is disposed to respond, say by applying or refusing to apply the concept *red* or *prime*, in an infinite number of slightly different circumstances. Kripke argues that this appeal to dispositions nevertheless does not suffice to rule out regularities that agree in all the cases one has dispositions with respect to, and differ in others so remote (perhaps, in the case of *prime*, because the numbers involved are so large, and in the case of *red* because surrounding

circumstances are so peculiar) that one does not have dispositions to treat them one way rather than another.

This last argument is controversial, but it is not a controversy that need be entered into here; however it may be with the finiteness objection to a dispositional account of the regularities that, according to the line of thought being considered, are to play the role of norms implicit in practice, there is another more serious objection to it. No one ever acts incorrectly in the sense of violating his or her own dispositions. Indeed, to talk of 'violating' dispositions is illicitly to import normative vocabulary into a purely descriptive context. Understanding the norms implicit in practice as descriptively adequate rules codifying regularities of disposition (even if a unique set of such rules is forthcoming) loses the contrast between correct and mistaken performance that is of the essence of the sort of normative assessment being reconstructed. If whatever one is disposed to do counts for that reason as right, then the distinction of right and wrong, and so all normative force, has been lost. Thus the simple regularity view cannot be rescued from the gerrymandering objection by appealing to dispositions in order to single out or privilege a unique regularity.

The problem that Wittgenstein sets up, then, is to make sense of a notion of norms implicit in practice that will not lose either the notion of implicitness, as regulism does, or the notion of norms, as simple regularism does. McDowell puts the point nicely: "Wittgenstein's problem is to steer a course between a Scylla and a Charybdis. Scylla is the idea that understanding is always interpretation. We can avoid Scylla by stressing that, say, calling something 'green' can be like crying 'Help' when one is drowning—simply how one has learned to react to this situation. But then we risk steering on to Charybdis—the picture of a level at which there are no norms . . . How can a performance be nothing but a 'blind' reaction to a situation, not an attempt to act on interpretation (thus avoiding Scylla), and be a case of going by a rule (avoiding Charybdis)? The answer is: by belonging to a custom (*PI* 198), practice (*PI* 202), or institution (*RFM VI-31*)."<sup>39</sup> The Scylla of regulism is shown to be unacceptable by the regress-of-rules argument. The Charybdis of regularism is shown to be unacceptable by the gerrymandering-of-regularities argument.

If anything is to be made of the Kantian insight that there is a fundamental normative dimension to the application of concepts (and hence to the significance of discursive or propositionally contentful intentional states and performances), an account is needed of what it is for norms to be implicit in practices. Such practices must be construed both as not having to involve explicit rules and as distinct from mere regularities. Wittgenstein, the principled theoretical quietist, does not attempt to provide a theory of practices, nor would he endorse the project of doing so. The last thing he thinks we need is more philosophical theories. Nonetheless, one of the projects pursued in the rest of this work is to come up with an account of norms implicit in

practices that will satisfy the criteria of adequacy Wittgenstein's arguments have established.

#### IV. FROM NORMATIVE STATUS TO NORMATIVE ATTITUDE

##### 1. Kant: *Acting According to Conceptions of Rules*

Two theses have so far been attributed to Kant. First, the sort of intentionality characteristic of us, exhibited on the theoretical side in judgment and on the practical side in action, has an essential *normative* dimension. Second, norms are to be understood as having the form of explicit *rules*, or principles. The first of these has been endorsed, as expressing a fundamental insight. The second has been rejected, on the basis of Wittgenstein's argument from the regress of rules as interpretations of rules. The conclusion drawn was that norms that are explicitly expressed in the form of rules, which determine what is correct according to them by *saying* or describing what is correct, must be understood as only one form that norms can take. That form is intelligible only against a background that includes norms that are *implicit* in what is *done*, rather than *explicit* in what is *said*.<sup>40</sup> At least the norms involved in properly understanding what is said by rules, or indeed in properly understanding any explicit saying or thinking, must be construed as norms of practice, on pain of a vicious regress.

In Kant's account of us as normative creatures, however, these two theses are inseparably bound up with a third. As has already been pointed out, Kant takes it that everything in nature happens according to rules. Being subject to rules is not special to us discursive, that is concept-applying, subjects of judgment and action.<sup>41</sup> What is distinctive about us as normative creatures is the way in which we are subject to norms (for Kant, in the form of rules). As natural beings, we act according to rules. As rational beings, we act according to our *conceptions* of rules.<sup>42</sup> It is not being bound by necessity, acting according to rules, that sets us apart; it is being bound not just by natural but by rational necessity. Kant's whole practical philosophy, and in particular the second Critique, is devoted to offering an account of this distinction between two ways in which one can be related to rules. Most of the details of his way of working out this idea are special to the systematic philosophical setting he develops and inhabits and need not be rehearsed here. Two fundamental features of his idea, however, must be taken seriously by any attempt to pursue his point about the normative character of concept-users.

The first of these has already been remarked on in connection with Frege. It concerns the distinction between the causal modalities and the more properly normative 'ought's whose applicability to us is being considered as a criterion of demarcation. This is the phenomenon distinguishing the force of causal 'must's from the force of logical or rational 'must's that Wittgen-

stein invokes in connection with his discussion of misunderstandings of the 'hardness' of the latter in relation to the former. It is an essential feature of the sort of government by norms that Kant is pointing to that it is compatible with the possibility of *mistakes*, of those subject to the norms going *wrong*, failing to do what they are obliged by those norms to do, or doing what they are *not* entitled to do. The 'ought' involved in saying that a stone subject to no other forces *ought* to accelerate toward the center of the earth at a rate of 32 feet per second per second shows itself to have the force of an attribution of natural or causal necessity by entailing that the stone *will* so act. The claim that it in this sense ought to behave a certain way is incompatible with the claim that it does not do so. In contrast, no such entailment or incompatibility is involved in claims about how we intentional agents ought to behave, for instance what else one of us is committed to believe or to do by having beliefs and desires with particular contents. Leaving room for the possibility of mistakes and failures in this way is one of the essential distinguishing features of the 'ought's that express government by norms in the sense that is being taken as characteristic of us, as opposed to it. The sense in which we are *compelled* by the norms that matter for intentionality, norms dictating what we are under various circumstances obliged to believe and to do, is quite different from natural compulsion.

The second feature of Kant's idea addresses precisely the nature of this normative compulsion that is nevertheless compatible with recalcitrance. For he does not just distinguish the sense in which we are bound by these norms from the sense in which we are bound by natural necessity in the purely formal terms invoked by this familiar point about the possibility of our going wrong. He characterizes it substantively as acting according to a *conception* or a *representation* of a rule, rather than just according to a rule. Shorn of the details of his story about the nature of representations and the way they can affect what we do, the point he is making is that we act according to our *grasp* or *understanding* of rules. The rules do not immediately compel us, as natural ones do. Their compulsion is rather mediated by our *attitude* toward those rules. What makes us act as we do is not the rule or norm itself but our *acknowledgment* of it. It is the possibility of this intervening attitude that is missing in the relation between merely natural objects and the rules that govern them. The slippage possible in our acting according to our conception of a rule is made intelligible by distinguishing the sense in which one is bound by a rule whose grip on us depends on our recognition or acknowledgment of it as binding from the sense in which one can be bound by a rule whose grip does not depend on its being acknowledged. This explanatory strategy might be compared to Descartes's invocation of intervening representations in explaining the possibility of error about external things—though Kant need not be understood as following Descartes's path from an implicit appeal to the regress that threatens such representationalist pictures of cognition to a diagnosis of the relation be-

tween the subject and those mediating, error-enabling internal representations as itself immediate and hence immune to the possibility of error.

The idea underlying the demarcational strategy Kant introduces when he defines us, denizens of the realm of freedom, as beings that are capable of acting according to a conception of a rule by contrast to the denizens of the realm of nature, is that natural beings, who merely act according to rules, that is, regularly, are capable of acknowledging norms *only* by obedience to them. We rational beings are also capable of grasping or understanding the norms, of making assessments of correctness and incorrectness according to them. Those assessments play a role in determining what we go on to do—the phenomenon Kant denominates “the rational will.” But for us, in contrast to merely natural creatures, the assessment of the propriety of a performance is one thing, and the performance itself is another. The possibility of not doing what we nevertheless count as bound or obliged to do arises out of this distinction. What is special about us is the sort of grasp or uptake of normative significance that we are capable of. To be one of us in this sense of ‘us’ is to be the subject of normative *attitudes*, to be capable of acknowledging proprieties and improprieties of conduct, to be able to treat a performance as correct or incorrect.

## 2. Practical Normative Attitudes

It is a challenge to retain this insight about the significance of our normative attitudes while accommodating Wittgenstein’s pragmatist point about norms (and so jettisoning the intellectualist insistence on the explicitness of norms that colors Kant’s treatment). In order to do so, it must be possible to distinguish the attitude of acknowledging *implicitly* or *in practice* the correctness of some class of performances from merely exhibiting regularities of performance by producing only those that fall within that class. Otherwise, inanimate objects will count as acknowledging the correctness of laws of physics, and the distinction Kant points out is lost. As before, the challenge is to reject intellectualist regulism about norms without falling into nonnormative regularism.

Consideration of this third thesis of Kant’s sharpens the point, however, by focusing it on the capacity to adopt a normative practical attitude—to act in such a way as to attribute a normative significance, without doing so by *saying* that that is what one is doing. The question now becomes, What must one be able to *do* in order to count as *taking* or *treating* a performance as correct or incorrect? What is it for such a normative attitude—attributing a normative significance or status to a performance—to be implicit in practice? The importance of this question is a direct consequence of Kant’s point, once his rendering has been deintellectualized by replacing grasp of principles with mastery of practices.

It is a consequence of this criterion of adequacy that the practical performances that are assessments cannot be just the same performances that are assessed. Addressing the simplest case first, treating a performance as correct cannot be identified with producing it. For according to such an identification, the only way in which a norm can be acknowledged in practice is by obeying it, acting regularly according to it. But then it is impossible to treat performances as incorrect: the norms one counts as acknowledging simply are whatever regularities describe what one does. Such an account would collapse Kant's distinction between the way in which we are governed by norms we acknowledge and the way in which we are governed by natural laws independently of our acknowledgment of them.

Kant's principle that we are the ones who act not only according to rules but according to a conception of them is the claim that we are not merely *subject* to norms but *sensitive* to them. This principle has been taken over here by saying that we are characterized not only by normative *statuses*, but by normative *attitudes*—which is to say not only that our performances are correct or incorrect according to various rules but also that we can in our practice *treat* them as correct or incorrect according to various rules. Using 'assessment' to mean an assignment of normative significance—in the most basic case taking as correct or incorrect—the point may be put by saying that Kant's principle focuses demarcational interest on the normative *attitudes* exhibited by the activity of *assessing*, rather than just on the normative *statuses* being *assessed*. In order to respect the lessons of Wittgenstein's pragmatism about the normative, assessing must be understood as something *done*; the normative attitude must be construed as somehow implicit in the practice of the assessor, rather than explicit as the endorsement of a proposition. Construed in these practical terms, a consequence of Kant's distinction is that mere conformity to a norm is not even a candidate as a construal of the normative attitude of assessing conformity, which expresses the sort of sensitivity to norms that characterizes us.

For brutes or bits of the inanimate world to qualify as engaging in practices that implicitly acknowledge the applicability of norms, they would have to exhibit behavior that counts as treating conduct (their own or that of others) as correct or incorrect. Of course such things do respond differentially to their own and each other's antics. The question is what role such a response must play in order to deserve to be called a practical taking or treating of some performance as correct or incorrect, perhaps in the way in which eating something deserves to be called a practical taking or treating of it as food. Any sort of reliable differential responsive disposition can be understood as inducing a classification of stimuli. Iron rusts in some environments and not others, and so can be interpreted as classifying its environments into two sorts, depending on which kind of response they tend to elicit. Such responsive classification is a primitive kind of practical taking of something as

something. It is in this sense that an animal's eating something can be interpreted as its thereby *taking* what it eats as food. The issue of current concern is what must be true of a behavioral response-kind for it to be *correct* or *appropriate* that something is taken as by being responded to in that way, rather than *food* or *wet*.

### 3. Sanctions

In part because much of the tradition of thought about normative status and attitudes has taken its departure from a legal model, it is natural to answer this question by invoking the notion of *sanctions*: of reward and punishment. According to such a *retributive* approach to assessment, one treats a performance as correct or appropriate by *rewarding* it, and as incorrect or inappropriate by *punishing* it. Such an account can take many forms, depending on how sanctions are construed. In the simplest case, applying a negative sanction might be understood in terms of corporal punishment; a prelinguistic community could express its practical grasp of a norm of conduct by beating with sticks any of its members who are perceived as transgressing that norm. In these terms it is possible to explain for instance what it is for there to be a practical norm in force according to which in order to be entitled to enter a particular hut one is obliged to display a leaf from a certain sort of tree. The communal response of beating anyone who attempts to enter without such a token gives leaves of the proper kind the normative significance, for the community members, of a license. In this way members of the community can show, by what they do, what they *take* to be appropriate and inappropriate conduct.

One example of this approach is Haugeland's account of practical norms in terms of social constellations of dispositions having a structure he calls "conformist." He asks us to imagine under this heading creatures who not only conform their behavior to that of other community members in the sense of imitating each other, and so tending to act alike (normally in the sense of typically) in similar circumstances, but also sanction each other's performances, making future behavior more likely to conform to ("cluster around") the emergent standards. "The clusters that coalesce can be called 'norms' (and not just groups or types) precisely because they are generated and maintained by censoriousness; the censure attendant on deviation automatically gives the standards (the extant clusters) a *de facto* normative force."<sup>43</sup> According to such an account the normative attitudes of taking or treating some performance as correct or incorrect are understood in terms of behavioral *reinforcement*, in the learning-theoretic sense. The advantage of such a way of putting things is that reinforcement is a purely functional descriptive notion, definable in abstraction from the particular considerations about familiar animals, in virtue of which beating them with a stick is

likely to function as negative reinforcement. Treating a performance as correct is taken to be positively sanctioning it, which is to say positively reinforcing it. Positively reinforcing a disposition to produce a performance of a certain kind as a response to a stimulus of a certain kind is responding to the response in such a way as to make it more likely in the future that a response of that kind will be elicited by a stimulus of the corresponding kind. Treating a performance as incorrect is taken to be negatively sanctioning it, which is to say negatively reinforcing it. Negatively reinforcing a disposition to produce a performance of a certain kind as a response to a stimulus of a certain kind is responding to the response in such a way as to make it less likely in the future that a response of that kind will be elicited by a stimulus of the corresponding kind.

The approach being considered distinguishes us as norm-governed creatures from merely regular natural creatures by the normative *attitudes* we evince—attitudes that express our grasp or practical conception of our behavior as governed by norms. These normative attitudes are understood in turn as *assessments*, assignments to performances of normative significance or status as correct or incorrect according to some norm. The assessing attitudes are then understood as dispositions to *sanction*, positively or negatively. Finally, sanctioning is understood in terms of *reinforcement*, which is a matter of the actual effect of the sanctioning or reinforcing responses on the responsive dispositions of the one whose performances are being reinforced, that is sanctioned, that is assessed.

Such a story is a kind of regularity theory, but not a simple regularity theory. It does not identify a norm wherever there is a regularity of behavior. In keeping with Kant's insight (as transposed from an intellectualist to a pragmatist key), norms are discerned only where attitudes—acknowledgments in practice of the bindingness of those norms—play a mediating role in regularities. Only insofar as regularities are brought about and sustained by effective assessments of propriety, in the form of responsive classifications of performances as correct or incorrect, are regularities taken to have specifically normative force. The possibility of incorrect, inappropriate, or mistaken performances—those that do not accord with the norm—is explicitly allowed for. Thus there is no danger of this sense of 'obligatory' collapsing into the sort of causal modality that governs merely natural happenings. Acknowledging a norm as in force is distinguishable in these terms from simply obeying it. A cardinal advantage of these theories is that while to this extent countenancing Kant's distinction between genuinely norm-governed and merely regular activity, they make intelligible how conduct that deserves to be called distinctively norm-governed could arise in the natural world.

The fundamental strategy pursued by such a theory is a promising one. As here elaborated, it involves three distinguishable commitments. First, Kant's distinction between acting according to a rule and acting according to a

conception of a rule is taken to express an important insight about the special way in which we are normative creatures. Second, the pragmatist regress-of-rules argument is taken to show that in order to make use of this insight, it is necessary that the sort of normative attitude that Kant takes to play an essential mediating role in our government by norms be understood as involving implicit acknowledgment of norms in practice. Specifically, it is necessary to make sense of the idea of practically taking or treating performances as correct or incorrect. Third, taking or treating performances as correct or incorrect, approving or disapproving them in practice, is explained in terms of positive and negative sanctions, rewards and punishments. This tripartite strategy is endorsed and pursued in the rest of this work. There are reasons not to be happy with the regularist way of working it out that has just been sketched, however.

Even the version of a regularity theory that Haugeland presents, which appeals only to patterns of positive and negative reinforcement to fund the notion of sanctions and thereby that of practical normative attitudes, merely puts off the issue of gerrymandering. Just as there is no such thing as *the* regularity of performance evinced by some actual course of conduct—because if there is one way of specifying it, there are an infinite number of distinguishable variants that overlap or agree about the specified performances and disagree about what counts as “going on in the same way”—so there is no such thing as *the* regularity that is being reinforced by a certain set of responses to responses, or even dispositions to respond to responses. The issue of gerrymandering, of how to privilege one specification of a regularity over equally qualified competitors, arises once more at the level of the reinforcing regularity. Again, simple regularity theories are subject to the objection that they conflate the categories of what is in fact done and what ought to be done, and hence that they fail to offer construals of genuinely normative significances of performances at all.

This is a way of failing to take sufficiently seriously Kant’s distinction between acting according to a rule and according to a conception of a rule. Sanctions theories fund this crucial distinction by means of the distinction between producing a performance and assessing it. But assessing, sanctioning, is itself something that can be done correctly or incorrectly. If the normative status of being incorrect is to be understood in terms of the normative attitude of treating as incorrect by punishing, it seems that the identification required is not with the status of *actually* being punished but with that of *deserving* punishment, that is, being *correctly* punished. Of course sanctioners can be sanctioned in turn for their sanctioning, which is thereby treated as itself correctly or incorrectly done. Nonetheless, if actual reinforcement of dispositional regularities is all that is available to appeal to in making sense of this regress, it might still be claimed that what is instituted by this hierarchy of regularities of responses to regularities of responses ought not to count as genuinely normative.<sup>44</sup>

#### 4. Regularities of Communal Assessment

There are other dimensions along which it is instructive to consider sanctions theories of implicitly normative practice. Accounts that endorse the tripartite strategy by rendering practical acknowledgment of normative significance in terms of reward and punishment understand normative status, paradigmatically the significance of a performance as correct or incorrect, in terms of normative *attitudes*, paradigmatically *taking* or *treating as* correct or incorrect. Their characteristic suggestion is that this sort of practical attitude of assessment—responsive classification as correct or incorrect—is to be understood in terms of the practice of sanctioning. Giving pride of place in this way to normative attitudes in the understanding of normative statuses involves emphasizing the distinction of perspective between *assessing* a performance and *producing* that performance. Any theory that reconstructs Kant's distinction by appealing to this difference of perspectives on normative significance—the difference, namely, between consuming and producing normatively significant performances—is in one important sense a *social* theory of the sort of norm-governedness distinctive of us. Haugeland's censorious herd animals shape *each other's* behavior by their capacity not only to perform but to censure performance. Each animal in the community that is thereby constituted may (and perhaps to be a full-fledged member must) be able to do both, but as he conceives it, each act of censure involves two organisms, the censuring and the censured, the reinforcer and the reinforced.<sup>45</sup>

There is another sort of theory that combines the idea that normative statuses cannot be understood apart from normative attitudes with the idea that the relation between them involves social practices in a different but perhaps even more robust sense of 'social'. It has often been noticed that simple regularity theories of implicit norms gain no ground by shifting from concern with regularities characterizing the behavior of individuals to regularities characterizing the behavior of groups.<sup>46</sup> As a response to this concern, the leading idea of this sort of the construal of norms as implicit in social practices is that of *communal assessment*. On this approach, the key to the importance of the social is taken to lie in the possibility that the performances individual community members produce are assessed, responded to, or treated in practice as appropriate or inappropriate by the community to which the individual belongs. An individual might be taken implicitly to endorse or treat a performance as correct simply by producing it. The community, unlike the individual, need not be counted as having taken up a practical attitude regarding the propriety of the performance just in virtue of that performance's having been produced by one of its members. The class of performances produced by its members, rather, determines which fall within the scope of communal attitudes, which are liable to communal endorsement or repudiation.

Looking to assessments by the community does provide further resources for regularity theories. For a regularity can govern (that is, be displayed or conformed to by) the assessments of the community without each individual performance counting as correct according to it, and in that case a distinction between correct and incorrect performances by individuals is underwritten. In this way, the communal assessment theories are just like the theories already considered, with Haugeland's as an example, in which norms are taken to consist in regularities of practical appraisal. (It is irrelevant for the present point whether such appraisals are then understood in terms of sanctions, and the sanctions in terms of reinforcement.) The difference is that where the theories previously considered look to regularities of appraisals by individuals, the approach now on the table looks to regularities of appraisals by the community as a whole.<sup>47</sup>

There are two sorts of objections not yet considered to accounts that construe the norms implicit in practices in terms of regularities or dispositions regarding communal assessments or attributions of normative significance. First, the idea of communal performances, assessments, or verdicts on which it relies is a fiction. Second, the approach smuggles normative notions illicitly into what purports to be a reductive, nonnormative regularity theory. On the first point, communal assessment theorists have a tendency to personify the community, to talk about it as though it were able to do the same sorts of things that individual community members can do—perform additions, apply rules, assess performances, and so on. Thus to pick one page almost at random, Kripke talks about “the community's accepting” conditionals codifying relations between attributions of intentional states and commitments to act, what “the community regards as right,” what “the community endorses,” and so on.<sup>48</sup> This is a typical passage from Wright: “None of us can unilaterally make sense of the idea of correct employment of language save by reference to the authority of securable communal assent on the matter.”<sup>49</sup>

The difficulty with this way of talking is that assenting, endorsing, accepting, and regarding as right are in the first instance things done by *individuals*. It is not the community as such that assesses applications of the concept *yellow*, say, but individual members of that community. Any account that seeks to extend these notions to include cases where the subject or agent is a community should say explicitly what sort of performance or speech act is envisaged. Some communities have meetings, authorized representatives, or other ways of officially settling on a communal view or act, for instance of disapproval or endorsement. But this sort of thing is the exception and could in any case hardly be what is wanted for explaining either norms in general or conceptual, intentional, or linguistic norms more particularly.

This tendency to talk of the community as somehow having attitudes and producing performances of the sort more properly associated with individuals is neither accidental nor harmless. This *façon de parler* is of the essence of

the communal assessment approach. It is a manifestation of the orienting mistake (about which more will be heard later) of treating *I-we* relations rather than *I-thou* relations as the fundamental social structure.<sup>50</sup> Assessing, endorsing, and so on are all things we individuals do and attribute to each other, thereby constituting a community, a 'we'. But this insight is distorted by *I-we* spectacles—perhaps the same that have always been worn by political philosophers in conceiving their topic. The pretense of communal assessment is not harmless because the easy ways of cashing out the metaphor of community approval and so forth present familiar dilemmas. A notion of communal endorsement or repudiation might be built out of regularities of endorsement and repudiation by individual community members. But universal agreement is too much to ask, and how is it to be decided what less ought to count?

In fact, the approval of some almost always matters more to the community than that of others—though this division may be different for different issues. We recognize experts. But being an expert is having a certain *authority*, and that is a normative matter. One might go on to give an account of the status of having the authority to speak for the community on some matters in terms of the attitudes of attributing or recognizing that authority on the part of other community members. (A story along these lines is endorsed in what follows.) One might go on to offer sufficient conditions for the attributions of such attitudes, and so such statuses, to a community, entirely in nonnormative terms. (This move is repudiated in what follows.) But without some such story, how is one to understand talk of what the community endorses or repudiates?

Connected with this question is the problem of how community membership can be understood, in line with a communal assessment theory. This in turn is closely related to the previously mentioned problem, peculiar to regularity theories, of distinguishing those in whose practice the norms are implicit from those on whom those norms are binding. If 'norm implicit in a practice' is understood just as 'regularity truly descriptive of actual performances (or performances there are dispositions to produce under suitable circumstances)', and those performances are also the ones 'subject to' or 'governed by' the practices comprising them, then there is no possibility of irregularity, of violating a norm. Being a member of a community is rather being one who *ought* to conform to the norms implicit in the practices of the community. Community membership has this normative significance; it is a normative status.

Unless this status is understood in some way other than as being one who in fact exhibits the regularity in question, it will be impossible to violate a norm, because impossible to act irregularly. I cannot be out of step with any regularity that characterizes the behavior of each and every one of us. Extending that argument, if an account is to be offered of norms as social regularities, it is not by itself enough to identify what is correct for *all*

community members with what accords with the regularities descriptive of the practice of *some* of those community members (the experts). Some account is required of how those members are picked out. The distinction between the experts, the ones who have authority, whose actions and assessments cannot fail to be correct, and those who are subject to that authority, bound by the norms instituted by the regularities of performance of the experts, is a normative distinction, a distinction of normative status. Unless this distinction is itself understandable as a matter of regularities descriptive of the performance of community members, the social regularity account explains only some normative notions, while appealing to others.

Just as having the status of a community member is something with normative consequences—for one thereby is *subject* to a certain sort of authority, a certain standard of correctness—so having the status of an expert or official is something with normative consequences, for one thereby *exercises* a certain sort of authority. It is appropriate to ask whether the circumstances of application of the concept ‘expert’ (in the sense of authoritative assessor) could be specified in terms of regularities of conduct specified in nonnormative terms. Is there a distinction between actually assessing and being entitled to assess, one’s assessments having authority? It is just in this sense that it is appropriate to ask more generally whether the circumstances of application of the concept ‘community member’ can be specified in terms of nonnormative regularities. Is there a distinction between actually being assessed and being properly assessed—being subject to the authority of assessments?

The point is that talk about the community to which a performer belongs is not obviously translatable into talk about regularities of individual performance. Belonging to the community is a concept used so as to have normative consequences of application, concerning the member’s being responsible to the assessments of the community, being subject to its authority. An understanding of norms implicit in practice in terms of regularities of communal assessment requires the idea of a regularity that the performer is somehow bound by or answerable to. Exactly how must the performer be related to other performers to be appropriately taken to be responsible to their assessments?<sup>51</sup> It is possible that appeal will have to be made to something other than regularities of performance to secure this connection. Understanding normative status, including the normative status of being a community member, in terms of some sort of liability to being rewarded or punished is one thing; understanding that liability nonnormatively is another. The claim is not that these difficulties—with specifying what it is for a community to endorse or repudiate a performance, and the related questions of how to pick out the community members, those over whom such assessments have authority, and perhaps how to pick out experts whose assessments have such authority—are insuperable. These difficulties show only that the invocation of communal assessments does not by itself provide

a recipe for construing norms implicit in practices in terms of nonnormative regularities of performance (even assessing performances).

To recapitulate: Regularity theories of norms implicit in practice, that is of practical normative status, no matter whether they invoke specifically *social* regularities, or look in addition at regularities of *attitude* or assessment, or combine those moves in a theory that relies on regularities of communal assessment, are subject to two general sorts of objection. First, the gerrymandering argument challenges them to pick out a unique specification of the regularity in question, sufficient to project it so as to determine the application of the norms to novel cases. Second, bracketing those difficulties, attempts to define normative statuses in terms of nonnormative regularities can be criticized from two directions, as either failing to reconstruct some essential features of genuine normative statuses, or as covertly appealing to normative notions.<sup>52</sup> Under the first heading, they may fail to make room for the crucial distinction between *performance* and the *normative status* or significance of a performance, that is between what is done and what is correct or ought to be done, or they may fail to make room for the equally crucial distinction between the normative *status* of a performance and a normative *attitude* toward that status, that is between what is correct and what is taken to be correct. (More will be said about this second sort of insufficiency below.) The danger of regularity theories smuggling in normative notions arises both in specifying what it is for someone to be a member of a community in the sense of being governed by its norms, properly subject to assessments according to them, and again in specifying what it is for something to count as a communal attitude, an assessment by a community of conformity with its norms. This latter may take the form of specifying a subclass of community members whose assessments are imbued with communal authority and so have the status of official or expert assessments.

This mention of one sort of social practice theory is entirely preliminary. The issue of reasons for understanding the sort of practice in which norms are implicit as *social* practice, and in what sense of 'social', is addressed further along. It is one of the prime tasks of this work to elaborate a suitable notion of the social practices that institute the norms underlying explicitly propositional attitudes. For the approach eventually to be endorsed concerning the social nature of those implicit norms to be intelligible, something must be said here about the relation between normative and nonnormative vocabularies. Regularity theories, of whatever stripe, are (to adapt a phrase of Dretske's) attempts to bake a normative cake with nonnormative ingredients. Gerrymandering aside, the objections just mentioned represent the two ways such an enterprise can go wrong: by failing to produce a genuinely normative product or by employing some already normative raw materials. The discussion here of accounts that identify norms implicit in social practice with regularities of communal assessment arose from consideration of ways of pursuing the strategy that starts with Kant's distinction between

acting according to rules and acting according to conceptions of rules, and adapts it to accommodate Wittgenstein's pragmatic point about norms by rendering the normative attitude of taking or treating as correct (which according to the Kantian line is essential to the characteristic sort of normative status we exhibit) in terms of sanctions, of reward for what is (thereby) taken to be correct and punishment for what is (thereby) taken to be incorrect. The sanctions approach, the broadly retributive rendering of attitudes of assessment, offers aid and comfort to those concerned to tell normative stories with nonnormative vocabulary.

The animating idea is that the classification of performances needed to get this two-stage scheme off the ground, their being positive or negative sanctions, is something that can be made available in perfectly naturalistic terms. One might try to define the two sorts of sanctions in terms of the production of benefit versus harm to the one whose status or performance is being assessed. Or one might try to define them in terms of the preferences and aversions of the one assessed—respecting, as it were, the views of the sanctioned one rather than the sanctioner, as to what counts as benefiting and harming the sanctioned one. Less subjectively, reward and punishment might be understood in terms of giving pleasure and inflicting pain. Most austere, one might define rewarding or punishing a certain kind of behavior or performance in functional terms, as positively or negatively reinforcing the reliable dispositions to respond differentially to stimuli that are being realized by the performances responded to. This version takes it that to treat a response to a certain stimulus as incorrect is just to punish it, in the sense of responding to it in a way that in fact decreases the probability that the one being assessed will respond in that way to that sort of stimulus in the future.

### *5. Normative Sanctions*

In this connection it is important to realize that it is one thing to understand practical assessment as sanctioning, and quite another to understand sanctioning in nonnormative terms such as reinforcement. A retributive approach to the normative need not be given a naturalistic turn at all. Defining normative attitudes in terms of dispositions to apply sanctions does not by itself reduce the normative to the nonnormative—it just trades off one sort of norm for another. At the most basic level, to reward someone is to offer some good (either objectively or subjectively), and to punish them is conversely to inflict something bad. Benefit and harm, desirable and undesirable, are concepts that also have normative senses. Indeed, these senses would seem to be primary, so that some sort of reductive hypothesis would be needed to naturalize them. To turn the retributive story about normative attitudes into a naturalistic one, an account might for instance understand what is good (and so rewarding) in terms of what is desirable, what is desir-

able in terms of what is desired, and what is desired ultimately in terms of what is pursued.

Commitment to such a reduction is optional. Positive and negative sanctions may consist in acclaim and censure that itself has only a normative significance. A correct action might be rewarded by the grant of some extraordinary privilege or by release from some onerous obligation—and the status of such a response as reward need not depend on whether the one rewarded would in fact have been disposed to refrain from acting without the boon of entitlement or would in fact have been disposed to act so as to fulfill the obligation had it not been lifted. An incorrect action might be punished by withholding a license to act in certain other ways or by imposing an extraordinary obligation—and the status of such a response as punishment need not depend on whether the one punished is in fact disposed to refrain from acting even without the boon of entitlement, or is in fact disposed to act so as to fulfill the obligation imposed. In such cases one is rewarded or punished for what one does "in another world"—by a change in normative status rather than natural state.<sup>53</sup>

Consider once again the case of discerning a practical norm in force in a community, according to which to enter a particular hut one is obliged to display a leaf from a certain sort of tree. As pointed out above, the assessing response constituting the community's acknowledgment of such a norm (the attitude corresponding to the status) might in some cases be describable in nonnormative terms—one who violates the norm is beaten with sticks, the norm-violating behavior is negatively reinforced. But other cases are possible, for instance ones in which the assessing response is to punish by making other actions inappropriate—one who violates the norm is not permitted to attend the weekly festival. In such a case, the normative significance of transgression is itself specified in normative terms (of what is *appropriate*, of the transgressor is *entitled* to do). The punishment for violating one norm is an alteration in other normative statuses. Acting incorrectly alters what other performances are correct and incorrect.

Once again, it need not be assumed that the alteration of status according to which it becomes inappropriate to attend the tribal festival has the actual effect of disposing the transgressor not to attend. The alteration of status need itself have no reinforcing function. This could be so even if the assessing attitude corresponding to the consequential norm forbidding attendance at the festival is itself enforced by actual reinforcing responses—that is, even if it is the case that an attempt to attend the festival by one who is not entitled will be punished by beating the offending community member with sticks. In such a case, the norm regarding entitlement to attend the festival is intelligible in terms of attitudes expressed by sanctions specifiable in non-normative terms, while the norm regarding entitlement to enter the hut is intelligible in terms of attitudes expressed by sanctions specifiable only in

normative terms of consequences for entitlement to attend the festival. Since the norms governing festival attendance are directly nonnormatively intelligible, the norms governing hut entrance are also nonnormatively intelligible—but indirectly, at one remove. In this way one norm can depend on another, as the sanctions expressing assessments of the normative significance of performances according to the first norm consist in alterations of normative status with respect to the second norm.

If what qualifies some response to a performance as a sanction—and therefore, according to the retributive line being considered, as an assessment—is specifiable only in normative terms, that is in terms of the correctness or incorrectness, (the normative status) of further performances according to other norms, that kind of sanction can be thought of as being *internal* to the system of norms being discerned.<sup>54</sup> If, by contrast, what qualifies a response as a sanction is specifiable in wholly nonnormative terms of what various community members do or are disposed to do, without reference to the specifically normative status of their performances, that kind of sanction can be thought of as being *external* to the system of norms being discerned. In this terminology, the simple sort of dependence of one norm on another just considered occurs when the attitudes corresponding to one kind of normative status (e.g. propriety of hut entering) are expressed by normatively internal sanctions, defined in terms of another sort of normative status (e.g. propriety of festival attending). In the case described, that second sort of status is itself made intelligible by normatively external sanctions, ones that can be specified in terms of the movement of sticks and consequent alterations in dispositions to attempt festival attendance. But internal sanctions can be defined in terms of normative statuses that themselves are defined by internal sanctions referring one to still further norms. Clearly this sort of dependence of one norm on another according to the retributive paradigm can be extended and ramified, making sense of complex webs of interdependent normative statuses.

In the cases so far imagined, these webs of norms linked by internal sanctions are anchored, as each chain of definitional dependence terminates in some normative status that is definable independently, by external sanctions specified in nonnormative terms. Even this restriction can be relaxed. The consequences of an assessment of a performance as correct or incorrect with respect to one norm may extend no further than other assessments of correctness, with respect to other norms. It is possible to interpret a community as instituting normative statuses by their attitudes of assessment, even though each such status that is discerned is responded to by sanctions that involve only other normative statuses. It is compatible with the sanctions paradigm of assessment, and so of normative attitude, that it should be “norms all the way down.” Such an interpretation would not support any reduction of normative status to nonnormatively specifiable dispositions, whether to perform or to assess, whether individual or communal.

Norms acknowledged by external sanctions can be attributed to a community one by one, in an atomistic way. Attribution of any norm whose acknowledgment by the community takes the form of assessments expressed by internal sanctions, however, commits the interpreter to attributing also the norms on which it depends. Such dependences introduce a holistic element into the attribution of normative significances to the performances of a community. Using the retributive paradigm of normative attitudes of assessment to structure an interpretation of a community as exhibiting practices in which interdependent norms are implicit does not require that there be some nonnormatively specifiable behavior associated with the acknowledgment of each sort of normative status or significance discerned by that interpretation. An example of a system of practices in which the normative significances of performances must be attributed in a holistic way and are not translatable into nonnormatively specifiable dispositions is offered in Chapter 3, which presents sufficient conditions for such a system of practices to confer *propositional* contents on the statuses and attitudes it institutes.

A strategic divide looms here. Wittgenstein argued that an unproductive regress results from conceiving explicit rules as the only form of the normative. The lesson drawn from the regress of rules interpreting rules is the pragmatic one, that there must be "a way of grasping a rule which is not an interpretation, but which is exhibited in what we call 'obeying the rule' and 'going against it' in actual cases."<sup>55</sup> That is, there must be such a thing as norms that are implicit in practice. But what is the relation between such norm-laden practices and nonnormatively describable regularities of performance? The division of explanatory strategies arises over the question of whether the practices invoked to halt the regress<sup>56</sup> can be analyzed in terms of regularities and dispositions characterized without the use of normative vocabulary.

In line with Kant's insight that normative *attitudes*—the sort of uptake of or sensitivity to norms that he talked about in terms of *conceptions* of rules—are essential to the way in which our conduct is governed by norms, two suggestions have been put on the table.<sup>57</sup> The first is the idea of constructing the normative attitude of taking or treating something as correct or incorrect in practice in terms of the application of positive and negative *sanctions*. The second is the idea that these assessing attitudes have a fundamentally social structure, so that the practices in which norms are implicit ought to be understood as essentially *social* practices. Each of these ideas could be given a reading in naturalistic or nonnormative terms, as part of a reductive explanatory strategy. Putting them together would then yield an approach that understands norms as implicit in regularities or dispositions regarding communal assessments of performances as correct or incorrect, and that understands such assessments in turn as behaviorally reinforcing rewards and punishments.

Wittgenstein certainly emphasizes the social nature of the practices un-

derlying the norms involved in discursive intentional states, and the importance to those practices of similarity or agreement of dispositions to respond to performances as correct or incorrect. There is accordingly a temptation to understand him as responding to the regress-of-rules argument by putting forward a reductive social regularity account of the practices in which norms are implicit.<sup>58</sup> But he need not be understood this way. His insistence that unless the responsive dispositions of a community are consilient, there can be no proprieties of practice is a point concerning presupposition, not reduction. Wittgenstein's is the somewhat delicate position first, that the usefulness of normative attributions, the viability of this stratum of discourse, presupposes a variety of regularities of performance and disposition; second, that those regularities obtain is *not* part of what is *asserted* by such attributions.

An analogy he comes back to again and again is the measurement of length, in which the possibility of practices of measurement presupposes features of the world such as the rigidity, spatial invariance under transportation, and temporal constancy of measuring rods, interpersonal comparability of measurements, the functional equivalence of various means of measuring the same length, the irrelevance to the result of such contextual features as whether the object measured is sacred or profane, to be used in sport or commerce, and so on. That we can be trained so as almost always to respond in the same way when applying concepts to novel cases, for instance, is a necessary condition of there being a practice determining what response is correct in such cases. But this is not to say that what it is for it to be correct consists in this agreement, as the reductive social regularity account of those norms would have it. There are three levels at which performances can be discussed: a level of norms explicit in rules and reasons, a level of norms implicit in practice, and a level of matter-of-factual regularities, individual and communal. To say that various claims made at the third level state necessary conditions for the applicability of vocabulary of the sort employed at the first two is not to make a reductive claim. The social regularity view conflates the second and the third levels, and so misunderstands Wittgenstein's remarks about the significance of matter-of-factual regularities, by taking them to involve commitment to the possibility of a reduction of the normative to the dispositional.<sup>59</sup>

## V. FROM ASSESSMENT TO THE SOCIAL INSTITUTION OF NORMS

### 1. *Pufendorf on the Institution of Norms by Attitudes*

As discursive beings whose characteristic activities are applying concepts, giving and asking for reasons, taking-true and making-true, we live and move and have our being in a space structured by norms. Yet we can describe, and largely successfully cope with, the not-us around us, while

restricting ourselves to a resolutely nonnormative vocabulary. In thinking about the relation between acting according to conceptions of rules (on a suitably pragmatic reading of what that consists in), as we do, and merely acting according to rules, as the rest of *it* does, it is important to distinguish two ways in which the normative significances we assign to things might be thought to be unnaturalized second-class citizens in an intrinsically insignificant natural world. These correspond to two different sorts of domestication to which normative statuses might be subjected. Couched in terms of supervenience, they are the claim that settling all the facts specifiable in nonnormative vocabulary settles all the facts specifiable in *normative* vocabulary, on the one hand, and the claim that settling all the facts concerning normative *attitudes* settles all the facts concerning normative *statuses*, on the other.

These are intimately related claims; the difference between them is subtle, and they are often run together. Each is the heir to a line of thought central to and characteristic of the Enlightenment project of disenchanting the natural world and humanizing values. The first can trace its origins to atoms-in-the-void physicalism—the conviction that a specification of the values of an appropriate range of dynamic variables for all the fundamental particles provides a *complete* description of everything that deserves to be called *real*. The second is animated by the humanistic thought that the merely natural world is devoid of values, that the worth of things and the fitness of actions is a product of our activity—that unlike natural properties, normative proprieties are in the eye of the human beholder. These ideas are of course at work in the thought of many Enlightenment philosophers. For present purposes it will suffice to consider briefly representative statements by one of the earliest.

The second line of thought emerges most clearly in the thought of the pioneering philosopher of law Samuel (Freiherr von) Pufendorf (1632–1694). Although not much read by philosophers today, his magnum opus *De Jure Naturae et Gentium* stands at the beginning of a tradition of Enlightenment thought about norms that culminates in Kant's practical philosophy (which was greatly influenced by Pufendorf). The relative unfamiliarity of these seminal views perhaps excuses quotation at greater than usual length. These passages all come from the opening of the work, in Chapter 1, entitled "On the Origin and Variety of Moral Entities": "It is for us to observe, how, chiefly for the direction of the acts of the will, a specific kind of attribute has been given to things and their natural motions, from which there has arisen a certain propriety in the actions of man, and that outstanding seemliness and order which adorn the life of men. Now these attributes are called Moral Entities, because by them the morals and actions of men are judged and tempered, so that they may attain a character and appearance different from the rude simplicity of dumb animals."<sup>60</sup> We are distinguished from the brutes by the fact that our actions are subject to assessment according to their

propriety, a special kind of attribute over and above the natural motions of things, an attribute that has somehow been given to them.

What is the source of those special normative attributes? "We seem able, accordingly, to define moral ideas most conveniently as certain modes [qualities], added to physical things or motions, by intelligent beings, primarily to direct and temper the freedom of the voluntary acts of man, and thereby to secure a certain orderliness and decorum in civilized life."<sup>61</sup> Intelligent beings add these properties to things by their activities. They are called 'moral' entities in virtue of their practical function as guides to action.

Now as the original way of producing physical entities is creation, so the way in which moral entities are produced can scarcely be better expressed than by the word *imposition*. For they do not arise out of the intrinsic nature of the physical properties of things, but they are super-added, at the will of intelligent entities, to things already existent and physically complete, and to their natural effects, and, indeed, come into existence only by the determination of their authors. And these authors give them also certain effects, which they can also remove at their own pleasure without any accompanying change in the object to which they had been added. Hence the active force which lies in them does not consist in their ability directly to produce any physical motion or change in any thing, but only in this, that it is made clear to men along what line they should govern their liberty of action.<sup>62</sup>

These norms are not part of the intrinsic nature of things, which is entirely indifferent to them. They are imposed by the will of intelligent beings and can affect things only through their effect on the will of such beings—beings who can act according to a conception of them. "Since, therefore, moral entities have been instituted to bring order into the lives of men, for which purpose it is required that they also, who must live according to their rule, should adopt a set standard in their relations toward one another, in determining their actions, and finally in fixing their attitude toward those things which are used in the lives of men; for this reason they are understood to be inherent primarily in men, but also in their actions, and even, to some extent, in things."<sup>63</sup> Our activity *institutes* norms, *imposes* normative significances on a natural world that is intrinsically without significance for the guidance or assessment of action. A normative significance is imposed on a nonnormative world, like a cloak thrown over its nakedness, by agents forming preferences, issuing orders, entering into agreements, praising and blaming, esteeming and assessing.<sup>64</sup>

One of the defining characteristics of early science is its disenchantment (*Entzauberung*, in the word we owe to Weber) of the world. The meanings and values that had previously been discerned in things are stripped off along with the supernatural and are understood as projections of human interests, concerns, and activities onto an essentially indifferent and insignificant mat-

ter. The Enlightenment disenchantment of the world and its assignment to us of responsibility for the norms, values, and significance we nonetheless find in the world are two sides of one coin. Meaningless objects and meaning-generating subjects are two aspects of one picture.<sup>65</sup> On this view, valuing is the source of values—a tradition carried on by figures as disparate as Mill and Nietzsche. Contractarian theories, and those that invoke positive law to explain various rights and obligations, are species of this genus. Each explains these deontic statuses in terms of what agents are doing in instituting or constitutively recognizing such entitlements and commitments.

Pufendorf does not suggest, and he does not believe, that the activity by which we institute norms is itself describable in the purely physical terms that suffice to describe the antics of merely natural objects. His claim is that the normative statuses of things, the normative significances we take them to have, are products of our practical normative attitudes, as expressed in our activity of imposing those significances and acknowledging them in assessments.<sup>66</sup> He does not conjoin this thesis with any sort of physicalism about the mechanism by which these moral secondary qualities arise from our practical activity. It is clearly possible to agree with the dictum of another Enlightenment thinker, Hamlet, that "There is nothing either good or bad, but thinking makes it so," without holding in addition a reductionist view about such thinking.

One who does conjoin these commitments is Hobbes. He explains, not good and evil, but *calling* things good and evil. He understands the use of the words 'good' and 'evil' as expressing appetites, desires, or aversions. He expresses his commitment to fundamental normative statuses being instituted by our attitudes this way: "But whatsoever is the object of any mans Appetite or Desire; that is it, which he for his part calleth Good: And the object of his Hate, and Aversion, Evill . . . For these words of Good, Evill . . . are ever used with relation to the person that useth them: There being nothing simply and absolutely so; nor any common Rule of Good and Evill, to be taken from the nature of the objects themselves; but from the Person of the man."<sup>67</sup> Gauthier comments on one difference between the sort of view Hobbes endorses concerning the relevant norm-instituting practical attitudes and the kind of view Hamlet endorses on this point: "If things considered in themselves are neither good nor bad, if there is no realm of value existing independently of animate beings and their activities, then thought is not the activity that summons value into being . . . Desire, not thought, and volition, not cognition, are the springs of good and evil."<sup>68</sup>

Pufendorf would not disagree. Where he does disagree is with Hobbes's subsequent endorsement of materialism about the will. The claim that normative statuses are instituted by our normative attitudes entails the claim that the normative proprieties so instituted are just natural properties of a special kind only in the context of a collateral claim that the norm-instituting practical attitudes can themselves be specified in nonnormative terms.

One can hold, as Pufendorf does, that there are no values apart from our acknowledgment or recognition of them, or more generally our attitudes toward them, without being thereby obliged in addition to understand those attitudes in terms of desires or preferences that can be characterized independently in value-free terms.

## 2. Kantian Autonomy: *The Authority of Norms Derives from Their Acknowledgment*

Pufendorf's idea that normative statuses are *instituted* by our practical attitudes makes a stronger claim than the idea previously extracted from Kant's demarcation of us as beings who act not only according to rules but according to our conceptions of rules. For the latter idea requires only that the normative statuses of demarcational interest essentially *involve* the uptake or grasp of such statuses, that is, our practical attitudes toward them. But normative statuses could be taken to be unintelligible apart from normative attitudes without thereby being taken to be instituted by and therefore in some sense to supervene on those attitudes. However Kant does in fact subscribe also to a version of the stronger thesis about attitudes instituting statuses, for the case of the genuinely moral normative statuses characteristic of us as agents.<sup>69</sup>

Kant's practical philosophy, his account of us as agents, takes its characteristic shape from his dual commitments to understanding us as *rational* and as *free*. To be rational, for him, means to be bound by rules. But Kant is concerned to reconcile our essential nature as in this way bound by norms with our radical autonomy. He combines the essential defining moment of our dependence on universals with that of our independence as particulars (as Hegel puts the point) in the thesis that the *authority* of these rules over us derives from our *acknowledgment* of them as binding on us. Our dignity as rational beings consists precisely in being bound only by rules we endorse, rules we have freely chosen (like Odysseus facing the Sirens) to bind ourselves with. We do not have the freedom to opt out entirely—choosing to be bound by no rules at all would be choosing to relinquish our rationality entirely. Yet if something other than our own attitudes and activity could bind us, we would not be free. Autonomy consists, as the etymology demands, in setting up laws for ourselves.

This view of Kant's inherits a venerable Enlightenment tradition.<sup>70</sup> It is based on a certain picture of the nature of the *authority* of rules or laws (the only form of norm considered). Pufendorf takes it that "since good repute, or moral necessity, and turpitude, are affections of human actions arising from their conformity to some norm or law, and law is the bidding of a superior, it does not appear that good repute or turpitude can be conceived to exist before law, and without the imposition of a superior."<sup>71</sup> More generally: "A

law may most conveniently be defined as a decree by which a superior obligates a subject to adapt his actions to the former's command.”<sup>72</sup>

The key notion is that of a *superior*, someone who has *right* to command.

An obligation is properly laid on the mind of man by a superior.<sup>73</sup>

The power of obligating, that is, the faculty of laying an intrinsic necessity on persons to do something, properly lies in him who has authority or sovereignty.<sup>74</sup>

To see the question of authority in terms of who can command, make a rule binding, or lay down the law is another bit of fallout from the origin of thought about norms in thought about the institution of explicit positive laws. The issue of sovereignty is just the issue of “who's to be master, that's all,” as the linguistic Leninist, Humpty Dumpty, says. The authority of norms depends on the nature of the author of the commands that make them explicit; their bindingness derives from the interpersonal relation of superior to subordinate.

The *consequences* of such a relation of authority being in force, what follows from a rule or law inheriting the authority of a superior lawgiver, Pufendorf conceives in terms of sanctions.<sup>75</sup> The *antecedents* or grounds on which such a relation is based, what makes one individual superior to a subordinate other in this normative sense, he conceives disjunctively: “Mere strength [to sanction] is not enough to lay an obligation on me at the desire of another, but that he should in addition have done me some special service, or I should of my own accord consent to his direction . . . But when a man of his own accord consents to the rule of another, he acknowledges by his own act that he must follow what he himself has decided.”<sup>76</sup> The “special service” clause is put in to allow our creator a special claim on our obedience.

That special pleading aside, the standard Enlightenment thought, common as well to Hobbes and Rousseau among the progenitors of Kant's theory, is that our own *acknowledgment* or *endorsement* of a rule is the source of its authority over us—in short that our normative statuses such as obligation are instituted by our normative attitudes. Authority is not found in nature. The laws of nature do not bind us by obligation, but only by compulsion. The institution of authority is human work; we bind ourselves with norms. Contract theories are the result of combining a conflation of norms with their explicit expression in rules or laws, an understanding of their authority or bindingness on the hierarchical model of superior/subordinate (each given aid and comfort by the tradition of legalism), and an insistence on rational dignity as demanding autonomy.

Kant's reconciliation of us as free in virtue of being rational, with us as bound by norms in virtue of being rational—and so of freedom as constraint by a special kind of norm, the norms of rationality<sup>77</sup>—accordingly involves treating the normative status of moral obligation as constituted by normative

attitudes. It is our attitude toward a rule, our acknowledgment or recognition of moral necessity alone, that gives it a grip on us—not just in terms of its effect on our actual behavior, but in terms of our liability to assessment according to the rule that expresses that necessity. In this sense the norms that bind us rational creatures are instituted by our practical attitudes and activity. They are what we bring to the party. But while Kant in this way endorses the supervenience of moral normative status on moral normative attitude, he does not endorse any sort of naturalism or reductionism about those attitudes. He does not take it that specifications of those normative attitudes supervene on specifications of the movements of particles, described exclusively in the vocabulary of natural science. Grounding normative status in normative attitude does not entail relinquishing the distinction between normative proprieties and natural properties.

### *3. Objectivity and the Social Institution of Conceptual Norms*

Kant also takes it that we are genuinely bound by the rules we endorse. This requires that once we endorse one, it is not up to us what it demands—there is some fact of the matter as to what we have thereby obliged ourselves to do. Although the status of being obliged to follow a particular rule is instituted by our attitudes, what is correct according to that rule is not simply determined by what we *take* to be correct according to it. The status of correctness of a performance according to a rule does not collapse into the attitude of assessing that performance as correct. Endorsing a rule gives it a grip on us. Part of that grip is that the rule does not mean just whatever we later might take it to mean. So Kant underwrites not only the possibility of *mistakes of performance*, which was already claimed to be essential to there being norms in play, but also the possibility of *mistakes of assessment*.

Wittgenstein appeals to this possibility as a criterion of adequacy for an account of norms being in force; talk of norms being implicit in a practice requires that there be room for a distinction between what is correct according to the norm and what the one whose performances are being assessed *takes* to be correct. He brings this consideration to bear against the possibility of setting up rules for oneself whose meaning is determined *only* by one's own actual dispositions to make assessments. This is the line of thought that concludes: "One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'."<sup>78</sup> The thought here is that the distinction between status and assessment (the attitude of taking or treating a performance *as* correct, appropriate, or in order) is essential to the notion of genuinely normative status. It is motivated by the idea that assessing is itself something that can be done correctly or incorrectly, and furthermore that it is the norm according to which perform-

ances are being assessed that determines which assessments are correct or incorrect. If there is no distinction to be made between correct and incorrect assessments, then there is no sense in which the performances being assessed are governed by a norm *according to which* they are being assessed.

One of Wittgenstein's most important claims is that the practices in which the norms that articulate meanings and their uptake in understanding are implicit must be *social* practices. It is clear that the emphasis on their social character emerges for him somehow from the need to keep the notion of what one is committed to by the application of a concept distinct in principle from what one *takes oneself* to be committed to thereby. One natural way of understanding how such considerations can lead to the conclusion that discursive practice must be social practice is elaborated by Crispin Wright.<sup>79</sup> He combines an understanding of conceptual norms as instituted by practical normative attitudes—taking or treating various uses as correct or incorrect—with a way of maintaining a distinction between conceptual commitments and individual assessments of those commitments. He does so by identifying the normative status of being a correct application of a concept with being taken to be such a correct application, not by an individual, but by the whole community. According to this story, although individual performances can be correct or incorrect, and assessments of correctness by individuals can be correct or incorrect, no such difference applies to communal assessments. The community is incorrigible about what is a proper application of a concept and what is not. Communally endorsed applications of a rule, or, in the idiom employed here, acknowledgments of a norm implicit in the practice of the community, cannot be mistaken. "For the community itself there is no authority, so no standard to meet."<sup>80</sup>

Thus Wright secures a distinction between the commitments one undertakes in employing a particular concept and any individual's attitudes toward or assessments of those commitments, but at the cost of obliterating any such distinction between normative status and the attitudes of the whole community. There clearly are socially instituted norms of this sort. Whatever the Kwakiutl treat as an appropriate greeting gesture for their tribe, or a correctly constructed ceremonial hut, is one; it makes no sense to suppose that they could collectively be wrong about this sort of thing. The question is whether *conceptual* norms ought to be understood as being of this type. There is good reason to think they ought not. It is a fundamental feature of our understanding of our concepts that they incorporate *objective* commitments. Thus, our use of the term 'mass' is such that the facts settle whether the mass of the universe is large enough that it will eventually suffer gravitational collapse, independently of what we, even all of us and forever, take those facts to be. We could all be wrong in our assessment of this claim, could all be treating as a correct application of the concepts involved what is objectively an incorrect application of them.

On Wright's view the objectivity we take our conceptual norms to have is

an illusion that must be surrendered if they are to be properly understood. The normative attitudes discussed here under the heading of assessment—taking or treating applications of concepts as correct or incorrect—he talks about in terms of “ratification.” Understanding conceptual norms as objective, in the sense that the whole community can coherently be conceived to be wrong in its assessment of the commitments involved in some applications of its concepts, is taking them to be ratification-independent, in his terminology. McDowell, whose insistence on the objectivity of conceptual norms was discussed in this connection in Section IV, summarizes his disagreement with Wright on this point in this way: “In Wright’s reading . . . Wittgenstein’s point is that the natural contractual conception of understanding should not be discarded, but purged of the idea—which it must incorporate if the intuitive notion of objectivity is to have application—that the patterns to which our conceptions oblige us are ratification-independent. I expressed a suspicion above that this purging would not leave a residue recognizable as a conception of meaning and understanding at all.”<sup>81</sup> Wright takes it that understanding how the normative statuses involved in concept use are instituted by practical normative attitudes of assessing or ratifying the propriety of particular applications of concepts, while keeping normative statuses from collapsing into normative attitudes in a way that obliterates the norms entirely, at once requires understanding the practices of concept use and its assessment as *social* practices and relinquishing the idea that conceptual norms are *objective*.

A central aim of the present study is to show, by contrast, how these criteria of adequacy can be satisfied without giving up the objectivity of conceptual norms. Indeed the primary explanatory challenge to a social practice theory of discursive commitments is to show how, starting with the sort of norms for which Wright’s analysis is correct—normative statuses about which the community’s all-inclusive practical assessment cannot be mistaken, such as who is really married or what obligations are incurred by spitting in front of the chief—genuine, and therefore *objective*, conceptual norms can be elaborated. These bind the community of concept-users in such a way that it is possible not only for individuals but for the whole community to be mistaken in its assessments of what they require in particular cases.

How does objectivity precipitate out of the social soup of norms that are whatever the community takes them to be? According to the answer elaborated in Chapter 8, it is precisely the objectivity of conceptual norms, when properly understood, that leads to the requirement that the practices in which such norms are implicit be *social* practices. The objective representational dimension of conceptual content—the kind of correctness of claiming or concept application that answers not to individual or communal attitudes or assessments but to the properties of the things represented—turns out to depend on the *social* articulation of the inferential practice of giving and asking for reasons. Focusing on the distinction of social perspec-

tive between *acknowledging* (and thereby undertaking) a commitment oneself and *attributing* a commitment to another makes it possible to understand the objectivity of conceptual norms that consists in maintaining the distinction between the normative *statuses* they incorporate and the normative *attitudes* even of the whole community—while nonetheless understanding those statuses as constituted by the practical normative attitudes and assessments of community members. Far from precluding the possibility of conceptual objectivity, understanding the essentially social character of the discursive practice in which conceptual norms are implicit is just what makes such objectivity intelligible.

## VI. FROM INTENTIONAL INTERPRETATION TO ORIGINAL INTENTIONALITY

### 1. *The Stance Stance*

The normative house has many mansions. The particular norms of concern in this work are *discursive* normative statuses, the sort of commitment and entitlement that the use of concepts involves. These norms, it will be claimed, are constituted by *social* practices. These are practices that incorporate the distinction of social perspective between two kinds of practical attitude one can adopt toward a commitment: *acknowledging* it (oneself) and *attributing* it (to another). Elaborating an account along these lines is pursuing three of Wittgenstein's grand themes: the insistence on the *normative* character of language and intentionality, the *pragmatist* commitment to understanding these norms in terms of practices rather than exclusively in terms of rules, and the recognition of the essentially *social* character of such norms. One way in which the significance of the social character of the attitudes that institute intentional norms can begin to be approached is by considering the relation between the practical activity of intentional interpretation and the intentional states that are attributed by such interpretations.

Dennett's original account of intentional systems and intentional explanations provides a useful place to start.<sup>82</sup> One characteristic feature of that account is the idea that intentionality ought to be understood in terms of ascriptions of intentionality. Explanatory pride of place is granted to a certain sort of *attitude*, what Dennett calls a "stance." To adopt the intentional stance toward some system is to offer an intentional explanation of its behavior, by attributing intentional states to it. Adopting the intentional stance toward something is *taking* or treating it in practice as an intentional system. The status of being an intentional system, of exhibiting intentional states, is constituted by this attitude or stance: "A particular thing is an intentional system only in relation to the strategies of someone who is trying to explain and predict its behavior."<sup>83</sup> Dennett's explanatory strategy is first

to define what it is to adopt the intentional stance, that is to offer an intentional explanation, then to explain when it is appropriate to adopt that stance, and finally to define an intentional system as whatever is appropriately treated as one by adopting the intentional stance toward it. A few words are in order about each of these moves.

The beginning of wisdom about intentional explanation lies in appreciating the *normative* significance of attributing intentional states (mentioned in Section II of this chapter). Attributing suitably related beliefs and desires is attributing a certain sort of *reason* for action. Taking someone (1) to believe that it is raining and that the only way to stay dry is to open an umbrella and (2) to desire to stay dry is taking that individual to have a reason to open an umbrella.<sup>84</sup> To say this is not yet to say that the one who has such a reason *will* act according to it, even in the absence of competing reasons for incompatible courses of action. What follows immediately from the attribution of intentional states that amount to a reason for action is just that (*ceteris paribus*) the individual who has that reason *ought* to act in a certain way. This ‘ought’ is a *rational* ought—someone with those beliefs and those desires is rationally obliged or committed to act in a certain way. The significance of the states attributed is in the first instance a matter of the force of the better reason, rational force. That, as previously remarked, is a normative affair. Intentional interpretations attribute normative statuses, whose significance concerns practical proprieties. This is not to deny that reasons can be causes. It is just to unpack slightly what is meant by saying that they are reasons. The relation of such normative attributions of status and propriety to attributions of natural states and properties is a further issue.

Dennett acknowledges the normative core of intentional attribution and the corresponding distinction between physical and intentional explanation: “Deciding on the basis of available empirical evidence that something is a piece of copper or a lichen permits one to make predictions based on the empirical theories dealing with copper and lichens, but deciding on the basis of available evidence that something is (may be treated as) an intentional system permits predictions having a normative or a logical basis rather than an empirical one.”<sup>85</sup> Attributing a natural state or property such as being copper supports *descriptive* conclusions about how the subject of those attributions *will* (in fact) behave. Attributing a normative status or propriety such as having beliefs and desires that amount to a reason for opening one’s umbrella supports *prescriptive* conclusions about how the subject of those attributions *ought* (rationally) to behave. Within Dennett’s project, however, the ultimate interest of intentional explanation lies in its use in deriving predictions concerning actual behavior. Some additional premise is required to get from the prescriptive conclusions that intentional attributions immediately supply to the descriptive predictions Dennett is concerned with.

He supplies the additional premise, in the form of a *substantive rationality assumption*, to the effect that agents generally do what one ought (ration-

ally) to do, what one is committed by one's intentional states to do. To be rational in Dennett's sense is to act as one rationally ought, to act as one's intentional states commit or oblige one to act. In order to derive predictions of actual behavior from attributions of intentional states, it is necessary to add the assumption that the subject to which those states are attributed is rational in this sense. In other words, intentional interpretation supplies a *primary* intentional explanation of the *normative* status of the one interpreted, an account of what performances are *appropriate* in the light of the beliefs and preferences attributed. Supplemented by a substantive rationality assumption, these normative characterizations can be used to ground predictions about actual performances, yielding a *secondary* intentional explanation of behavior described in nonnormative terms. The substantive rationality assumption provides the bridge that connects the normative significance of intentional attribution with the actual dispositions of the subject of such attribution.

Dennett's most controversial claim is his stance stance—his claim that there is no room for a distinction between actually being an intentional system and being appropriately treated as one. Intentional systems, things that have intentional states, just are whatever things it is predictively useful to adopt the intentional stance toward. The point of the stance idiom is that the notion of someone viewing or treating something as an intentional system is to be prior, in the order of explanation, to that of being an intentional system. The only notion of intentional system Dennett permits himself is "what one is treating something as when one offers intentional explanations of its behavior." Intentionally interpreting, adopting an intentional interpretive stance, is a practical attitude, and proprieties governing that practical attitude institute intentional states and hence normative statuses. Intentional states and intentional systems are, if not in the eye of the beholder, in the successful explanatory strategies of the theorist.

In the same way, Dennett distinguishes the significance of *ascribing* intentional states to a system from that of *describing* the system. Intentional ascriptions are appropriate according to their predictive utility, not their descriptive accuracy. The appeal to stances or attitudes as prior in the order of explanation to intentional states or normative statuses need not be read this way, however. For the contrast between talking about something in intentional vocabulary and talking about it in physical vocabulary is not for Dennett a distinction between adopting a stance and doing something else. The physical stance is also a stance. What appears in the instrumentalist reading as a distinction between what is really out there and what it is convenient from the point of view of prediction to attribute (the naturalized version of the proprieties of takings as intentional attributions), between describing and ascribing, between representing and adopting a stance or attitude, is, for Dennett, not itself a factual matter, a matter of what is out there, a matter of whether our representings do or do not correspond as they ought

to what is represented. It is rather a distinction between two stances that one may adopt. Taking there to be a physical fact of the matter determining the proprieties of our takings is adopting one stance; taking it that there are only predictive conveniences determining those proprieties is another. It is stances all the way down.

For Kant the difference between the realm of Nature and the realm of Freedom, and hence in the ordinary sense the distinction between facts and norms, is itself not a factual but a normative difference (the difference between acting according to rules and acting according to conceptions of rules). So one might say that for Dennett the difference between physical systems and intentional systems is itself a normative difference, a matter of the *propriety* of adopting different explanatory-predictive stances to the system in question. After all, for something to be a sample of copper is just for it to be proper or correct to treat it as one, in one sense of 'proper or correct' (the objective representational sense discussed at the end of the previous section), just as for something to be an intentional system is for it to be proper or correct to treat it as one, in another sense of 'proper or correct'. The question is how to understand the relation between the kinds of norms that govern the adoption of these different sorts of stance or attitude. It follows that Dennett's strategy of treating the normative significances of intentional states as instituted by the attitudes of interpreters does not by itself involve a commitment to reducing the normative to the nonnormative, insofar as it is *proprieties* of attitudes that are invoked. That reductive commitment comes in later, in explaining those proprieties. Understanding those proprieties in terms of predictive success, as Dennett does (a strategy different from that to be pursued here) gives an objective basis to the norms governing the adoption of the intentional stance. It puts Dennett in a position to say that talk of the predictive utility of adopting that stance is just a way—indeed, the only one available to us—of specifying an important kind of objective pattern of behavior. Thus the normative status of being an intentional system does not collapse into the adoption of normative attitudes of intentional interpretation.

## *2. Different Stances and Kinds of Intentionality*

Understanding being an intentional system in terms of being appropriately taken or treated as an intentional system by being intentionally interpreted is not as such a reductive strategy for understanding intentionality in nonintentional terms. Offering intentional explanations of the behavior of others is something that only intentional systems can do. What is the relation between the intentionality that an intentional interpreter and attributor attributes, and that which the interpreter exhibits or possesses? To attribute beliefs (and desires and intentions), to adopt the intentional stance, one must have the concept of belief (desire, intention) and the capacity to

acquire beliefs concerning the appropriateness of applying that concept in understanding the behavior of candidate intentional systems. According to Dennett, intentional systems that can take up the intentional stance toward other systems have a special kind of intentionality. Intentional interpreters belong to "the subclass of intentional systems that have language, that can communicate": "Just as not all intentional systems currently known to us can fly or swim, so not all intentional systems can talk, but those which can do this raise special problems and opportunities when we come to ascribe beliefs and desires to them. That is a massive understatement, since without the talking intentional systems, of course, there would be no ascribing beliefs, no theorizing, no assuming rationality, no predicting."<sup>86</sup> Clearly, then, it is not possible to understand the second-class sort of intentionality attributed by creatures who offer intentional explanations of others, without understanding the first-class sort of intentionality those attributors themselves display. Dennett's assumption that possessing intentional concepts and attributing intentional states such as belief—that theorizing, predicting, assuming, and explaining—all presuppose specifically linguistic capacities is not universally shared, although good reasons for it will emerge. For present purposes, what matters is the distinction between first- and second-class intentionality—the kind possessed by attributors of intentionality, and the kind possessed by those to whom intentionality is attributed, rather than the specific characterization of the former. For it now appears that the intentionality of relatively simple systems such as animals and chess-playing computers, toward which Dennett takes it to be appropriate to adopt the intentional stance, can be understood only against the background of an understanding of the more complex systems capable of adopting that explanatory stance.

Thus it must be asked whether the fact that something is an intentional system in the first-class sense of attributing intentionality is a fact of the same general sort as the fact that something is an intentional system in the second-class sense of attributed intentionality. Dennett says of the second, attributed variety of intentionality, that the only facts in the vicinity are facts about the propriety of adopting a certain kind of stance toward it. Is the same thing true of the ascriber? Are the only facts about whether what one is doing is ascribing or attributing intentionality facts about the practical propriety of adopting a certain stance toward the interpreter, treating it in a certain way? Is adopting the intentional stance something one really does, or is the taking of a stance merely something that is sometimes appropriately attributed, so that it can be correct to adopt the stance that someone is adopting the intentional stance? Is it in this sense stances all the way down?

Distinguishing *simple intentional systems*, which are merely intentionally interpretable, from *interpreting intentional systems*, systems toward which the intentional stance can be adopted from systems that can adopt that stance toward others, is distinguishing instituting intentionality from instituted intentionality. Simple intentionality, which on this line is in the

eye of the beholder, is for that reason *dependent* on and in an important sense *derivative* from the intentionality exhibited by interpreters. The clearest examples of the derivative character of some intentionality or conceptual content are those in which interpreters explicitly assign some meaning to an intrinsically meaningless event by deciding to understand it in a certain way. They *make* an event mean something ("One if by land, and two if by sea . . .") by *taking* it to mean that, by understanding it that way. The meaning is conferred on the occurrence by the response to it that becomes appropriate, by the conclusions that are drawn from it (" . . . And I on the opposite shore will be"). The intentional content of the signal derives from the intentional content of the beliefs it makes appropriate for its audience. Noises and marks on paper do not mean anything all by themselves. Meaning is correlative with understanding, and they understand nothing. It is the possibility of *our* understanding them as expressing a content involving the application of concepts that makes *them* mean anything. Our understanding, our practices of interpretation institute that meaning, which derives from them.

The intentionality, the conceptual content, of noises and marks is borrowed from and dependent on that of the thoughts and beliefs that interpret them, the takings, or practical attitudes that attribute such content. On pain of an infinite regress, it seems necessary to distinguish the *derivative intentionality* such merely interpretable items display from the *original intentionality* their interpreters display. Clearly the simple intentionality of systems that can be interpreted as having and acting according to beliefs and desires is not derivative in the same sense in which that of inanimate marks and noises is. Nonetheless, on a view such as Dennett's the intentional contentfulness of the states of such systems depends on their interpretability by other, more capable systems. In the case both of inanimate and animate interpretables, the attempt to understand the sort of intentionality they display drives one back to the practical attitude or activity of interpreting. They exhibit no intentionality intelligible in its own right, apart from the practical attitudes of the interpreting systems.

Here, then, is a challenge: to maintain the stance toward both simple and interpreting intentional systems—that is, to acknowledge that the normative status of being such intentional systems is intelligible only by reference to the normative attitude of taking or treating something as such a system, that is interpreting it as one—while at the same time securing the distinction between original and derivative intentionality—and so not allowing the notion of intentional normative status to collapse into that of the normative attitude of intentional interpretation. This ought to seem hard to do. Indeed, Searle claims in effect that it is impossible<sup>87</sup>—that if derivative intentionality is to be intelligible, so must a sort of 'intrinsic' intentionality possessed by intentional interpreters, which can be made sense of quite apart from any reference to anyone taking or treating the states of those interpreters as intentionally contentful. From the point of view of the present project,

the regress argument he employs to derive that conclusion is flawed by its dependence on an inappropriate model of what it is to take or treat something as intentionally contentful. For following his treatment in *Speech Acts*,<sup>88</sup> he understands taking or treating a mark or noise as expressing a certain claim or proposition as depending on propositionally explicit beliefs and intentions regarding it—as interpreting it in Wittgenstein's sense. A version of the regress-of-rules argument then shows that those beliefs and intentions cannot have their meaning conferred on them in the same way. But this leaves out the possibility of conferral of such content by *implicit* practical taking or treating of states, performances, and expressions as intentionally contentful. This is the possibility pursued in Chapter 3.

The theory developed in this work can be thought of as an account of the stance of attributing original intentionality. It offers an answer to the question, What features must one's interpretation of a community exhibit in order properly to be said to be an interpretation of them as engaging in practices sufficient to confer genuinely propositional content on the performances, statuses, attitudes, and expressions caught up in those practices? The key to the account is that an interpretation of this sort must interpret community members as taking or treating each other in practice as adopting intentionally contentful commitments and other normative statuses. If the practices attributed to the community by the theorist have the right structure, then according to that interpretation, the community members' practical attitudes institute normative statuses and confer intentional content on them; according to the interpretation, the intentional contentfulness of their states and performances is the product of their own activity, not that of the theorist interpreting that activity. Insofar as their intentionality is derivative—because the normative significance of their states is instituted by the attitudes adopted toward them—their intentionality derives from each other, not from outside the community. On this line, only communities, not individuals, can be interpreted as having original intentionality.

For this to work, the practices that institute the sort of normative status characteristic of intentional states must be *social* practices. Those practices essentially incorporate a distinction of social perspective between the attitudes of *undertaking* a commitment, as someone who believes that a bear is approaching might be taken to be committed to believing that an animal is approaching, and *attributing* a commitment, as the one who interprets another as having such a belief might do. The first sort of attitude toward a normative status must be attributed even to simple intentional systems—the rationality that is for Dennett the mother of intention is a way of talking about the sense of 'ought' in which one who believes a bear is approaching ought to believe that an animal is approaching. Just for that reason, the second sort of attitude is an implicit version of adopting the intentional stance. According to the account offered in Chapter 3 of the practices that confer distinctively propositional contents (and accordingly underlie all dis-

cursive intentionality, the conceptual contentfulness of expressions, performances, attitudes, and statuses), the practical normative attitudes of undertaking and attributing commitments come as a package—neither is intelligible apart from the other. Undertaking a commitment just is doing something that makes it appropriate for that commitment to be attributed. Normative statuses of the sort whose paradigm is provided by the inferentially articulated commitments constitutive of rationality are instituted by constellations of socially perspectival normative attitudes of attributing and undertaking such commitments. This is the *I-thou* structure of norm-instituting social practices that was contrasted above with the *I-we* sociality many theorists appeal to, and which is understood here as arising out of the more primitive perspectival variety.

### 3. Summary

The point of this chapter is to motivate the criteria of adequacy governing the account of discursive practice presented in Chapter 3, as well as the basic raw materials deployed there to satisfy those conditions. The first major point is the *normative* significance of intentional states, meanings, and the sort of understanding that is the uptake of those meanings. The second point is that norms that are *explicit* in the form of rules, principles, or claims (Wittgenstein's "interpretations") depend for their intelligibility—their determining a distinction between performances that are correct and incorrect, appropriate and inappropriate—on a more fundamental form of norms that are *implicit* in practice—in what is *done* rather than what is *said*. Making this distinction raises the question of how to understand the practice of making propositionally explicit claims (formulating principles, promulgating rules, and so on) in terms of norms that are implicit in practices.

The third point is that the attempt to understand norms implicit in practices by identifying the correct/incorrect distinction with the regular/irregular distinction (one strategy for reducing the normative to the nonnormative) will not work, for a reason parallel in form to the one that shows the need for a notion of norms implicit in practice in the first place. The regress-of-rules or regress-of-interpretations argument against *regulism* is that if rules were the only form of norms, they would fail to sort performances into those that are correct according to the rule and those that are not. For applying the rule is itself something that can be done correctly or incorrectly, and any performance that is correct according to one interpretation is incorrect according to others. The gerrymandering argument against *regularism* is that if norms are understood as regularities, they fail to sort performances into those that are correct (regular) and those that are not. Any course of conduct embodies many regularities, and any candidate performance that is regular according to one of them is irregular according to others. The two strategies

do not provide the resources to *privilege* one of the competing interpretations or regularities.

The fourth point, then, is that there is another move available for understanding what it is for norms to be implicit in practices. This is to look not just at what is *done*—the performances that might or might not accord with a norm (be appropriate or inappropriate)—but also at *assessments* of propriety. These are attitudes of taking or treating performances as correct or incorrect. If such attitudes are themselves understood on the model of propositionally explicit beliefs or commitments—as responding to a performance as correct by *saying* of it that it is correct—then the regress objection to regulism about norms reappears. But such assessing attitudes can also be understood as implicit in practice.

One way of doing that is to look to *sanctions*—treating a performance as correct by responding in practice with a reward (or the withholding of punishment) and treating it as incorrect by responding in practice with a punishment (or the withholding of a reward). What counts as a reward or punishment might be construed naturalistically, for instance as any response that positively or negatively reinforces the behavior responded to. Or it might be construed normatively, for instance in terms of the granting of special rights or the assignment of special obligations. Again, the assessing attitudes taken to be relevant to normative statuses can be taken to be implicit in the responses of other individuals, or of responses associated in some way with the whole community. In any of these cases, if the normative status of being a correct performance were identified solely by appeal to *regularities* exhibited by assessments, then the gerrymandering objection would be reinstated.

The fifth point, then, is that one way to demystify norms is to understand them as *instituted* by the practical attitudes of those who acknowledge them in their practice. Apart from such practical acknowledgment—taking or treating performances as correct or incorrect by responding to them as such in practice—performances have natural properties, but not normative properties; they cannot be understood as correct or incorrect without reference to their assessment or acknowledgment as such by those in whose practice the norms are implicit. At the same time, it must be acknowledged that a cardinal criterion of adequacy of any account of the *conceptual* norms implicit in discursive practice is that it make intelligible their *objectivity*. Doing so requires that the normative status of being a correct application of a concept not collapse into normative attitudes, as construing correctness as consisting just in being *taken* to be correct conflates them. The objectivity of conceptual norms requires that any attitude of taking, treating, or assessing as correct an application of a concept in forming a belief or making a claim be coherently conceivable as mistaken, because of how things are with the objects the belief or claim is about.

The next chapter opens the discussion of the propositional contents that are conferred on expressions, performances, attitudes, and statuses by their

playing a suitable role in a system of discursive normative social practices. The idea of normative statuses as *instituted* by practical attitudes, which has been put on the table in this chapter, should be distinguished from the idea of their intentional contents as *confferred* by the social practices in which those statuses and attitudes play a role. As the terms are used here, the institution of status by attitude has to do solely with pragmatics, the study of the practices in which discursive norms are implicit. The conferral of content by practice has to do with the relation between such pragmatics and semantics, which is the study of conceptual contents.<sup>89</sup> The raw materials for a pragmatics that have been assembled here are employed, in Chapter 3, to contribute to both explanatory projects. The next chapter accordingly begins the investigation of concept use and intentional contentfulness.

### Appendix: Wittgenstein's Use of *Regel*

It should be admitted that Wittgenstein's own terminology in some ways obscures the very point he is after in the regress-of-rules argument. For he uses "rule" extremely broadly, to cover much more than is allowed in the usage endorsed here. According to this latter usage, rules are discursively articulated and propositionally contentful; they determine what is correct by describing the correct performances, *saying* what must be true of a performance for it to be correct. By contrast, Wittgenstein uses "rule" (*Regel*) in at least three importantly distinct senses. First is the sense that coincides with the usage preferred here: rules explicitly say what one is to do and are consulted as such by those who follow them—the rule followers' performance is governed by their understanding of the concepts used to characterize what they ought, according to the rule, to do. Second, he sometimes uses "rule" to describe *whatever* guides or is consulted by those whose behavior is being assessed, whether or not it is discursively or conceptually articulated. Finally, he even sometimes talks about following a rule whenever someone's behavior is subject to normative assessment, whenever responsibility to proprieties of conduct is attributed, regardless of whether there is anything the one "following the rule" is aware of or consulting, or being guided by in determining what to do.

In one central text, Wittgenstein offers two senses in which games may be said to be played according to rules: "The rule may be an aid in teaching the game . . . Or it is an instrument of the game itself—Or a rule is employed neither in the teaching nor in the game itself; nor is it set down in a list of rules. One learns the game by watching how others play. But we say that it is played according to such-and-such rules because an observer can read these rules off from the practice of the game—like a natural law governing the play."<sup>90</sup> The first is following a rule in the sense in which that phrase is used

here, according to which one must understand what the rule says and then try to produce performances that the concepts it employs properly apply to. The other, corresponding to the third of the senses distinguished above, is totally external, involving norms that are only in the eye of the beholder, as the remark about natural laws indicates. These are the two senses that Kant distinguishes as acting according to a *conception* of a rule, as agents do, and merely acting according to rules, as inanimate objects do.

In the context of the regress-of-rules argument, this third sense of rule-following, in which it coincides with simple regularity, must be marginal—the question of how to understand a way of grasping a rule that is not an interpretation hardly arises for inanimate objects as they act according to the laws of physics. This is the sense that seems to be involved in the discerning of rules wherever it would be correct to apply ‘same’ or ‘agreement’.<sup>91</sup> Wittgenstein is even willing to appeal to this sense in such outré (according to the usage preferred here) cases of “rule-following” as those involving rules relating pain to pain-expressing behavior.<sup>92</sup> These would seem to be cases in which the rule is entirely in the eye of the beholder, who takes there to be a regularity. Insofar as they are not, these are cases of the second sort, where the performer is being guided by something, but not by something explicit and articulate. It is in this sense that he is willing to call tables of colors and even signposts “expressions of rules.”<sup>93</sup> He seems to call it “following a rule” wherever there is some object whose features it would be appropriate to cite in justifying one’s performance, exhibiting it as appropriate or correct. Though in one place he seems to be careful not to call a map a rule,<sup>94</sup> in others he is even willing to say a line can function as a rule,<sup>95</sup> and a line is clearly not a saying of any sort.

This multiplicity of senses cries out for the distinctions to be explicitly marked terminologically, which is the intent of the relatively more regimented uses of ‘rule’ and ‘practice’ that is employed here. The important point is that there is a way of grasping a rule that is not an interpretation. This should be talked about, as Wittgenstein at least sometimes does, in terms of *practices*—grasping a rule without interpreting it is grasping it in practice, rather than by substituting one expression of a rule for another. Most cases of understanding explicit claims and obeying explicit orders should be understood in this way. Such application of a rule is something that can be done correctly or incorrectly. Practices in this sense are the primitive sort of acknowledgment that performances are governed by norms. But according to this way of regimenting the idiom, not all practices are graspings of rules. There are practices that involve the acknowledgment of norms without involving rules at all, except in the sense that others, looking on, may be able to state rules—whose expressions are not available to the practitioners.

According to this way of using the term, rules are explicit statements that specify what is correct and incorrect by *saying* (describing) it. Obeying an

order (one of the activities Wittgenstein most often links with rule-following in the *Investigations*) thus counts as a kind of rule-following. But one ought not to say that there are rules involved at all in any practice that does not involve claiming, judging, and describing, though of course there are proprieties of practice in more primitive 'games'. That 'practice' is not to be restricted to "ways of grasping rules that are not interpretations" for Wittgenstein seems to be clear from one of the passages cited above: "To obey a rule, to make a report, to give an order, to play a game of chess, are customs (uses, institutions)."<sup>96</sup> Here rule-following is explicitly just one example. Making a report is not following a rule, though it is governed by proprieties of practice, nor, typically, is giving an order.

## Toward an Inferential Semantics

---

Sure, he that made us with such large discourse,  
Looking before and after, gave us not  
That capability and god-like reason  
To fust in us unused.

SHAKESPEARE, Hamlet

---

### I. CONTENT AND REPRESENTATION

#### 1. *Intentionality: Propositional and Object-Representing Contentfulness*

Taking or treating someone as one of us may be called *recognizing* that individual.<sup>1</sup> According to the construal of recognition being developed here, taking or treating as one of us is adopting a certain kind of intentional stance. It requires first of all interpreting the one it is directed at as the subject of intentional states. But taking someone as one of us also requires, it was suggested, interpreting that individual as an intentional interpreter—as able to *attribute* intentional states, and so as able to adopt toward others just the same sort of attitudes out of which that very stance is constructed. The previous chapter assembled some raw materials for an account of the normative significance of the intentional states we attribute to each other—and take each other to attribute to each other—in adopting the attitudes of mutual recognition that institute the status of community membership, of being one of us.

Before such an account is presented, in the next chapter, it is necessary to look more closely at the sort of *content* that sets apart—as distinctively *intentional*—the states and statuses (and therefore the attitudes) that are attributed when we recognize someone. For intentional states are intention-

ally contentful states, and the theoretical job of the contents they are taken to have is precisely to determine, in context, the particular significance of being in or attributing the states those contents are associated with. As the terms are used here, *semantics* is the study of such contents, and *pragmatics* is the study of the force or significance of the states, attitudes, and performances that have those contents. Accordingly, to fill in the details of a story about the normative character of the *pragmatic significance* of intentionally contentful states, attitudes, and performances, an inquiry into the nature of their *semantic contents* is called for.

Brentano, who brought the term 'intentionality' back into modern usage, defines it this way: "Every mental phenomenon is characterized by what the scholastics of the Middle Ages called the intentional (also mental) inexistence of an object, and what we could call, although not in entirely unambiguous terms, the reference to a content, a direction upon an object (by which we are not to understand a reality in this case), or an immanent objectivity. Each one includes something as object within itself, although not always in the same way. In presentation, something is presented, in judgment something is affirmed or denied, in love loved, in hate hated, in desire desired, etc."<sup>2</sup> "Not always in the same way" indeed. "Intentional object" as used here involves assimilations along two dimensions. First is a dimension most clearly picked out in terms of grammatical categories: what is affirmed or denied in a judgment is something expressible by the use of a declarative sentence, while what is loved or hated is something referred to by the use of a singular term. The former may be called 'propositional contentfulness'. It is typically expressed by the use of a declarative sentence and is ascribed by the use of a 'that' clause appended to a specification of the contentful state or attitude, as in "the belief *that* Carlyle wrote *Sartor Resartus*" or "desiring *that* Pufendorf's reputation be rehabilitated." The latter may be called 'object-representing contentfulness'. It is typically expressed implicitly by the use of a singular term as a grammatical direct or indirect object, and it is attributed explicitly by using terms such as 'of' or 'about', as in "a belief *about* Carlyle" or "desiring something *of* Pufendorf's reputation" (for example that it be rehabilitated).

Putting these two sorts of contentfulness in a box together is not just an idiosyncrasy of Brentano's. Searle, for instance, offers this pretheoretical delineation of the subject matter of his book *Intentionality*: "If a state *S* is Intentional then there must be an answer to such questions as: What is *S* about? What is *S* of? What is it an *S* that?"<sup>3</sup> To insist on distinguishing these sorts of content in the way indicated above is not yet to diagnose a confusion in remarks like this. There is no confusion insofar as propositional and object-representing contentfulness ought to be understood as species of a genus.

An approach to the characterization of that genus is not far to seek. Stalnaker speaks for the dominant tradition in offering this formulation:

"The problem of intentionality is a problem about the nature of representation. Some things in the world—for example pictures, names, maps, utterances, certain mental states—represent or stand for, or are about other things—for example people, towns, states of affairs."<sup>4</sup> The genus, it is suggested, is *representational* content. Indeed, Stalnaker, like others, is comfortable talking interchangeably about "intentional or representational states." This basic insight should be accepted to this extent: it is clear that intentionality has a representational dimension and that to understand intentional contentfulness one must understand representation.

A common response to this insight is to envisage an explanatory strategy that starts with an understanding of representation and on that basis explains the practical proprieties that govern language use and rational action. It is not clear, however, that a suitable notion of representation can be made available in advance of thinking about the correct use of linguistic expressions and the role of intentional states in making behavior intelligible. The temptation to think otherwise is connived at by insufficient appreciation of some of the fundamental criteria of adequacy to which an account of the representational dimension of intentional contentfulness must answer. It is important to keep in mind the explanatory challenges faced by a semantic theory that appeals to representation as its basic concept, and some of the ways in which those explanatory obligations are liable to be unobtrusively shirked. To point these out is not to show that they cannot be satisfied—that representational explanatory strategies are in principle broken-backed. It is merely to guard against the danger that such an explanatory starting point may recommend itself in virtue of its apparent immunity to difficulties it has not squarely confronted.

A particularly unhelpful way of pursuing the representational semantic explanatory strategy is to model representation on *designation*. The designational model is objectionable on two grounds connected with the distinction of grammatical category between sentences and subsentential expressions such as singular terms. First, it assumes that the relation between a singular term and the object it picks out or refers to, for instance that between a name and its bearer, is antecedently intelligible—that the notion of tagging or labeling something can be made sense of before one considers the use of such tags or labels in *saying* something (paradigmatically, in making a claim). In this way, the strategy runs afoul of the principle of the pragmatic priority of the propositional, which is discussed further along.

Second, it assumes that the notion of representation as reference picked out in this way for the category of singular terms and predicates can be univocally and unproblematically extended to apply to the category of sentences. Sentences are understood as representing states of affairs, in the same sense that singular terms represent objects (and in the same sense that predicates represent properties or sets of objects). The notion of representation, conceived as designation, is then supposed to make the grammati-

cal distinction between singular terms and sentences intelligible by appealing to the ontological distinction between objects and states of affairs. Even if it is granted that there is a clear sense in which singular terms such as names and marks on maps represent particular objects, for instance individual people and cities, it does not follow that it is possible to introduce the category of states of affairs as what is in the same sense represented by declarative sentences and 'that' clauses. Nor ought it to be assumed that the ontological category of states of affairs can be made intelligible apart from and in advance of explaining the use of declarative sentences and the 'that' clauses used to report such uses in indirect discourse.

## 2. Two Senses of 'Represents'

Introducing the notion of states of affairs as the kind of thing represented by declarative sentences requires sensitivity to the second dimension of assimilation involved in Brentano's idiom. For one must be careful not to confuse what is *represented* by sentences with what is *expressed* by them. This is a familiar point, but it is worth emphasizing. As Brentano acknowledges by appending to his phrase "direction upon an object" the qualification "by which we are not to understand a reality in this case," 'represent' is ambiguous between two intimately related but importantly distinct senses. Searle puts the point this way: "'About' . . . has both an extensional and an intensional-with-an-s reading. In one sense (the intensional-with-an-s), the statement or belief that the King of France is bald is about the King of France, but in that sense it does not follow that there is some object which they are about. In another sense (the extensional) there is no object which they are about because there is no King of France. On my account it is crucial to distinguish between the content of a belief (i.e. a proposition) and the objects of belief (i.e. ordinary objects)."<sup>5</sup> Thus as Searle sets things up, for a statement or belief to have content is for it to represent or be about something in the 'intensional' sense, while for it to have an object or objects is for it to represent something in the 'extensional' sense.<sup>6</sup> The relation between the two senses emerges more clearly if one or the other is taken as primary and the remaining one specified in terms of it. Thus if 'represent' is reserved for the sense in which one can represent only what in fact exists, whether it be in the category of objects or of states of affairs—actual objects corresponding to singular terms and actual states of affairs corresponding to true claims—then the other sense can be picked out as *purporting* to represent. The other way to do things would be to use 'represent' even in the cases where nothing exists to be represented, where there need be no object or state of affairs as represented for there nonetheless to be a representing. When something does exist as represented, the representation might be called *successful* or *correct*.

An account of contentfulness in terms of representation needs to explain

both of these senses in which something can be a representing, and it needs to explain their relation to one another. It is clear that if contentfulness in general is to be identified with representational contentfulness, that is, with being a representing, then 'representing' should be understood as purported representing in a sense that contrasts with successful representing. For it makes sense to wonder whether, or to believe that, there is a present king of France or Schelling was the greatest German philosopher, even if it turns out that no object or state of affairs corresponds to that contentful state. A theoretical idiom that shrinks the scope of purported representing until it coincides with that of successful representing has no room for the notion of error, of representation that is incorrect or mistaken; and a notion of representation so thin as to preclude assessments of correctness provides no basis for any recognizable concept of intentional content.

A theoretical idiom that, on the contrary, expands the scope of successful representing until it coincides with that of purported representing is equally unpromising. The result of holding purported representing fixed and failing to distinguish successful representing from it is Meinongianism—commitment to a vast realm of entities, most of which do not exist, including many that *could* not exist. The trouble with taking it that there is something that is successfully represented by every purported representing is not just that it involves commitment to a luxuriant ontology; ontological self-indulgence is a comparatively harmless vice. But it can be symptomatic of a failure to shoulder an explanatory burden. In this case it evidently (and ultimately unhelpfully) transforms the demand for an account of the relation between correct and incorrect, unfulfilled or merely purported and actually successful representing, into a demand for an account of the relation between the statuses of what is represented in the two cases: between mere subsistence and robust existence. Ontological postulation can no more provide an explanation by itself in this case than it could in the one just considered, where the issue was an account of the relation between the sense in which singular terms are representationally contentful and the sense in which sentences are. (Of course, no more in this case than in that one does a commitment to taking the representational dimension of intentional content seriously entail going on to make such a mistake; it is important to recognize the temptation in order to resist it.)

Brentano, who did not make the mistake of his student Meinong, indicates some of the difficulties faced by such an attempt to ontologize the distinction between correct and incorrect representation by holding to a univocal sense of 'represents' and construing the distinction as a difference between two different sorts of representable: "It would be paradoxical to the highest degree to suppose that you could promise to marry an *ens rationis* and then keep the promise by marrying an actual, concrete particular."<sup>7</sup> It is disastrous to put the notion of successful representing in place of that of purporting to represent, that is, to have it play the role of necessary condition for content-

fulness. But while the two senses of ‘represent’ or ‘about’ must not be run together (from either direction), there is also reason not to want them to be driven too far apart. Purporting to represent is intelligible only as purporting to represent *successfully* or *correctly*. If what would make the representings successful has no part to play in determining the purport or content of those representings, it is hard to see how assessments of correctness could even get a grip on them. The trouble then is not just that of skepticism about justification, in the Cartesian mode. If all our ideas could have just the content-as-representational-purport that they do, even though the rest of the world, the representeds those ideas purport to represent successfully, were entirely different from what it is represented (purported) to be, how could we ever be justified in taking ourselves to be correct? The difficulty that looms is more serious still, threatening not just the cogency but even the comprehensibility of the picture of states and attitudes as contentful in virtue of their representing or being about the way things are. For the very notion of representings so much as purporting to be about representeds becomes unintelligible.

Acknowledging this distinction between representational purport and representational success is one of the theoretical jobs Frege assigns to his paired semantic concepts *Sinn* and *Bedeutung*. A sign is contentful insofar as it *expresses* a sense. A thought is the sense, the propositional content expressed by a declarative sentence. To say that it is true or false—to assess it along the dimension of correctness semantically relevant to thoughts—is to classify it in terms of the result of applying a function to objects serving as the arguments of that function, where both the function and the objects are picked out as those referred to by components of that sense. The structure of the later Frege’s semantic project accordingly encompasses accounts both of what it is to express a sense and of what it is for that sense to be *correct* in terms of how things are with what it represents. An utterance or inscription *expresses* a sense, for example a thought, and it is the sense expressed that then *refers* to objects, the thought that represents them as instantiating properties and standing in relations.<sup>8</sup> This idiom avoids the dangerous ambiguity inherent in talking about propositions as *represented* by sentences. For that way of talking is liable to be misunderstood as involving the identification of propositions with the facts or states of affairs successfully represented by true claims (according to the representational model of contentfulness) rather than with the claims or purported representations expressed by sentences.

### 3. *Representational Uptake*

The notion of representational *purport* implicitly involves a notion of representational *uptake* on the part of some consumer or target of the purporting. It is only insofar as something can be *taken* to be a representation

that it can purport to be one. For purporting to be something is putting oneself forward as aptly or appropriately taken to be that. The purport is veridical or spurious (for instance the representation is successful or misleading) accordingly as the taking it invites is correct or incorrect. That grasp of something as a representation is coordinate with representational purport is the point Dennett is making when he says: "Something is a representation only for or to someone; any representation or system of representations thus requires at least one user or interpreter of the representation"<sup>9</sup> (using "representation" to mean purported, not necessarily successful, representation). It was pointed out in the previous paragraph that according to the representational model of contentfulness being considered, representational purport is what is *expressed* by a representing, for instance a sign design, rather than what is *represented* by it if it is successful. The present point is then that talk of what is expressed is intelligible only in the context of talk of the activity of grasping what is expressed. By widening the focus a bit, this can be seen to be the manifestation (within the representational construal of contentfulness) of the general point that *meaning* and *understanding* are coordinate concepts. The notion of representational purport is one way of rendering what must be understood in grasping the content of an intentional state, attitude, or performance.<sup>10</sup> Representational purport and the understanding that is its uptake must both be explained in order to make an account of intentional contentfulness in terms of representation work. As Kant says: "The understanding, as a faculty of knowledge which is meant to refer to objects, requires quite as much an explanation as the possibility of such a reference."<sup>11</sup>

Looking back from the vantage point won for us by the later Wittgenstein, it is possible to see that one of the unfortunate emphases that Descartes imposed on the representationalist tradition is the privileging of *knowledge*, and therefore *successful* representation, as a topic of inquiry, over *understanding*, and therefore *purported* representation. For Descartes, representational purport, being "as if of" something, is an intrinsic and characteristic property of *pensées* (that is, specifically mental acts). He does not offer an account of what it is for a mind to grasp such purport, for it to take or treat an idea as being of or about something. He is concerned with how one might become entitled to a commitment to something that has objective (in his, neo-Scholastic sense) reality in one's thoughts having also formal reality outside them. He is not concerned with what the mind's taking one thing or sort of thing rather than another (or rather than nothing at all) as having objective reality in one's thoughts itself consists in. Representational purport, "the objective reality of things in thoughts," and its corresponding uptake by the mind whose thoughts they are serve Descartes as unexplained explainers. So the content of the representational commitments to which the mind's entitlement is at issue is never clarified. A representational model of contentfulness cannot rest with an account of

successful representation—not even if it is accompanied by a vindication of the right to believe that purported representation is often or even generally successful. It requires also an account of representational purport, and that requires an account of the uptake, grasp, or understanding of such purport.

It would of course be a blunder, of a familiar kind, to understand that uptake in general as consisting in *interpreting* something as a representation, in Wittgenstein's sense of 'interpreting'. Taking something as a representation must not be parsed in terms of the adoption of explicitly contentful attitudes or intentional states such as belief. If being a consumer of representational purport, taking something as a representation of something, is understood as believing of it that it correctly represents (or equally if the purport is understood as intending that it do so), then an infinite explanatory regress is generated by the possibility of querying the nature of the representational purport ('that . . .') and success ('of . . .') such a belief exhibits. There must be some way of understanding something as a representation that consists not in interpreting it (in terms of something else understood as a representation) but in taking, treating, or using it in practice as a representation. To understand what it is for red dots on a map to purport to represent cities and wavy blue lines to purport to represent rivers, the theorist must look to the practice of using a map to navigate. If such purport is to provide a model applicable to representational purport in general, that practice must admit of construals that do not appeal to the formation of propositionally contentful beliefs. The practice must be intelligible in terms of what counts as following it or going against it in what one actually does: the way it guides the behavior of those who can use maps.

The absence of a nonregressive account of what it is to take, treat, or use something as a representation of something else is the source of another traditional sort of dissatisfaction with the representationalist paradigm of contentfulness. It lies behind Rebecca West's irritated response to the "mind as the mirror of nature" model that it is hard to see why one would want a copy of the universe: "One of the damn things is enough." Progress in understanding intentional contentfulness is made by invoking representational relations only in the context of an explanation of what it is that makes representings graspable or intelligible as representings in a way in which what is represented is not. That is a matter of the uptake or consumption of representational contentfulness. Apart from the representational purport it expresses, and which is there to be grasped, a representing is just another bit of worldly furniture, like what it represents. Why is not confronting a map as well as terrain just adding one more thing to be baffled about? Invoking a relation (for instance some sort of isomorphism) between representing and represented does not by itself contribute to the task of explaining what the intelligibility of the representing consists in—why one of the damn things is not enough.

#### 4. Expression and Representation

Restricting attention for the moment to the propositional contents characteristic of intentional states such as belief, it has been suggested that it is no use asking what a proposition (or propositional content) is, without asking what it is for a sentence to *express* a proposition, or for a state to have one as its content. Just so it is no use asking what it is for a proposition to be *true*, or a representation to be *successful* or *correct*, without asking what it is to express one—what *purporting* to represent consists in. And it is no use asking what it is to *express* a proposition or other content (to *purport* to represent), without asking what it is to *grasp* or *understand* such purport. An account is required not only of how representings are distinguished from and related to representeds (in successful representing). An account is needed also of the representational content representings express—their representational purport. And that requires an account of the *attitude* of grasping such purport: of taking, treating, or using a representing as a representing, of acknowledging or attributing to it in practice its representational purport.

The treatment of representational content in upcoming chapters centers on an account of this practical attitude. Becoming entitled to use a concept of intentional content involves a twofold explanatory task: to say what it is to *express* a *propositional* content in general, and then to say what more is required specifically for the content expressed to *represent* something *objective*, in the way that matters for empirical science. Furthermore, each of these must include an account of what those who exhibit and attribute states, attitudes, and performances with such contents must do in order to count as taking or treating them in practice as contentful in those ways.

This is a request that can sensibly be addressed to Wittgenstein, as well. Even his sustained, penetrating discussions do not offer an account of what distinguishes language games within which states and performances acquire specifically *propositional* significances (the only ones that, by the lights of this work, deserve the title ‘Sprachspiel’), nor of what distinguishes those within which states and performances acquire specifically *representational* significances. He argues against understanding the contents determining the significances of all states and performances in terms of representational content. For one ought not simply to presume on syntactic grounds that terms are used to refer (or fail to refer), predicates are used to describe or characterize (or misdescribe), and sentences are used to claim (truly or falsely). Instead of asking what object is being referred to by the term, what property is being ascribed by the predicate, and what fact would make the sentence true, one ought first to look at the *use* of the expressions, to see if the putative referring, characterizing, and claiming in fact play a practical role that is best understood in terms of such contents. For many expressions

that might have been thought to be doing the jobs just mentioned ('sensation', 'intending', 'beetle', 'must', 'true', 'I am in pain', . . .), consideration of their use shows that another account of the content of the putative referrings, characterizings, and claimings is more appropriate.

The idea that not all contentful expressions play a straightforwardly representational role is a development of a line of thought that is already important in the *Tractatus*. Some previous varieties of logical atomism had distinguished themselves by their insistence that the only way any expression, sentential or not, could have content or contribute to the content of an expression of which it is a part is by standing for or representing something.<sup>12</sup> Thus, not only did these views grasp the nettle of commitment to negative and conditional facts, they also were committed to 'not' and 'if . . . then . . .' standing for some element in a complex state of affairs. The undertakers of such commitments are admirable more for their conceptual heroism than for their good sense.

In the *Tractatus*, Wittgenstein showed that one could best treat logically compound propositional contents as representing states of affairs by not treating every contentful expression (every one whose occurrence is significant for determining the state of affairs represented by the whole) as itself having its content in a representational way, by standing for something. Purely formal vocabulary, paradigmatically logical vocabulary, is contentful but does not itself stand for anything. (Kant and Frege had of course earlier shown the possibility of this sort of approach.) The opening sections of the *Investigations* argue along just these lines: not every piece of a representation contributes to its content by itself representing, and not every move in a language game is a representing of something. But Wittgenstein does not explain what one must do to be using an expression to refer, characterize, or claim (the features of use he associates with representational content), nor does he explain what is required for something caught up in a language game to express a specifically propositional content.

The notion of expression—of making propositionally explicit—shows up at two different levels in what follows. First, one who adopts the pragmatist's approach to intentionality owes an account of the practices that ultimately confer explicitly *propositional* content on the states, attitudes, and performances that play appropriate roles in those practices. This is an account of the implicitly normative practices in virtue of which anything at all can be made explicit as the content of a possible claim or belief *that p*. Such a theory should explain what it is for a performance, paradigmatically but not exclusively the tokening (by speaking or writing) of some linguistic item, to express an intentional content. And it should explain the relation between such expressions and the possession of content by states or attitudes somehow related to them. That is, it should explain what it is to express or exhibit a specifically *propositional* content—intuitively, one that could be true or false. Furthermore, it should explain what it is to express or exhibit a content

that purports to *represent* something, and it should explain the relation between representing states of affairs and representing particular objects. As has been pointed out, a necessary part of explaining the expression of contents with representational purport is explaining the *grasping* of such contents, the uptake that is the other side of such purport.

The second level concerns not the making of ordinary claims but the formulation of rules or principles. The regress-of-interpretations argument shows that the intellectualist tradition erred in treating the explicit form of norms as fundamental. But once a notion of propositional explicitness has been brought onboard in terms respectable according to pragmatist scruples, the fact that the contentful norms implicit in practical doings can be expressed in rules, claims, and interpretations that say or state explicitly what is implicit in those practical proprieties itself still stands in need of explanation. An account is needed of what it is to make explicit in the form of something that can be *said* or *thought* what is otherwise merely implicit in what is *done*. At this level, the implicit proprieties of practice that make it possible to make propositionally explicit claims are themselves made propositionally explicit in the form of rules or principles. A theory of *expression* accordingly is to explain how what is *explicit* arises out of what is *implicit*. In the first instance, it must explain how propositional content (the form of the explicit) is conferred by norms that are implicit in discursive practice—that is, what proprieties of use having such a content consist in. Then it must show how those same implicit, content-conferring norms can themselves be made explicit in the form of rules or principles.

### 5. From Practice to Content

These two challenges are addressed in the rest of the work. First, what role must states, attitudes, and performances play in (as it turns out, social) practice for it to be correct to interpret them as being *propositionally* contentful? That is, how are propositional contents conferred by practice? What proprieties of practical employment does possession of such content consist in? As already suggested, any answer must specify what it is for the practitioners themselves practically to take or treat states, attitudes, and performances of others and of their own as having such contents, and thereby to confer those contents on them. Chapters 3 and 4 develop a response to these questions. Second, what must be true of such contentful states, attitudes, and performances for it to be correct to interpret them as *representing* objects and objective states of affairs? Again the answer must specify what it is for the practitioners themselves in practice to take or treat those states, attitudes, and performances as having such contents, and so by their practice to confer such contents on them. Part 2—particularly in Chapter 6 (on the representation of objects by singular terms) and Chapter 8 (on objective

representation)—presents an account of these phenomena, within the framework introduced in Chapter 3.

The practical uptake of specifically representational purport must include normative assessment of states, performances, and expressions—assessment of their specifically representational correctness. (Of course, on pain of the familiar regress, such assessment must not be understood as in every case consisting in judging *that* a representation is correct; besides such propositionally explicit attitudes there must be practically implicit ones.) Treating something as a representation involves acknowledging the possibility that it *misrepresents*—that the representational taking is a *mistaking* (the object represented does not exist, the state of affairs represented does not obtain). It is these attitudes of distinguishing in practice between representations that are taken to be correct and those taken to be incorrect that forge the connection between the notions of representational purport and representational success.

Practical representational uptake of representings—treating objects, states, or performances as purporting to be correct representations of objects and facts—consists in taking them to be takings: taking them to express attitudes concerning what there is and how things are. That they are accordingly essentially liable for assessment as to their representational success (that they in a characteristic way answer to how things actually are for their correctness) means that such uptake incorporates an implicit distinction between representational *attitude* (how things are taken to be by what is treated as a representation) and representational *status* (how things actually are, which determines the success or correctness of that attitude). Thus the normative pragmatic distinction between status and attitude is central to the intelligibility of fundamental semantic concepts. It is reflected in the distinction between representational purport and representational success.

The *objectivity* of representational content is a feature of the practices of assessing the correctness of representations. The status of representings as correct or incorrect, successful or unsuccessful, depends on how things are with what is represented, rather than on the attitudes of representers. What is distinctive of specifically *representational* correctness is this objectivity—the way in which assessments of representational correctness take representings to answer to what is represented, rather than to how what is represented is *taken* to be. It is the way in which the status being assessed outruns any particular attitude toward it. Understanding the objectivity of representational content requires understanding this particular structure of authority and its acknowledgment—what it is for those assessing the correctness of representings to cede authority over them to what is represented, to treat their correctness in practice as determined by those representeds. Again, one lesson is that the representational dimension of semantic content cannot be understood apart from the normative pragmatic context in which it is embedded and in which it is accorded its characteristic significance.

It should be clear that the remarks in this section are not meant to have the force of arguments against treating representation as a central semantic category. Rather, they present some general criteria of adequacy for an account of this important semantic notion. By doing so, however, they do offer reasons not to treat representation as a semantic *primitive*, as an unexplained explainer. The next section shows why the role of semantic concepts in pragmatics (the proper use of language and the appropriate role of intentional states in rational action) dictates approaching semantics in the first instance through the notion of *propositional* contentfulness. The rest of the chapter then motivates an approach to propositional contentfulness that begins with the *inferential* articulation of the social practice of giving and asking for reasons. The following chapter presents a particular model of those social practices (in terms of deontic scorekeeping) and shows how they can be understood as at once instituting discursive commitments and conferring propositional contents on them. In Part 2, that framework is extended to include *representational* content, both of the sort expressed by sentences and that expressed by subsentential expressions. It concludes with a discussion of the social and inferential articulation of discursive practice in virtue of which the contents it confers are properly understood as involving an *objective* representational dimension.

## II. THE PRIORITY OF THE PROPOSITIONAL

### 1. Kant on Judgment as the Form of Awareness

It is appropriate to begin by addressing propositional contents because of what can be called the *pragmatic priority of the propositional*. The pre-Kantian tradition took it for granted that the proper order of semantic explanation begins with a doctrine of *concepts* or *terms*, divided into singular and general, whose meaningfulness can be grasped independently of and prior to the meaningfulness of judgments. Appealing to this basic level of interpretation, a doctrine of *judgments* then explains the combination of concepts into judgments, and how the correctness of the resulting judgments depends on what is combined and how. Appealing to this derived interpretation of judgments, a doctrine of *consequences* finally explains the combination of judgments into inferences, and how the correctness of inferences depends on what is combined and how.

Kant rejects this. One of his cardinal innovations is the claim that the fundamental unit of awareness or cognition, the minimum graspable, is the *judgment*. "As all acts of the understanding can be reduced to judgments, the understanding may be defined as the faculty of judging."<sup>13</sup> For him, interpretations of something as classified or classifier make sense only as remarks about its role in judgment. A concept just is a predicate of a possible judgment,<sup>14</sup> which is why "the only use which the understanding can make of

concepts is to form judgments by them.”<sup>15</sup> Thus for Kant, any discussion of content must start with the contents of judgments, since anything else only has content insofar as it contributes to the contents of judgments. This is why his transcendental logic can investigate the presuppositions of content-fitness in terms of the categories, that is, the “functions of unity in judgment.”<sup>16</sup>

The understanding is the active cognitive faculty, the faculty of spontaneity—understanding is something *we do*. “We have before given various definitions of the understanding, by calling it the spontaneity of knowledge (as opposed to the receptivity of the senses), or the faculty of thinking, or the faculty of concepts or of judgments; all of these explanations, if more closely examined, coming to the same.”<sup>17</sup> What we do is synthesize, bring things into a unity—that is, subject them to rules or concepts. What we do, as opposed to what happens to us, is to judge. Although synthesis happens at other levels than that of judgment (there is synthesis in intuition and imagination also), that synthesizing activity is an aspect of judging. “The same function which imparts unity to various representations in one judgment imparts unity likewise to the mere synthesis of various representations in one intuition, which in a general way may be called the pure concept of the understanding. The same understanding, and by the same operations by which in concepts it achieves through analytical unity the logical form of a judgment, introduces also, through the synthetical unity of the manifold in intuition, a transcendental element into its representations.”<sup>18</sup> Thus all our cognitive activity consists of judgment and aspects of that activity. Any content that can be discerned in any category is derivative from the content of possible judgments, that is, from propositional content. Kant’s *pragmatics*, or theory of cognitive *activity*, determines the fundamental unit of his *semantics*, or theory of the *contents* of cognitions.

## 2. Frege and Wittgenstein

This insight into the fundamental character of judgment and so of judgeable contents is lost sight of by Kant’s successors (indeed it could be argued that appreciation of it is still missing from such broadly semantic traditions as semiotics and structuralism). It is next taken up by Frege. Looking back over his lifework in 1919, he picks out this point as basic to his orientation: “What is distinctive about my conception of logic is that I begin by giving pride of place to the content of the word ‘true’, and then immediately go on to introduce a thought as that to which the question ‘Is it true?’ is in principle applicable. So I do not begin with concepts and put them together to form a thought or judgment: I come by the parts of a thought by analysis [*Zerfällung*] of the thought.”<sup>19</sup> Already in 1870 in the *Begriffsschrift*, Frege introduces “contents of possible judgment” or “judgeable contents” in the second paragraph and subsequently defines other sorts

of contents in terms of them. In an essay explaining the *Begriffsschrift* he summarizes this approach: "I start out from judgments and their contents, and not from concepts . . . instead of putting a judgment together out of an individual as subject and an already previously formed concept as predicate, we do the opposite and arrive at a concept by splitting up the content of a possible judgment."<sup>20</sup> The concept of a function, which stands at the center of Frege's technical contribution to semantics, is introduced in the *Begriffsschrift* as an element in his substitutional methodology for decomposing contents of possible judgment.<sup>21</sup> In the *Grundlagen* Frege continues to follow this Kantian line in insisting that "we ought always to keep before our eyes a complete proposition. Only in a proposition have the words really a meaning . . . It is enough if the proposition taken as a whole has a sense; it is this that confers on its parts also their content."<sup>22</sup> Frege holds this view because of the importance he assigns to the concept of *truth*; to talk about an expression as contentful is to talk about the contribution it makes to the truth-value of thoughts or propositions in which it occurs.

It is sometimes thought that Frege gave up his commitment to the primacy of the propositional by the late 1880s, when he began to assimilate sentences technically to singular terms under the heading *Eigennamen*, which includes everything except functional expressions. Such a view overlooks the very special role that sentences, as 'names' of truth-values, continue to play for him, even in the *Grundgesetze*. The importance of truth, and therefore of thoughts (the contents expressed by declarative sentences), continues to be emphasized at every stage in Frege's development. In his long 1914 essay entitled "Logic in Mathematics," he is still maintaining "that the name should designate something matters to us if and only if we are concerned with truth."<sup>23</sup> This is the same view that he had endorsed in his classic essay "Über Sinn und Bedeutung": "But now why do we want a proper name to have not only a sense, but also a reference [*Bedeutung*]? Why is the thought not enough for us? Because, and to the extent that, we are concerned with its truth value . . . It is the striving for truth that drives us always to advance from the sense to the reference."<sup>24</sup> In the context of such a view it is clear that the assimilation of sentences to singular terms as both having objects as *Bedeutungen* can in no way undercut the fundamental role played by truth-values, and so by the propositional contents that bear them. In that same essay he says that what is needed for a name to have content (express a sense) is that it "belong to a sufficiently complete totality of signs."<sup>25</sup> Given his views about identity, this means a system of signs that includes sentences in which the name occurs, and also further sentences that result from them by substituting other names for the ones in question. The totality of signs must include sentences, because to have a sense is to purport to have a *Bedeutung*, and as just indicated, such purport arises only in the context of concern with truth, because "anyone who seriously took the sentence to be true or false would ascribe to the name . . . a *Bedeutung*".<sup>26</sup> It is because the

point of deploying concepts in thought and talk is to judge, that is, take or treat judgeable contents as true, that such contents are given pride of place in Frege's scheme. As he says in the 1897 fragment on logic: "Every act of cognition is realized in judgments."<sup>27</sup>

Indeed, it can be misleading to focus on the concept of *truth* as what enforces attention to sentences. Frege takes this position because it is only to the utterance of sentences that pragmatic force attaches, and the explanatory purpose of associating semantic content with expressions is to provide a systematic account of such force. "'True' only makes an abortive attempt to indicate the essence of logic, since what logic is really concerned with is not contained in the word 'true' at all but in the assertoric force with which a sentence is uttered . . . the thing that indicates most clearly the essence of logic is the assertoric force with which a sentence is uttered."<sup>28</sup> Talk about the cardinal importance of concern with truth is a dispensable *façon de parler*. What actually matters is the pragmatic attitude of *taking*-true or putting forward *as* true, that is, judging or asserting. Semantic vocabulary is used merely as a convenient way of making explicit what is already implicit in the force or significance that attaches to the content of a speech act or attitude. (An account of just how this explication works is offered in Chapter 5, where specifically semantic vocabulary, paradigmatically 'true' and 'refers', is discussed.)

The point that the contents expressed by sentences must play a privileged explanatory role because it is to sentences that pragmatic force attaches has been brought home most forcefully by the later Wittgenstein. The use of sentences is prior in the order of explanation to the use of subsentential expressions because sentences are the only expressions whose utterance "makes a move in the language game." Sentences are expressions whose unembedded utterance performs a speech act such as making a claim, asking a question, or giving a command. That is why even when such a speech act is performed by an utterance that does not manifest the syntactic complexity typical of sentences (a shout of "Rabbit!" or "Fire!" for instance), the utterance should nonetheless be interpreted as a one-word sentence, as meaning what we might express by "Look at the rabbit!" or "There is a fire!"

Referring to something, indicating or naming it, is also something one can do with linguistic expressions; it is a speech act one can perform. But these belong to a class of speech acts that is in an important sense derivative from or parasitic on speech acts involving sentences, paradigmatically claiming, asserting, or putting forward *as* true. In order to use an expression as a name, to refer to or pick out an object with it, one must be able to use the name to say something (paradigmatically, to assert something) about the object referred to, indicated, or named. The significance of taking or treating something as a name, as purporting to refer to an object, consists in how one takes it to be proper to use the expression, and the use of expressions as names is unintelligible except in the context of using expressions containing them as sentences.<sup>29</sup>

### 3. Semantics Must Answer to Pragmatics

The primacy of propositional intentional contents also shows up if one considers cases in which the use of language is not to the fore. Intentional interpretation of nonlinguistic organisms—intentional explanation of their behavior by attributing beliefs and desires that make what they do intelligible—also depends on attributing *propositionally* contentful states, attitudes, and performances. Behavior is made intelligible by exhibiting it as rational, given various beliefs and pro-attitudes, and to do that is to exhibit a piece of practical reasoning that is taken somehow to stand behind or be implicit in the behavior. The imputed reasoning shows why an organism with the states or attitudes that provide the premises *ought*, rationally, to behave in the way specified by the conclusion. But *what can serve as a premise in reasoning must have a propositional content*. This point is so important to the present project that the rest of this chapter is devoted to motivating the treatment of this feature, in the next chapter, as a defining characteristic distinctive of the propositional. The intentional interpreter attributes to the cat the belief *that* there is a mouse around the corner from it, and the desire *that* it catch the mouse, and so on. Attributing intentional states so as to render behavior intelligible in the light of them requires attributing propositional contents to them.<sup>30</sup> So propositional contents have a pragmatic priority, not only in the setting of assessments of the significance of speech acts, but also in the setting of attributions of intentional states that do not evidently depend on linguistic practices.

Semantics must answer to pragmatics. The theoretical point of attributing semantic content to intentional states, attitudes, and performances is to determine the pragmatic significance of their occurrence in various contexts. This means settling how linguistic expressions of those contents are properly or correctly used, under what circumstances it is appropriate to acquire states and attitudes with those contents, and how one then ought or is obliged to go on to behave. It is specifically *propositional* contents that determine these pragmatic significances, so it is specifically propositional contents that it is the task of semantic explanatory theories to attribute. Semantic contents corresponding to *subsentential* expressions are significant only insofar as they contribute to the determination of the sorts of semantic contents expressed by full sentences. The pragmatic priority of sentence-use to name-use enforces a certain semantic explanatory priority of the contents expressed by sentences to those expressed by names. The task of the next chapter is to develop an account of the practices of using expressions as sentences—paradigmatically to make claims and so to confer specifically propositional contents on those expression uses and on the states and attitudes associated with them (to use them as having such contents).

What the theorist associates with states and expressions deserve to count as *semantic* contents only insofar as they play the right sort of role in determining the proprieties of practice governing those states and expres-

sions. It is possible to associate all sorts of abstract objects with strings of symbols in formalized languages, from sets of models to Gödel numbers. Such an association amounts to specifically *semantic* interpretation just insofar as it serves to determine how those strings are correctly used. For example, Tarski's mapping of well-formed formulas of the first-order predicate calculus onto topological domains qualifies as a semantic interpretation of them only because he can derive from it a notion of valid inference, a way of telling what follows from what—that is, a notion of their correct use. Apart from that, it would just be one more algebraic homomorphism.

#### 4. Two Mistakes the Designational Model Invites

An account of content in terms of representation must satisfy the requirement that it must show how semantic content so construed matters for the pragmatic significance of what it is associated with. For the reasons indicated above, this demand focuses attention to begin with on the representational rendering of specifically propositional contents. Two difficulties arise at this point: it is not clear how to derive a notion of propositional contentfulness from the designational representational model, and construing content in representational terms requires supplementation by a further story to get to the proper use of contentful expressions and the correct circumstances and consequences of being in contentful states. (In contrast, the explanatory strategy pursued in Chapter 3 begins with an account of the practices within which producing a performance or altering an attitude can have the pragmatic force or significance of making a claim or judgment; the notion of propositional contentfulness is then understood as what is expressed by such acts.)

On the first point, the pre-Kantian representationalist tradition offers no useful account of what is represented by judgments. For this tradition, representational relations hold between *things*. This categorial nominalism of the designational model extends even to predicates, which are understood as 'general names,' standing for universals in the same sense in which singular terms stand for particular objects. Not until Frege's semantic interpretation of predicates as corresponding to functions—and hence as not being names of any sort—would the idea of semantic relations that are not assimilable to the name/named model enter the tradition.

Applied to propositional contents, the hegemony of the designational semantic model results in two characteristic mistakes: assimilating sentences to complex names, and assimilating judging to predication. Kant provides the raw materials needed to move beyond these conceptions, but even he is not able to free himself entirely from them. That the first is a mistake becomes clear in the context of an attempt to explain the difference between *referring to a complex object*, for instance a squiggly blue line between a round dot and a square one, and *stating a fact* about its components, for instance *saying*

that the squiggly blue line is between a round dot and a square one.<sup>31</sup> This crucial difference can be elided by an incautious assimilation of each to a generic notion of *representing*, for in each case the speaker can be said to be representing something.

This difficulty is merely relocated by the introduction of a notion of state of affairs defined as the sort of thing that is represented by utterances that purport to state facts. Making this move is a version of the attempt to solve ontologically the problem of distinguishing referring from saying or stating; the idea is that each is representing, and the specific differences between them are a matter of the kind of thing represented. At the least, such a strategy demands a careful account of the relation between complex objects and the corresponding states of affairs. Any account along these lines of discourse that purports to state facts by the assertive utterance of declarative sentences is also obliged to tell a story about the states of affairs corresponding to normative claims—for instance to the claim that Kant ought not to have written *Dreams of a Spirit-Seer*, or that anyone committed to the claim that snow is white is committed to the claim that snow is spatially extended.<sup>32</sup>

The second mistake mentioned above as consequent on the unfortunate sway of designational semantic models corresponds to one way in which the representational tradition in semantics has attempted to acknowledge the special role played by the propositional. This strategy depends on a distinction on the side of the activity of representing, rather than (just) the category of thing represented, by distinguishing between representing as referring or naming, on the one hand, and representing as predication, on the other. The notion that making a claim can be modeled on representing something (particular) as something (general), in the linguistic case picking out an object with a singular term and predicating something of it with a general term, has a distinguished history. The next chapter discusses Frege's decisive demonstration that this approach is a mistake. It is not simply a mistake, however. Looking more closely at what is right about this broadly *classificatory* model of consciousness and at how it can be fixed up in response to some fundamental difficulties provides a way into an idiom for talking about semantic content that does not employ representational vocabulary at the outset.

### III. CONCEPTUAL CLASSIFICATION AND INFERENCE

#### 1. *Classification*

An ancient tradition insists that cognition essentially involves generality or universality. Particulars are not directly intelligible as such. Knowing or understanding something particular requires assimilating it to others, taking it to be like them in some way, and so to be an instance of a kind. Kant's account of cognition as beginning with the classification of

intuitions under concepts is a particularly well-developed representative of this tradition. Believing or judging, taking-true in general—for Kant the central sort of cognitive intentional state or act—has as its most basic form subsuming something particular under a universal.<sup>33</sup> In conceiving judgment (the activity of the cognitive subject, the exercise of its faculty of spontaneity, namely understanding) in terms of the *classificatory* employment of concepts, Kant adopts a model that animates as well the thought of the pre-Kantian tradition he inherits—a tradition that had not yet achieved his insight into the privileged role of judgment as the preeminent form of cognitive activity.

That model evidently underlies the epistemologies of both his empiricist and his rationalist predecessors. It forms the common background of their dispute over the source of the universals or concepts by relation to which particulars become intelligible; it is what makes urgent the question whether those universals are formed by abstraction from a more primitive kind of nonconceptual awareness of particulars, or whether on the contrary a grasp of such concepts is a precondition of anything recognizable as awareness at all. Kant follows the rationalists in treating the classificatory account of cognition as a classificatory account of consciousness generally. All awareness is understood as exhibiting the classificatory structure of universal or repeatable concepts subsuming particulars. Where earlier empiricists admit varieties of conscious apprehension short of conceptual comprehension—immediate, nonclassificatory awareness of determinate sense repeatables, for instance—Kant denies apprehension without comprehension, insisting that there must be conceptual classification wherever there is any sort of awareness. Awareness of what is classified and of how things can be classified derives from awareness that consists in classifying.<sup>34</sup>

A pragmatic version of this classificatory model results if it is de-intellectualized, stripped of residual commitments to understanding concepts as *explicit* to the mind—whether in the Kantian form of rules or recipes for it to follow in its synthesizing activity or in the pre-Kantian form as objects of its direct, nonclassificatory awareness. From such a perspective, the roots of conceptual classification are to be found in treating something in practice as being of a certain kind—taking something (particular) *as* something (universal), by behaving toward it in a way that assimilates it to others. Particular objects are classified as belonging together in some respect by being responded to alike in practice. A respect of similarity in what is responded to then corresponds to a repeatable response. Hegel develops such a pragmatic, indeed naturalized, version of Kant's account in the form of an *erotic* theory of the origins of awareness, an account of animal *desire* as the source of classification. As he puts it, an animal classifies some particular as food when it “falls to without further ado and eats it up.”<sup>35</sup> Eating something is treating it, responding to it, classifying it in practice *as* food. It exhibits a kind of practical, pre-Cartesian awareness of it *as* being of a certain kind. That

repeatable activity on the part of the organism induces a repeatable respect of similarity among the things that tend to elicit that activity.

On this account, classification of particular stimuli as instances of a general kind is *implicit* in what the responding organism *does*. So to each sort of thing that it does, there corresponds a different sort of repeatable proto-concept under which things can be classified: as food, sexual partner, prey, or predator, and so on. That no sort of explicit awareness is presupposed by this sort of implicit practical classificatory awareness or understanding is clear from the fact that all that the latter requires is a reliable differential responsive disposition. For any concrete object displays such dispositions. A chunk of iron reliably responds to some environments by melting, to others by rusting, to still others by falling. In each case it can be understood as classifying that environment, treating it in practice as being of a certain kind, assimilating it to some other possible environments and distinguishing it from others, by responding to it in a certain way.<sup>36</sup>

The Kantian rationalistic strategy of demarcation by *sapience*, awareness, and consciousness in a sense that requires the application of *concepts* would be trivialized by a classificatory model of the use of concepts that indiscriminately discerns classification according to concepts in the responsive regularities exhibited by the antics of every physical system whatsoever. Classification by the exercise of regular differential responsive dispositions may be a necessary condition of concept use, but it is clearly not a sufficient one. Such classification may underlie the use of concepts, but it cannot by itself constitute discursiveness. The chunk of iron is not conceiving its world as wet when it responds by rusting. Why not? What else must be added to responsive classification to get to an activity recognizable as the application of concepts? What else must an organism be able to do, what else must be true of it, for performances that it is differentially disposed to produce responsively to count as applications of concepts to the stimuli that evoke those responses? One dimension of a reply was indicated in the previous chapter—a normative dimension is required, which can underwrite a distinction between correct and incorrect applications of concepts. But many things can be done correctly or incorrectly. The question being asked now is what it is for what is subject to such assessment to be concept use (rather than, say, hammer use, or tooth use).

## 2. *Inferential Demarcation of the Conceptual*

An easy answer is that the response must be the forming of a belief or the making of a claim, acquiring a state or attitude or producing a performance that has an intentional content. This is of course correct, but unhelpful in the current setting. For the question is precisely what is required for a response to count as contentful in this sense. What is wanted is a characterization that does not appeal to semantic concepts such as *content*

and *concept*. If the issue is put in terms of the semantic concept of representation, it takes the form of inquiring as to what more is needed, beyond being a representation in the responsive-classificatory sense, to be a discursive or intentional representation, one that is conceptually contentful. (According to the idiom being employed, implicit grasp of such contents, of the representational purport they consist in—a grasp to be conceived of as some sort of practical mastery, as a kind of know-how—would then in favored cases count as sapient consciousness or awareness of what is represented as exhibiting a certain character.)

A more concrete way to put the question is to ask, What are the salient differences between a measuring instrument, such as a thermometer or spectrophotometer, and an observer who noninferentially acquires beliefs or makes claims about environing temperatures and colors? Artificial instruments differ from other physical systems, such as chunks of iron, only in having been constructed so that some subset of the partition of possible stimuli into equivalence classes according to the distinguishable responses the instruments are disposed to produce corresponds to some distinction of practical or theoretical significance to the user, who thereby attaches some significance to them. Suppose a spectrophotometer is hooked up to a tape recorder in such a way that it produces a noise of the acoustic type "That's red" when and only when it is irradiated with light of the proper frequency. And suppose that a fanatical human red-reporter nearby has just the same responsive dispositions to produce those noises. That is, the two systems are disposed to respond in the same way to the same stimuli, exhibiting the same noninferential circumstances of application for their responsive classifications of things as red. What makes the noise the one produces merely a signal on the basis of which someone else might conclude that something red is present, while the very same noise, reliably elicited under just the same circumstances from the other, counts as a noninferential report, expressive of a perceptually acquired belief, with an intentional content that includes the concept *red*? To vary the case, suppose the reporter's differential responsive dispositions to call things red are matched by those of a parrot trained to utter the same noises under the same stimulation. What practical capacities of the human distinguish the reporter from the instrument or the parrot? What, besides exercise regular differential responsive dispositions, must one be able to do, in order to count as having or grasping concepts, and so as able to perform not only classification but specifically *conceptual* classification?

Putting things this way makes it clear that what is at issue is a kind of *understanding*. The reporter's response is meaningful—not just, as in the case of the measuring instrument or the parrot, to others, but to the responding reporter personally. The spectrophotometer and the parrot do not understand their responses; those responses mean nothing to them, though they can mean something to us. The reporter understands the response he or she makes, attributes to it a kind of significance that the measuring instrument

and the parrot are oblivious to. The challenge is to explain what sort of practical capacity the relevant kind of understanding consists in, without an ultimately circular appeal to semantic concepts such as intentional content, concept-use, or the uptake of representational purport (treated as an explanatory primitive).

The leading idea of the approach to content and understanding to be developed here is due to Sellars. Sellars's suggestion is that the key element missing from the parrot and the measuring instrument—the difference between merely *responsive* classification and *conceptual* classification—is their mastery of the practices of giving and asking for *reasons*, in which their responses can play a role as *justifying* beliefs and claims. To grasp or understand a concept is, according to Sellars, to have practical mastery over the *inferences* it is involved in—to know, in the practical sense of being able to distinguish, what follows from the applicability of a concept, and what it follows from. The parrot does not treat "That's red" as incompatible with "That's green," nor as following from "That's scarlet" and entailing "That's colored." Insofar as the repeatable response is not, for the parrot, caught up in practical proprieties of inference and justification, and so of the making of further judgments, it is not a conceptual or a cognitive matter at all. What the parrot and the measuring instrument lack is an appreciation of the significance their response has as a reason for making further claims and acquiring further beliefs, its role in justifying some further attitudes and performances and ruling out others. Concepts are essentially inferentially articulated. Grasping them in practice is knowing one's way around the proprieties of inference and incompatibility they are caught up in. What makes a classification deserve to be called *conceptual* classification is its *inferential* role. It is practical mastery of the inferential involvements of a response, the responder's understanding it in this sense, that makes the response an intentional state or performance—one having a content for the one whose state or performance it is, and not merely for those using it as an indicator.

### 3. Holistic Consequences of Inferential Approach to Concepts

One immediate consequence of such an inferential demarcation of the conceptual is that one must have many concepts in order to have any. For grasping a concept involves mastering the proprieties of inferential moves that connect it to many other concepts: those whose applicability follows from the applicability of the concept in question, those from whose applicability the applicability of the target concept follows, those whose applicability precludes or is precluded by it. One cannot have just one concept. This holism about concepts contrasts with the atomism that would result if one identified concepts with differential responsive dispositions. The capacity to

treat some things as food by eating them need have no particular connection to the capacity to treat other things as dangerous by fleeing them. To treat states or performances as intentionally contentful in the sense of being conceptually articulated involves treating them as situated in a web of proprieties of inferential transitions from one content to another. Knowing one's way around the bit of the web centered on one conceptual content, being able to tell in practice which moves to it and from it are permitted or required and which forbidden, accordingly requires mastery of the proprieties of inference that govern the use of other concepts and contents as well.

By contrast, there is *prima facie* no reason why the fact that some object or property is represented by one simple idea, term, or predicate should be relevant to what is represented by others. Representational relations between nonintentional objects or properties and the intentional representings of them might be treated (as the empiricists in fact treat them) as separate building blocks that, when properly put together, determine what inferences are good in the sense of preserving accuracy of representation. Serving this role seems compatible with these representational relations being quite independent of one another. Knowing what one state or expression represents need convey no information at all about what anything else might represent.

But the inferential notion of semantic content is essentially holistic. Inferences involve both premises and conclusions. The inferential role of one of the premises essentially depends on that of the conclusions, and vice versa. One could not know something about the inferential role of one content without knowing at least something about the inferential roles of others that could be inferred from it, or from which it could be inferred. Contents understood in terms of inferential roles are evidently interdefined in a way in which contents understood in terms of representational purport need not be.

In his masterwork, "Empiricism and the Philosophy of Mind," Sellars exploits this consequence of his insight into the significance of inferential connections to concept-use, even in cases of responsive classification. He argues there that noninferential reports, by which perceptual states are made explicit, cannot constitute an autonomous fragment of a language—one that might be understood though no others are. Observation reports do indeed have a certain priority in the order of *justification* of empirical claims. But they cannot be accorded a similar priority in the order of *understanding* of those claims. Since knowledge requires not only justification but grasp or understanding of the content being justified, there can be no observational knowledge without inference. There can be no purely observational language or set of concepts with respect to which one could then ask whether the decision to append an inferential superstructure is a rational or justifiable one. The rock on which foundationalism founders is accordingly its incapacity to explain what it is to *understand* the significances of elements in the observational justificatory basis. For in order to be able to apply one concept

noninferentially, exercising a disposition to respond differentially to nonlinguistic stimuli, one must be able to apply others inferentially. Unless the response has such an inferential significance, it is not a conceptually contentful response. So the idea of an autonomous language game (or set of practices of applying concepts) consisting entirely of noninferential reports (even of purely mental happenings) is a radical mistake.

The argument does not rule out the possibility of languages or conceptual schemes that are devoid of *theoretical* claims and concepts—that is, that lack concepts that are applicable only as conclusions of inferences. One can have a scheme in which all the concepts have reporting uses and so are in this sense concepts of observables. But they must also have inferential uses. *Red* can be applied either noninferentially, as a response the reporter has been trained to make to a certain kind of visual stimulus, or inferentially, on the basis of entitlement to a prior application of the concept *scarlet*. The conclusion that there can be no conceptually articulated observation apart from inferential capacities holds equally whether what is being reported consists of external observable situations or internal, purely mental happenings. It is this argument that lies at the base of Sellars's critique of broadly Cartesian philosophies of mind.

#### 4. Inference and Practice

As ought to be expected from his discussion of the regress-of-rules argument, it is important to Sellars that the inferential conception of concepts connects the grasp or understanding of concepts (the uptake of conceptual content) with a certain kind of practical activity. Inferring is a kind of doing. Acknowledgment of inferential proprieties need not be explicit in the endorsement of rules or principles of inference but may remain implicit in the capacity to take or treat inferential transitions as correct or incorrect in practice. Inferential relations among concepts are implicit in the practice of giving and asking for reasons. The norms that govern these justificatory practices can be understood to confer inferentially articulated contents on the states, attitudes, and performances subject to them: for something to have such content just is for such norms to determine how it is correctly used or manipulated. The status of inference as something that can be *done* accordingly holds out the promise of securing an appropriate relation between pragmatics, the study of the practices, and semantics, the study of the corresponding contents.

Furthermore, because the activity through which the norms get their grip on conceptual contents is construed as *inference*, it is specifically *propositional* contents that in the first instance count as conceptually articulated. Inferential relations hold, in the paradigm case, between contents that are expressed explicitly by declarative sentences. The premises of inferences, and in the central cases their conclusions as well, must be understood to have a

propositional form. Insofar as an independent theoretical grip is possible on the notion of inference, propositional contents can be picked out by appeal to this property. (This is the strategy pursued in the next chapter.) So on an inferential rendering of the conceptual, the sort of doing that inferring is yields in a natural way the priority of propositional conceptual contents.

The pragmatic turn aside, this view too is due to Kant. As Sellars puts it: "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."<sup>37</sup> The subtlety and sophistication of Kant's concept of representation is due in large part to the way in which it is integrated into his account of the inferential relations among judgments. It remained for Hegel, however, to complete the inversion of the traditional order of semantic explanation by beginning with a concept of experience as inferential activity and discussing the making of judgments and the development of concepts entirely in terms of the roles they play in that inferential activity. Although something like this point had been implicit in Kant's notion of reason as systematicity, it was the young Hegel who first appreciated the line of reasoning, made familiar to us by Quine in "Two Dogmas"—namely, that if the content of a claim must at least determine what follows from it (what else it commits one to), then since what a claim commits one to depends on what collateral commitments are available to serve as additional premises ("auxiliary hypotheses"), the significance of undertaking any particular commitment cannot be determined without appeal to the contents of all those collateral commitments.

Hegel's two central semantic concepts in the *Phenomenology* are both inferential notions. "Mediation," his term for inferential articulation, is derived from the role of the middle term in syllogistic inference. "Determinate negation" is his term for material incompatibility, from which, he takes it, the notion of formal negation is abstracted. The contents of concepts are identified and individuated by the functional roles they play in historically evolving webs constituted by relations of mediation and determinate negation, that is, by their material inferential and incompatibility relations to each other. Hegel's interest in the significance of inference in semantics does not (as with Kant) arise primarily in the investigation of how it might be combined with representationalist insights (although he has something to say about that too). It arises rather in the investigation of how this rationalist insight might be combined with the insights of the Romantic expressivists.

The Romantics are perhaps best known for their rejection, not just of the Enlightenment's representationalism, discussed above, but also for their rejection of the significance it assigns to reason. They sought to displace the general demarcational emphasis on giving and asking for reasons or inquiring after truth, not just the specific version that sought to understand these matters in representational terms. The Romantic recoil from understanding us as representers overshot that mark and came to rest in an esteem for

feeling and inarticulate empathy and enthusiasm. Hegel saw in inferential notions of content a way to join the Romantics in rejecting representationalism, while parting company with them in their hostility to reason. The result is a synthesis of Enlightenment inferentialism and Romantic expressivism.<sup>38</sup>

### *5. Inferentialism and Representationalism*

Kant, however, did not originate the inferentialist line of semantic thought that Sellars appropriates from him, and that Hegel develops. In a discussion of his break with traditional empiricism, prompted by the issue of the sort of content that ought to be associated with logical, causal, and deontological modalities, Sellars puts the idea that stands at the center of his systematic thought in the form in which it originally occurred to him in the 1930s: "What was needed was a functional theory of concepts which would make their role in reasoning, rather than supposed origin in experience, their primary feature."<sup>39</sup> Put this way, the idea forms one of the mainstays of classical rationalism, even in the absence of Kant's insight about the privileged role that must be assigned to judgments on such an inferential-functional approach. Pre-Kantian empiricists and rationalists alike were notoriously disposed to run together causal and conceptual issues, largely through insufficient appreciation of the normative character of the "order and connection of ideas" that matters for the latter. But there is another, perhaps less appreciated, contrast at work here, besides that of the causal and the conceptual. Enlightenment epistemology was always the home for two somewhat uneasily coexisting conceptions of the conceptual. The fundamental concept of the dominant and characteristic understanding of cognitive contentfulness in the period initiated by Descartes is of course *representation*. However there is a minority semantic tradition that takes *inference* rather than representation as its master concept.

Rationalists such as Spinoza and Leibniz accept the central role of the concept of representation in explaining human cognitive activity, but they are not prepared to accept Descartes's strategy of treating the possession of representational content as an unexplained explainer. Each of them develops instead an account of what it is for one thing to represent another, in terms of the *inferential* significance of the representing. They are explicitly concerned (as Descartes is not) to be able to explain what it is for something to be understood, treated, or employed in practice *as* a representing by the subject—what it is for it to be a representing *to* or *for* that subject. Their idea is that the way in which representings point beyond themselves to something represented is to be understood in terms of *inferential* relations among representings. States and acts acquire content by being caught up in inferences, as premises and conclusions.<sup>40</sup> Thus a big divide within Enlightenment epistemology concerns the relative explanatory priority accorded to the concepts of representation and inference.

The British empiricists are more puzzled than Descartes about representational purport, the property of seeming to be about something. But they are clear in seeking to derive inferential relations from the contents of representings, rather than the other way around. In this regard they belong to the still-dominant tradition that reads inferential correctnesses off from representational correctnesses, which are assumed to be antecedently intelligible. The post-Cartesian rationalists, the claim is, give rise to a tradition based on a complementary, semantically reductive order of explanation. These *inferentialists* seek to define representational properties in terms of inferential ones, which must accordingly be capable of being understood antecedently. They start with a notion of content as a matter of what is a reason for what and understand truth and representation as features of ideas that are not only manifested in, but conferred by their role in reasoning. This is the tradition that Sellars inherits and builds on by developing a notion of conceptual content that starts with inferential roles.

#### IV. MATERIAL INFERENCE, CONCEPTUAL CONTENT, AND EXPRESSION

##### 1. Frege on Conceptual Content

The rationalists' inferential understanding of conceptual content, which Kant inherits and which remains one of the strands from which his systematic semantic tapestry is woven, provides the starting point as well for Frege's semantic investigations. Frege may seem an unlikely heir to this inferentialist tradition. After all, he is usually thought of as the father of the contemporary way of working out the representationalist order of explanation. Its strategy is to start with an independent notion of relations of reference or denotation obtaining between mental or linguistic items and objects and sets of objects in the largely nonmental, nonlinguistic environment. Then it determines from these in the familiar fashion: first truth conditions for the sentential representings built out of the subsentential ones, and then, from these, a notion of goodness of inference understood in terms of set-theoretic inclusions among the associated sets of truth conditions. But insofar as it is appropriate to read this twentieth-century story back into Frege at all (a dangerous and potentially misleading enterprise), it would be possible only beginning with the Frege of the 1890s. He starts his semantic investigations, not with the idea of *reference*, but with that of *inference*. His seminal first work, the *Begriffsschrift* of 1870, takes as its task the explicit expression of *inferential* roles: "In my formalized language [*Begriffsschrift*] . . . only that part of judgments which affects the possible inferences is taken into consideration. Whatever is needed for a correct [richtig] inference is fully expressed; what is not needed is . . . not."<sup>41</sup>

These inferential roles form the basis of his notion of *content*. It is because the sorts of contents that are associated with expressions are to be defined

in the first place in terms of inference that Frege must insist on the distinction between the sorts of contents that can, and those that cannot, serve as premises and conclusions of inference, and so play the basic sort of inferential roles. "We distinguish contents that *are*, and contents that *are not, possible contents of judgment*."<sup>42</sup> Frege's Kantian insistence on the priority of the propositional, of judgeable contents, is an aspect of his pursuit of the rationalists' inferentialist order of semantic explanation. He embraces Kant's insight that the notion of content must be made intelligible first for judgments, which alone can figure as premises and conclusions of inference, and only then extended to the contents expressed by fragments of declarative sentences. Recall the passage (already quoted in Section II of this chapter) in which he contrasts his procedure with that pursued by others in the tradition: "In Aristotle, as in Boole, the logically primitive activity is the formation of concepts by abstraction, and judgment and inference enter in through an immediate or indirect comparison of concepts via their extensions . . . I start out from judgments and their contents, and not from concepts . . . I only allow the formation of concepts to proceed from judgments . . . Instead of putting a judgment together out of an individual as subject and an already previously formed concept as predicate, we do the opposite and arrive at a concept by splitting up the content of a possible judgment."<sup>43</sup> It is for this reason that the fundamental definition introducing the notion of "conceptual content" (*begriffliche Inhalt*) (for which, as its name implies, the *Begriffsschrift* is supposed to supply a means of explicit expression) applies only to the contents of possible judgments. It will have to be extended later, by Frege's substitutional methodology, to allow the assignment of *indirectly* inferential roles to subsentential expressions, according to the contribution their occurrence makes to the directly inferential role (as premise or conclusion) of judgment-expressing sentences in which they occur. The substitutional strategy that Frege devised for quarrying subsententially expressed contents from sententially expressed ones is of the first importance for carrying out the inferentialist semantic explanatory program. Much is made of it in subsequent chapters of this work. Before Frege, one could only hope that there was some way of bridging this gap.

That the target notion of content is specifically *conceptual* content is accordingly not to be understood in terms of some antecedent notion of concepts. Rather, the conceptual is explicitly construed in *inferential* terms: "There are two ways in which the content of two judgments may differ; it may, or it may not, be the case that all inferences that can be drawn from the first judgment when combined with certain other ones can always also be drawn from the second when combined with the same other judgments. The two propositions 'the Greeks defeated the Persians at Plataea' and 'the Persians were defeated by the Greeks at Plataea' differ in the former way; even if a slight difference of sense is discernible, the agreement in sense is preponderant. Now I call that part of the content that is the same in both the

conceptual content. Only this has significance for our symbolic language [*Begriffsschrift*.]<sup>44</sup> Two claims have the same conceptual content if and only if they have the same inferential role: a good inference is never turned into a bad one by substituting one for the other. The fundamental semantic assignment of conceptual content to judgments is derived from the ultimately pragmatic notion of *correctness of inference*. This derivation is the first application of the substitutional methodology: semantically assimilating expressions accordingly as substitution of one for another preserves some semantically relevant property. In this case (prior to the others in the order of explanation), the semantically relevant invariant is propriety of inference.

This way of specifying the explanatory target to which semantic theories, including referential ones, are directed is picked up by Frege's student Carnap, who in *The Logical Syntax of Language* defines the content of a sentence as the class of nonvalid sentences that are its consequences (that is, can be inferred from it). Sellars in turn picks up the idea from him, as his references to this definition indicate.<sup>45</sup> As will emerge, an important feature of Carnap's definition is the appeal to *nonvalid* consequences. In this way what pertains to the *content* of a claim is distinguished from what pertains to its *form*.

This distinction is operative in the *Begriffsschrift* as well. Yet when Frege wants to be clear about what is expressed by even the purely formal assertions appearing in proofs about the expressive capacity of the *Begriffsschrift* itself, he does so by specifying their inferential role, restricting himself in this case to inferences whose propriety is underwritten by their form alone. So each assertion is introduced by displaying a proof of it from already-established assertions, thereby exhibiting the premises from which it follows as conclusion. Showing what a claim follows from is not sufficient to specify its inferential role, however. It matters as well what follows from it. Indeed, Frege often complains (for instance in the *Grundlagen*) about systems that introduce definitions that are never then employed in subsequent demonstrations. These provide a case where looking at inferential consequences is particularly important; since definitions do not have inferential antecedents, if their inferential consequents are not specified, their content is left entirely indeterminate. In order to complete the specification of the inferential roles of the assertions of the system he presents, Frege appends to the *Begriffsschrift* a list indicating for each assertion all of the subsequent assertions in whose proof it is used as premise. That is, he specifies for each assertion what follows from it (together with other assertions, of course) as well as what it follows from. In this way he makes explicit the inferential roles, and so the conceptual contents, conferred on the judgments he puts forward by the purely formal reasoning involving them that is displayed in his book.

In contrast to his original procedure, the tradition Frege initiated in the 1890s makes truth, rather than inference, primary in the order of semantic explanation. Dummett says of this shift: "In this respect (and in this respect

alone) Frege's new approach to logic was retrograde. He characterized logic by saying that, while all sciences have truth as their goal, in logic truth is not merely the goal, but the object of study. The traditional answer to the question what is the subject-matter of logic is, however, that it is, not truth, but inference, or, more properly, the relation of logical consequence. This was the received opinion all through the doldrums of logic, until the subject was revitalized by Frege; and it is, surely, the correct view.<sup>46</sup> And again: "It remains that the representation of logic as concerned with a characteristic of sentences, truth, rather than of transitions from sentences to sentences, had highly deleterious effects both in logic and in philosophy. In philosophy it led to a concentration on logical truth and its generalization, analytic truth, as the problematic notions, rather than on the notion of a statement's being a deductive consequence of other statements, and hence to solutions involving a distinction between two supposedly utterly different kinds of truth, analytic truth and contingent truth, which would have appeared preposterous and irrelevant if the central problem had from the start been taken to be that of the character of the relation of deductive consequence."<sup>47</sup> The important thing to realize is that the Frege of the *Begriffsschrift* has not yet made this false step. Of course, adopting a semantic order of explanation that begins with proprieties of inference requires both an account of those proprieties (that is, an account of its raw materials) and an account of how talk about truth is eventually to be construed in these terms (that is, an account of its consequences). This is the strategy pursued in this work. The first of these challenges is responded to in Chapters 3 and 4, and the second in Chapter 5.

There are two further points to keep in mind regarding this passage of Dummett's. First, shifting from concern with *inference* to concern with *truth* is one move; understanding *truth* in terms of prior primitive *reference* relations involving objects and properties is another. Since the mature Frege treats truth as indefinable and primitive, the extraction of a representationalist commitment even from the texts of the 1890s requires further showing (compare Davidson's truth-without-reference view in our own day). Second, understanding the topic of logic in terms of inference is not the same as seeing it in terms of *logical inference*, or of "deductive consequence," as Dummett puts it (see the discussion of "formalism" about inference, below). The view propounded and attributed to Frege below is a different one—and from the contemporary vantage point it is a more surprising one than the one that Dummett endorses here.

## 2. Material Proprieties of Inference and the Dogma of Formalism

The kind of inference whose correctnesses essentially involve the conceptual contents of its premises and conclusions may be called, following Sellars, "*material* inference."<sup>48</sup> 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. Endorsing these inferences is part of grasping or mastering those concepts, quite apart from any specifically logical competence. From the point of view of a familiar sort of semantics (different from that to be explored here), one could say that the set of possible worlds in which the premises of these inferences are true is a subset of the set of possible worlds in which their conclusions are true. Since neither the premises nor the conclusions of such inferences employ logical concepts, it seems appropriate to distinguish them from inferences whose correctness depends only on logical form.

Often, however, *inferential* articulation is identified with *logical* articulation. Material inferences are then treated as a derivative category. The idea is that being *rational*—mastering proprieties of inference and so being subject to the force of the better reason—can be understood as a purely *logical* capacity. In part this tendency is encouraged by merely verbally sloppy formulations of the crucial difference between the inferential force of reasons and the physically efficacious force of causes: formulations that render it as the difference between 'logical' and 'natural' compulsion. Mistakes ensue, however, if the concept *logical* is employed with these circumstances of application conjoined with consequences of application that restrict the notion of the logical force of reasons to *formally* valid inferences. The substantial commitment that is fundamental to this sort of approach is what Sellars calls "the received dogma . . . that the inference which finds its expression in 'It is raining, therefore the streets will be wet' is an *enthymeme*."<sup>49</sup>

According to this line of thought, wherever an inference is endorsed, it is because of belief in a conditional. Then the instanced inference is understood as implicitly involving the conditional "If it is raining, then the streets will be wet." With that "suppressed" premise supplied, the inference is an instance of the formally valid scheme of conditional detachment. The "dogma" expresses a commitment to an order of explanation that treats all inferences as good or bad solely in virtue of their form, with the contents of the claims they involve mattering only for the truth of the (implicit) premises. According to this way of setting things out, there is no such thing as material inference. This view—which understands "good inference" to mean "formally valid inference," postulating implicit premises as needed—might be called a *formalist* approach to inference. It trades primitive goodnesses of inference for the truth of conditionals. Doing so is taking a retrograde step that corresponds to the one Dummett complains about. The grasp of logic

that is attributed must be an *implicit* grasp, since it need be manifested only in distinguishing material inferences as good and bad, not in any further capacity to manipulate logical vocabulary or endorse tautologies involving them. But what then is the explanatory payoff from attributing such an implicit logical ability rather than just the capacity to assess proprieties of material inference?

It is worth considering an example of how formalist presuppositions can be embodied misleadingly in vocabulary. Here is Dennett in "Intentional Systems":

Earlier I alleged that even creatures from another planet ["in virtue of their rationality"] would share with us our beliefs in logical truths; light can be shed on this claim by asking whether mice and other animals, in virtue of being intentional systems, also believe the truths of logic. There is something bizarre in the picture of a dog or mouse cogitating a list of tautologies, but we can avoid that picture. The assumption that something is an intentional system is the assumption that it is rational; that is, one gets nowhere with the assumption that entity  $x$  has beliefs  $p, q, r \dots$  unless one also supposes that  $x$  believes what follows from  $p, q, r \dots$ ; otherwise there is no way of ruling out the prediction that  $x$  will, in the face of beliefs  $p, q, r \dots$  do something utterly stupid, and, if we cannot rule out that prediction, we will have acquired no predictive power at all. So whether or not the animal is said to believe the truths of logic, it must be supposed to follow the rules of logic.<sup>50</sup>

Dennett understands intentionality in terms of rationality (as the view being developed here does), and understands rationality in terms of the discrimination in practice of good inferences ("what follows") from bad ones (as the view being developed here does). But there is a slide here from "follows" to "logically follows." No justification is offered for the move, first, from discriminating good from bad inferences to the need for any specifically logical capacity or, second, for the move from logical capacity to belief in logical truths. Perhaps appropriate (even logically valid) inferences can be endorsed without commitment to the corresponding (logical) conditional truths.

On the first point: perhaps there are good nonlogical inferences, and rationality consists in the way discriminating them matters to one's deliberations and assessments. Why should "following the rules of logic" be either necessary or sufficient for this discrimination? In any case, it was argued in Chapter 1 that one ought to distinguish both exhibiting a regularity and acknowledging a norm implicitly in one's practice (two construals of discriminating good from bad inferences) from following a rule. On the second point, Dummett was cited above as pointing out that defining logical consequence in terms of logical truth is neither a trivial nor a harmless move.

In fact Dennett (and in this regard he is typical) thinks of this way of putting things as a harmless *façon de parler*, warranted by a general inter-

changeability of talk of endorsing inferences and talk of believing conditionals. The “belief in logical truths,” or even, less committally, endorsement of logically good inferences, that he has in mind is implicit in practical discriminations. The passage continues: “Surely our mouse follows or believes in modus ponens, for we ascribed to it the beliefs: (a) there is a cat to the left, and (b) if there is a cat to the left, I had better not go left, and our prediction relied on the mouse’s ability to get to the conclusion.” What was actually attributed to the mouse is a belief with content (a) and a desire to avoid the cat. Citing its intelligent behavior licenses the attribution of a practical inference. It does not, by itself, tell for or against expressing that inference as a material inference or as detachment from an endorsed conditional. Why should all inferences be assimilated to detachments, or other formal logical rules of inference? Dennett’s justification is that “in general there is a trade-off between rules and truths; we can suppose  $x$  to have an inference rule taking  $A$  to  $B$  or we can give  $x$  the belief in the ‘theorem’: if  $A$  then  $B$ . As far as our predictions are concerned, we are free to ascribe to the mouse either a few inference rules and belief in many logical propositions, or many inference rules and few if any logical beliefs.”<sup>51</sup>

The conditional beliefs that can be traded off for endorsements of inferences should not be called “logical” beliefs simply because they concern inferences. Though they involve logical concepts, namely the conditional, they are not in general logically true. Indeed, Dennett continues: “We can even take a patently nonlogical belief like (b) and recast it as an inference rule taking (a) to the desired conclusion.” To do so would be to establish or endorse a *material* correctness of inference, what Dennett calls “a set of nonlogical inference rules.” Once the possibility of this sort of inference is acknowledged, inferential formalism surrenders a priori privileges and must contend with inferential materialism for privileges of explanatory priority. According to the famous argument of Lewis Carroll in “What the Tortoise Said to Achilles,” as Dennett acknowledges, some inferential commitments (“rules of inference”) must be attributed if any consequences are to be licensed by the attribution of beliefs, even conditional beliefs. So there must be “rules” as well as “truths.” However, once the purely formal-logical inferences are allowed (paradigmatically detachment inferences licensed by conditionals), accounts of rational performance can take the form either of attributions of endorsements of material inferences or of conditional propositions, as might be theoretically convenient for other reasons. Either decision ought to be justified.

What considerations ought to persuade a theorist to accord explanatory priority to the attribution of material inferential commitments or to the attribution of conditional propositional commitments, and so to treat material or formal inference as fundamental? Dennett’s answer is: “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.<sup>152</sup> This is a formalist position, in that all inferences are assimilated to detachments and are understood as involving, at least implicitly, endorsements of conditionals whose logical content explicitly relates premises and conclusions. The only concession to material inferences arises in the possibility of licensing detachment in a retail, content-respecting fashion, rather than wholesale, in a purely formal logical way. But why should this model be employed? Why should all goodness of inference be seen as logical goodness, even at the cost of postulating "implicit" premises involving logical concepts?

What is at issue is two different ways of understanding the relation between something implicit and an explicit expression of it. It is possible to agree with the formalist in understanding conditionals as inference licenses, which make explicit in the content of a claim what is implicit in the endorsement of an inference, without going on to construe all inferences as involving the use of conditionals. The question is how one ought to construe the relation between what is explicit in the form of a rule or principle (in this case a conditional claim) and what is implicit in proprieties of practice (in this case in the endorsement of an inference). The formalist line of thought begins with explicit propositional licenses that license inferences in virtue of their logical form. Material inferences (say from rain to wet streets or vice versa) are understood privatively: as enthymemes resulting from the suppression or hiding of one of the premises required for a proper warrant. Opposed to this might be a pragmatist line of thought, beginning with material inferences—that is, nonlogical, content-based reasoning. It would then be necessary to explain how logical vocabulary such as the conditional is to be understood as permitting the expression of those implicit inferential commitments in an explicit fashion—that is, as judgeable, claimable, believable contents, as the contents of potential propositional commitments.

There are general reasons to prefer an order of explanation that begins with what is implicit in practice (what people do) and proceeds to an account of what they explicitly believe or say, over one taking the opposite tack. Only in this way can one hope to understand believing or saying in terms of more primitive capacities (knowing-that in terms of knowing how). That asymmetry manifests itself in this case in the question of how one understands logical concepts or the use of logical vocabulary. On the formalist line, anything that has any inferential capacities at all is credited with mastery of a battery of logical concepts and the corresponding inference rules without which they would be without content. These can be thought of as introduction and elimination rules, of which detachment is a cardinal example. Logical concepts are quite different from others in being presupposed by all contentful concepts and inferences. It is a short step from treating mastery of these 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. Kant rescued them by insisting that it is the formality of logical (and, more controversially, transcendental) concepts that entitles them to a special status that would indeed be absurd for ordinary contentful concepts. Assessing the Kantian formalist move requires looking more closely at what is being said when an inference is described as being valid in virtue of its logical form.

### *3. Conceptual Content and Material Inference*

Before looking at how logical concepts might function to make explicit conceptual contents that are implicit in practical proprieties of inference, however, it is worth looking more closely at the relation between inference and content. The picture being developed is one according to which materially good inferences correspond to the conceptual content of nonlogical expressions, while inferences valid in virtue of their logical form alone correspond to the conceptual content of purely logical expressions. This can be approached by considering, to begin with, the notion of *material* inferences: inferences whose propriety essentially involves the nonlogical conceptual content of the premises and conclusions. The approach Sellars endorses is best understood by reference to the full list of alternatives he considers:

We have been led to distinguish the following six conceptions of the status of material rules of inference:

(1) Material rules are as essential to meaning (and hence to language and thought) as formal rules, contributing to the architectural detail of its structure within the flying buttresses of logical form.

(2) While not essential to meaning, material rules of inference have an original authority not derived from formal rules, and play an indispensable role in our thinking on matters of fact.

(3) Same as (2) save that the acknowledgment of material rules of inference is held to be a dispensable feature of thought, at best a matter of convenience.

(4) Material rules of inference have a purely derivative authority, though they are genuinely rules of inference.

(5) The sentences which raise these puzzles about material rules of inference are merely abridged formulations of logically valid inferences. (Clearly the distinction between an inference and the formulation of an inference would have to be explored.)

(6) Trains of thought which are said to be governed by “material rules of inference” are actually not inferences at all, but rather activated associations which mimic inference, concealing their intellectual nudity with stolen “therefores.”<sup>53</sup>

His own position is that an expression has conceptual content conferred on it by being caught up in, playing a certain role in, material inferences: “It is

the first (or 'rationalistic') alternative to which we are committed. According to it, material transformation rules determine the descriptive meaning of the expressions of a language within the framework provided by its logical transformation rules . . . In traditional language, the 'content' of concepts as well as their logical 'form' is determined by the rules of the Understanding."<sup>54</sup>

Sellars, in arguing that material inferences are essential to the meaning (content) of nonlogical locutions, cites a phenomenon that is as important to the expressivist picture of logical concepts as it is to the materialist conception of inference presupposed by inferentialist approaches to conceptual content. Sellars's argument that material inferences are essential to the meaning (content) of nonlogical locutions depends on a central conceptual phenomenon. He argues for the theoretical indispensability of a conception of material inferences in terms of the practical indispensability of what is made explicit by a certain familiar kind of vocabulary. His argument is attributed to an interlocutor who maintains that:

such subjunctive conditionals as "If I had released this piece of chalk, it would have fallen," and "If there were to be a flash of lightning, there would be thunder" . . . [must be interpreted] as expressions of material rules of inference . . . He therefore claims to have shown beyond reasonable doubt not only that there are such things as material rules of inference, but, which is far more important, that they are essential to any conceptual frame which permits the formulation of such subjunctive conditionals as do not give expression to logical principles of inference. Since we are all conscious of the key role played in the sciences, both formal and empirical, in detective work and in the ordinary course of living by subjunctive conditionals, this claim, if substantiated, would indeed give a distinguished status to material rules of inference.<sup>55</sup>

He concludes: "Now, unless some other way can be found of interpreting such subjunctive conditionals in terms of logical principles of inference, we have established not only that they are the expression of material rules of inference, but that the authority of these rules is not derivative from formal rules. In other words, we have shown that material rules of inference are essential to the language we speak, for we make constant use of subjunctive conditionals."<sup>56</sup>

The point is not the indispensability of the vocabulary of conditionals that permit detachment inferences even with counterfactual premises. It is the indispensability of what those conditionals express: the implicit proprieties of material inference that they help make explicit. "Even though material subjunctive conditionals may be dispensable, permitting the language to be extensional, it may nevertheless be the case that the function performed in natural languages by material subjunctive conditionals is indispensable."<sup>57</sup>

The material inferences codified in subjunctive conditionals are inferential involvements that are essential to the contents of the concepts used in science and everyday life. These are not logically valid inferences. But logical

vocabulary, subjunctive conditionals, can be used to express these material inferential relations. Without such vocabulary, the inferences can still be endorsed. With it, those content-generating inferential endorsements can be made explicit as the content of a claim or propositional endorsement.

#### 4. From Material to Formal Properties of Inference

Should inferentialist explanations begin with inferences pertaining to propositional *form*, or those pertaining to propositional *content*? One important consideration is that the notion of formally valid inferences is definable in a natural way from that of materially correct ones, while there is no converse route. For given a subset of vocabulary that is privileged or distinguished somehow, an inference can be treated as good in virtue of its form, with respect to that vocabulary, just in case it is a materially good inference and cannot be turned into a materially bad one by substituting nonprivileged for nonprivileged vocabulary, in its premises and conclusions. This is another application of the substitutional methodology Frege employs in individuating the conceptual contents of judgments, and again in discerning indirectly conceptually contentful components within them. All it requires is a partition of vocabulary into two kinds: those that are to be held fixed and those that are to be regarded as replaceable. Call the kind of vocabulary that is to be held fixed the *K*-vocabulary. The general structure of formality definitions is then that the set of *K*-valid inferences (those that will be understood as good in virtue of their *K*-form alone) comprises those that meet the two conditions of being inferences that (1) are good inferences and (2) cannot be turned into bad inferences by substituting non-*K* for non-*K* vocabulary.

Clearly, what inferences are treated as valid in virtue of their form by such a procedure depends on how the vocabulary is divided into the two kinds. In the limit, if *all* the vocabulary were treated as irreplaceable, no substitutions of non-*K* for non-*K* vocabulary would be possible, and a fortiori none could turn a correct inference into one that is not correct. So *all* materially good inferences would count as good in virtue of their *K*-form, in the case where *K* comprises the whole vocabulary over which the field of inferences is defined. At the opposite end of the spectrum, if *no* vocabulary is treated as irreplaceable, then if there were any bad inferences at all, none of the good inferences would count as good in virtue of their *K*-form. For all could be turned into bad inferences by some substitution or other.

If the *K*-vocabulary (that which is not substituted for) is *logical* vocabulary, then the good inferences whose correctness is invariant under substitution of non-*K* for non-*K* vocabulary (nonlogical for nonlogical vocabulary) are the logically valid inferences—namely those that are good in virtue of their *logical* form. (Quine recommends this Fregean substitutional way of thinking about logical form, although he appeals to *truth* rather than propriety of

*inference* as the semantically relevant invariant whose preservation is at issue.) But this substitutional conception of what it is for an inference to be good in virtue of its form is not essentially restricted to a notion of *logical* form. If one picks out specifically zoological vocabulary or moral vocabulary or theological vocabulary to play the role of the distinguished *K*-vocabulary, the substitutional mechanism will take as its input a practical classification of inferences into good or bad, correct or incorrect, and yield as its output a distinguished set of inferences that are not just good, but are good in virtue of their zoological, moral, or theological form. The mechanism is perfectly general.

It follows that on this way of thinking about things, logical vocabulary cannot be picked out by appeal to its formality or by its involvement in formal proprieties of inference. If it is specifically logical form that is of interest, then one must antecedently be able to distinguish some vocabulary as peculiarly logical. That done, the Fregean semantic strategy of looking for inferential features that are invariant under substitution yields a notion of logically valid inferences. So the formal goodness of inferences derives from and is explained in terms of the material goodness of inferences, and so ought not to be appealed to in explaining it. And logical vocabulary must be picked out in some way that does not appeal to inferences that are formally valid or good in virtue of their form. Frege's way of specifying the characteristic linguistic role in virtue of which vocabulary qualifies as logical is discussed below.

### 5. Sellars on Expressive Rationality

So far two related claims have been introduced: that conceptual contents are inferential roles, and that the inferences that matter for such contents in general must be conceived to include those that are in some sense materially correct, not just those that are formally valid.<sup>58</sup> It will be argued in a moment that a commitment to the second of these, no less than the first, is to be found already in Frege's early writings, though not in the developed form to which Sellars brings it. But in both thinkers these ideas are combined with a third, which makes this line of thought especially attractive. In one of his early papers, Sellars introduces the idea this way: "In dealing with such situations [attempts to justify acceptance of a law 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* causally necessitates *B* as the expression of a rule governing our use of the terms '*A*' and '*B*'.<sup>59</sup> Sellars understands such modal statements as inference

licenses, which formulate as the content of a claim the appropriateness of inferential transitions. More than this, he understands the function of such statements to be making explicit, in the form of assertible rules, commitments that had hitherto remained implicit in inferential practices. Socratic method is a way of bringing our practices under rational control, by expressing them explicitly in a form in which they can be confronted with objections and alternatives, a form in which they can be exhibited as the conclusions of inferences seeking to justify them on the basis of premises advanced as reasons, and as premises in further inferences exploring the consequences of accepting them.

In the passage just cited, Sellars tells us that the enterprise within which we ought to understand the characteristic function of inductive inference is a form of rationality that centers on the notion of expression: making explicit, in a form that can be thought or said, what is implicit in what is done. This is a dark and pregnant claim, but it epitomizes a radical and distinctive insight. What follows is intended to shed some light on it and its role in an inferentialist vision of things. The general idea is that the paradigmatically rational process that Sellars invokes under the heading of "Socratic method" depends upon the possibility of making *implicit* commitments *explicit* in the form of claims. Expressing them in this sense is bringing them into the game of giving and asking for reasons as playing the special sort of role in virtue of which something has a conceptual content at all—namely an inferential role, as premise and conclusion of inferences.

This is distinct from (but obviously related to) the sort of rationality that then consists in making the appropriate inferential moves. Even totalitarian versions of the latter—for instance those that would assimilate all goodness of inference to logical validity, or to instrumental prudence (that is, efficiency at getting what one wants)—depend upon the possibility of expressing considerations in a form in which they can be given as reasons, and reasons demanded for them. All the more does Socratic reflection on our practices, particularly on those material-inferential practices that determine the conceptual contents of thoughts and beliefs, depend on the possibility of their explicit expression. Here is another early (perhaps equally dark) statement of this important Sellarsian theme:

Now, among the linguistic activities which can be discriminated are the 'explicative' or 'analytic' which, to use Ayer's phrase 'elucidate the proper use' of linguistic expressions. Furthermore the anthropologist . . . can distinguish within language activity between that which "deals directly with the environment" and that which attempts to mirror, within language itself, the relation of language to the world. In connection with this Fichtean self-diremption, the language user makes use of such words as 'means', 'true', 'verified' and so on. This is linguistic activity as semantic and pragmatic metalanguage. But the language

activity of human organisms can achieve an even greater degree of internal complexity, such as comes out most clearly in the 'explicative' metalinguistic activity of the logician and epistemologist, but is also to be found, highly confused, in more practical beings.<sup>60</sup>

### 6 . *The Expressive Project of the Begriffsschrift*

To begin to explicate this notion of explication, it is helpful to return to the consideration of the young Frege's inferentialist program. Frege's *Begriffsschrift* is remarkable not only for the inferential idiom in which it specifies its topic but equally for how it conceives its relation to that topic. The task of the work is officially an *expressive* one—not to *prove* something, but to *say* something. Frege's logical notation is designed for expressing conceptual contents, making *explicit* the inferential involvements that are *implicit* in anything that possesses such content. As the passage quoted above puts it: "Whatever is needed for a correct inference is fully expressed." Talking about this project, Frege says: "Right from the start I had in mind the expression of a content . . . But the content is to be rendered more exactly than is done by verbal language . . . Speech often only indicates by inessential marks or by imagery what a concept-script should spell out in full."<sup>61</sup> The concept-script is a formal language for the explicit codification of conceptual contents. In the preface to the *Begriffsschrift*, Frege laments that even in science, concepts are formed haphazardly, so that the ones employing them are scarcely aware of what they mean, of what their content really is. When the correctness of particular inferences is at issue, this sort of unclarity may preclude rational settlement of the issue. What is needed, he thinks, is a notation within which the rough-and-ready conceptual contents of the sciences, beginning with mathematics, can be reformulated so as to wear their contents on their sleeves. His explanatory target avowedly concerns a sort of inference, not a sort of truth, and the sort of inference involved must be content-conferring material inferences, whose propriety is determined before logical vocabulary comes on the scene, not the derivative formal ones whose propriety is underwritten by the use of that vocabulary.

Frege explicitly contrasts his approach with that of those, such as Boole, who conceive their formal language only in terms of formal inference, and so express no material contents: "The reason for this inability to form concepts in a scientific manner lies in the lack of one of the two components of which every highly developed language must consist. That is, we may distinguish the formal part . . . from the material part proper. The signs of arithmetic correspond to the latter . . . In contrast, Boole's symbolic logic only represents the formal part of the language."<sup>62</sup> Frege's own project is to express the contents that make up the material part of the language, not just the "formal cement that can bind these stones together": "My concept-script has a more far-reaching aim than Boolean logic, in that it strives to make it

possible to present a content when combined with arithmetical and geometrical signs . . . It is in a position to represent the formation of the concepts actually needed in science.”<sup>63</sup> It is the wider domain to which his expressive ambition extends that Frege sees as characteristic of his approach.

Since contents are determined by inferences, expressing inferences explicitly will permit the expression of any sort of content at all: “It seems to me to be easier still to extend the domain of this formula language to include geometry. We would only have to add a few signs for the intuitive relations that occur there . . . The transition to the pure theory of motion and then to mechanics and physics could follow at this point.”<sup>64</sup> Indeed, he goes on to suggest that for this reason, “by laying bare the misconceptions that through the use of language often almost unavoidably arise concerning the relations between concepts and by freeing thought from that with which only the means of expression of ordinary language, constituted as they are, saddle it . . . my ideography [*Begriffsschrift*], further developed for these purposes, can become a useful tool for the philosopher.”<sup>65</sup>

### 7. Frege’s Expressive Conception of Logic

Frege’s early understanding of logic offers some specific content to the notion of explicitly expressing what is implicit in a conceptual content. That is what is required to fill in a notion of expressive or elucidating rationality that might be laid alongside (and perhaps even be discovered to be presupposed by) notions of rationality as accurate representation, as logically valid inference, and as instrumental practical reasoning. Before he makes the fateful step from seeing logic as an attempt to codify inferences to seeing it as the search for a special kind of truth (which Dummett bemoans, and to which we owe much of contemporary logic), Frege’s aim is to introduce vocabulary that will let one say (explicitly) what otherwise one can only do (implicitly). Consider the conditional, with which the *Begriffsschrift* begins. Frege says of it: “The precisely defined hypothetical relation between contents of possible judgments has a similar significance for the foundation of my concept-script to that which identity of extensions has for Boolean logic.”<sup>66</sup> Prior to the introduction of such a locution, one could *do* something, one could treat a judgment as having a certain content (implicitly attribute that content to it), by endorsing various inferences involving it and rejecting others. After conditional locutions have been introduced, one can *say*, as part of the content of a claim, that a certain inference is acceptable. One is able to make explicit material inferential relations between an antecedent or premise and a consequent or conclusion. Since according to the inferentialist view of conceptual contents, it is these implicitly recognized material inferential relations that conceptual contents consist in, the conditional permits such contents to be explicitly expressed. If there is a disagreement about the goodness of an inference, it is possible to say what the dispute

is about and to offer reasons one way or the other. The conditional is the paradigm of a locution that permits one to make inferential commitments explicit as the contents of judgments.

The conditional ("the precisely defined hypothetical relation between contents of possible judgments"), rather than inclusion relations among extensions of concepts, plays the central role in Frege's logic because of two cardinal features of his view that distinguish it from the modern set-theoretic interpretations that develop from Boole's approach. First, he understands the content of nonlogical concepts in terms of their inferential role, rather than in terms of their extensions. Second, he understands the task of logical vocabulary to be expressing explicitly what is implicit in those material conceptual contents. What is implicit in those contents, according to the first or inferentialist commitment, is proprieties of inference. Making what follows from what explicit, as itself a *judgeable* content, one that can itself appear as a premise or conclusion in inference, is exactly the job of the conditional.

Frege's overall project for his *Begriffsschrift* is to use conditionals to make it possible to *say* explicitly what the inferential role of ordinary, nonlogical concepts is. Where, as he thinks is often the case in natural language, the content expressed by words is unclear, the project of expressing them explicitly will show where they need or can use clarification. The project is the rectification of concepts: clarifying them by explicating their contents. It is *saying* what their inferential role is: what follows from the applicability of each concept and what its applicability follows from. Employing the explicating logical locutions of which the conditional is the paradigm is to enable what Frege calls "the scientific formation of concepts." Such concepts will wear their contents on their sleeves; the inferential proprieties in virtue of which they mean what they mean are written down for all to read. The particular sciences can then proceed with their reasonings according to the same standards of rigor in the definition and use of their concepts that nineteenth-century mathematics finally came to aspire to. Although the application of this expressive methodology to the special case of mathematics always was closest to Frege's heart and occupied the greatest part of his energies, right from the beginning he had wider expressive ambitions. So the later writings on the sorts of content to be associated with nonmathematical concepts ("On Sense and Reference" prime among them) ought not to be seen to represent any change of interest or detour from his primary project.

Frege is not as explicit about the role of materially correct inferences as Sellars is, but his commitment to the notion is clear from the relation between two of the views that have been extracted from the *Begriffsschrift*: semantic expressivism about logic and inferentialism about content. Expressivism about logic means that Frege treats logical vocabulary as having a distinctive expressive role—namely making explicit the inferences whose goodness is implicit in the conceptual contents of nonlogical concepts. Infer-

entialism about those conceptual contents is taking them to be identified and individuated by their inferential roles. Together these views require that it be coherent to talk about inference prior to the introduction of specifically logical vocabulary, and so prior to the identification of any inferences as good in virtue of their logical form.<sup>67</sup> In the context of an inferential understanding of conceptual contents, an expressivist approach presupposes a notion of nonlogical inference, the inferences in virtue of which concepts have non-logical content. So the early Frege envisages a field of material inferences that confer conceptual content on sentences caught up in them. Although Frege does not offer an explanation of the concept, in the *Begriffsschrift* his expressive, explicating project commits him to something playing the role Sellars later picks out by the phrase "material inference."

There is a sense, then, in which the early Frege does see endorsement of conditional judgments as *implicit* in endorsement of the correctness of inferences. It is implicit in exactly the sense that what one is committed to by endorsing an inference as correct, and so by associating a certain conceptual content (that is inferential role) with the premises and conclusion, can be *made explicit* by expressing it in the form of a conditional judgment. The point of introducing logical vocabulary is precisely to make it possible to trade hitherto merely implicit inferential commitments for explicit assertional commitments to conditionals. And the payoff from expressing explicitly (in the form of judgments) the content-constitutive commitments that were implicit in prior inferential practice is the clarification and rectification of those conceptual contents. Formalism about inference—denying the existence of materially good inferences by assimilating all good inferences to logically good inferences, understanding all proprieties of inference as always already underwritten by logical form—turns things on their head. It misses the point of the process of explication that Frege puts at the center of the logical enterprise. It is a form of intellectualism, platonism, or *regulism* in the sense defined in the previous chapter. For it sees *rules* or *principles* as already standing behind every propriety of (in this case inferential) practice.

Frege's primary interest is in the process of explication: of expressing what is implicit in a practice, formulating it as an explicit rule or principle. This pragmatist project of explaining how knowing-that is founded on knowing-how, of explaining the codification in (conditional) principles of (inferential) practice, is unintelligible from any theoretical standpoint that acknowledges only the explicit form of propriety. Frege's fundamental insight into the expressive role of logical vocabulary (above all the conditional) is not incompatible with claiming that commitment to a conditional is implicit in endorsement of an inference, provided one is careful about what is meant by 'implicit'—provided, that is, that it is understood as making reference to the possibility of engaging in the substantive activity of making it explicit in the form of a claim or principle. But if one goes on to treat all reasoning as explicitly involving *detachment* from conditionals, and there-

fore implicit endorsement of logical truths involving conditionals (including tautologies involving nested or iterated conditionals), then the line has been crossed and Frege's expressive insight has been lost. As should become clearer from the discussion of Section V below, one of the casualties of the inferential formalist's inversion of the significance of the role of conditionals in making explicit in the form of a principle what is implicit in an inferential practice is a proper understanding of the way in which the contents of inferentially articulated concepts evolve and are clarified as they are expressed with the help of logical locutions.

### *8. Expressive Completeness and the Two-Valued Conditional*

Various special features of Frege's presentation of his conditional, and of the use he goes on to make of it in the *Begriffsschrift*, tend to obscure the crucial expressive role in explicating inferences (and therefore conceptual contents) that he assigns to it. These are picked up and emphasized, to some extent even in his own later work, but especially in the subsequent logistical tradition to which he gave birth, and make it difficult to work back to an appreciation of his original logical project. The difficulties stem from his use of the now-classical two-valued conditional.

It is difficult now to read the definition by which he introduced his conditional (in the fifth paragraph of the *Begriffsschrift*) without being blinded by hindsight—in particular by the glare of the truth-tabular tautology formulation presented by Wittgenstein in the *Tractatus*. Frege does define what has come to be called (by the lights of this work ludicrously inappropriately) the "material" conditional. He does so, however, not in terms of a semantic distinction between judgeable contents that are true and those that are false, but rather in terms of a pragmatic distinction between those that are affirmed (*bejaht*) and those that are denied (*verneint*). This is his invariable practice in the *Begriffsschrift*, although in later years he is happy enough to recast these claims in terms of truth (as part of the reorientation of his thought toward logical truth that Dummett rightly complains about). Putting things in terms of truth rather than affirmation pushes into the background (though it does not abolish) the way in which the semantic notion of content is beholden to the pragmatic notion of force, in the explanation of which it serves, and which is the source of the priority of judgeable contents and so, even in the later work, of the special central and ineliminable role played by the True as *Bedeutung*. It is worth recalling in this connection Frege's formulation of his view in 1915, already quoted in Section II above: "'True' only makes an abortive attempt to indicate the essence of logic, since what logic is really concerned with is not contained in the word 'true' at all but in the assertoric force with which a sentence is uttered . . . the thing that indicates most clearly the essence of logic is the assertoric force with which a sentence is uttered."<sup>68</sup>

In fact Frege's view is that 'true' is a bit of logical vocabulary, which serves to express explicitly what is done implicitly in asserting. This is why by the time of the *Grundgesetze* (1893) he has adopted a regimentation in which all claims are expressed explicitly in the form of identities that have a sentence on one side and the canonical name 'the True' on the other. It is with that regimentation in mind, in turn, that he claims that the True is an object that must be recognized, at least implicitly, by anyone who makes judgments at all. (Identity claims are explained as the explicit expression of "recognition judgments" in the *Grundlagen*; see 7.1 below.) Assimilating all assertions to assertions of identities permits the use of his (ultimately substitutional) semantics for identity statements (forwarded in "*Über Sinn und Bedeutung*") in general application to all claims, which is his strategy in the *Grundgesetze*.

The key point is that explication is not explanation. Proprieties of inference are not *explained* in terms of something more primitive by being expressed in the explicit form of claims by the use of conditionals; the force of asserting or judging is not *explained* by expressing it explicitly as a saying of a sentence that it is (a name of the) true. This is why Frege always insists that truth is indefinable, something the understanding of which is always already implicit in claiming. "The True" is not a name whose sense one can grasp first, and then appeal to in explaining what it is to make a claim; its use merely makes explicit what is implicit in claiming. It has an expressive, not an explanatory role. Thinking of it the other way around is making a mistake with respect to 'true' and claiming that is strictly analogous to the inferential formalist's mistake regarding the conditional and inferring. A version of the preferred pragmatic direction of explanation is presented below, where Chapter 5 discusses the role of 'true' in terms of the expressive, explicating function it performs with respect to claiming, according to the account of that practice offered in Chapter 3.

Bracketing subtleties regarding the relation between truth and commitment or affirmation, the fact remains that the conditional Frege actually defines and employs rules out only the case in which the consequent is denied or taken to be false while the antecedent is affirmed or taken to be true. This form of conditional, whatever its compositional virtues, is an extremely impoverished resource for the expression of proprieties of inference. The job that has been attributed to the conditional for Frege is that a conditional be affirmable or taken to be true just in case the inference from its antecedent to its consequent is endorsed or taken to be correct. Using the two-valued conditional to establish the connection between the correctness of an inference and the truth or endorsement of the claims that are its premises and conclusions has unpalatable results. Frege clearly has in mind a fundamental semantic principle regarding this connection: a good inference never takes one from premises that are true to a conclusion that is not true. This is a way of thinking about inferences as commitment-preserving: if one is committed to the premises of a good inference, in the sense of taking them

to be true (the sense that matters for assertion and judgment), then one is committed in the same sense to the conclusion.<sup>69</sup>

Such a principle could be agreed to both by those who adopt the traditional order of semantic explanation—by understanding the principle as explicating the correctness of inference in terms of a prior notion of truth (or taking-true)—and by those who adopt the converse order of semantic explanation (pursued in this work)—by taking truth or the sort of commitment involved in taking-true (and hence in asserting and judgment) to be explicated as what is preserved by good inferences.<sup>70</sup> But in either case, that truth or commitment is preserved by an inference ought to be taken to be a *necessary* condition of its being a good inference, not a *sufficient* condition. Affirming or taking-true both the claim that Hegel was Hölderlin's roommate and the claim that 43 is prime, and so being committed to the inference from the one claim to the other preserving truth and commitment, does not involve endorsing the propriety of that inference.

The two-valued conditional is subject to this familiar sort of complaint about fallacies of irrelevance precisely because the inferences it codifies explicitly are those that result from implausibly treating the plausible semantic preservation principle as, not only a necessary condition of good inference, but also as a sufficient one. It follows that the two-valued conditional Frege actually defines is an alarmingly bad choice for making explicit actual proprieties of inference. That fact in turn seems to cast doubt on the expressive understanding of his project. If he really wants logical vocabulary to make inferences explicit—because he wants to make nonlogical contents explicit and understands them in inferential terms—why does he employ the blunt, crude tool that is the two-valued conditional, whose expressive powers are hopelessly inadequate for the task of expressing the material inferences that might plausibly be identified with conceptual contents?

The answer is that although he hopes eventually to be able to use logical vocabulary to make explicit the inferential involvements in virtue of which nonlogical claims have the conceptual contents they do, the task Frege actually undertakes in the text of the *Begriffsschrift* is much less ambitious. The only concepts whose inferential role he actually makes explicit there are the logical concepts themselves, and those mathematical concepts that turn out to be definable from them. The concepts of geometry and mechanics—and indeed, the rest of the nonlogical concepts that philosophers might be interested in clarifying by expressing them explicitly—are to be *expressible* by means of logical vocabulary, together with other primitive signs. They are not understood to be, as some of the mathematical concepts (but not, for instance, those of geometry) are, *definable* by means of the logical vocabulary alone.

The first stage of Frege's grand project of clarification of nonlogical concepts through their explication in logical terms is to make explicit the conceptual contents of the logical expressions that are to be employed in that project. These concepts must themselves be “formed scientifically.” This is

why he is proud to display, for each of the official propositions of the *Begriffsschrift* (couched entirely in logical vocabulary), what it follows from (in the proof of the proposition) and what follows from it (in the appendix). Doing so specifies the inferential role of those propositions, and so, indirectly, the conceptual content of the subsentential logical vocabulary that occurs in them.

Thus the only inferences Frege makes explicit in the *Begriffsschrift* are the inferences that are good in virtue of their logical form—for these determine the conceptual content (in his sense) of his logical vocabulary. He finds, in the two-valued conditional, an expressive equilibrium: the inferences in virtue of which that conditional means what it means can themselves be expressed and codified by the use of that conditional. Frege's logical vocabulary is potentially (and he makes it actually) self-explicating.<sup>71</sup> The official propositions of the *Begriffsschrift* explicitly specify the inferential roles of the logical vocabulary, and the inferential roles of those propositions can be expressed explicitly in terms of that vocabulary. Fascinated by how much of mathematical vocabulary turns out to be logical vocabulary in this sense, Frege does not in this work pursue the question of the expressive adequacy of his conditional for material, nonlogical, conceptual contents. He devotes most of the rest of his life to exploring the conceptual contents that can be made explicit by the use of this extensional conditional.<sup>72</sup>

The results he achieves with the poor expressive resources of the two-valued conditional deserve our awe and admiration. Nevertheless, the motivations remain for the grander semantic expressive aspirations that the young Frege contributes to the inferentialist tradition. It was pointed out above that distinguishing a privileged class of good inferences as good in virtue of their logical form, that is, as logically valid inferences, requires being able to pick out some vocabulary as distinctively *logical* vocabulary. Then the logically valid inferences are just those good inferences that remain good on all substitutions of nonlogical for nonlogical vocabulary. The demarcational question of how logical locutions ought to be identified has received various influential answers. The current suggestion is that Frege's early work is predicated on the idea that what distinguishes vocabulary as specifically *logical* is its expressive role in making conceptual content explicit. Vocabulary deserves the appellation 'logical' just in case it serves to make *explicit*, as the content of a claim, proprieties concerning the use of the expression that otherwise remain *implicit* in practice, specifically the proprieties in virtue of which it has the conceptual content that it does. It is because Frege understands those content-conferring practical proprieties to be in the first instance proprieties of *inferential* practice that the paradigmatic sentential logical locution for him is the conditional. One of the central tasks of the rest of this work is to show how this semantic expressive paradigm can be extended to other logical and semantic locutions.

In the next chapter it is argued that a key link connecting the implicit

norms governing the use of expressions with the conceptual content those practices confer on them is provided by the notion of the *incompatibility* of commitments. In practical terms of normative status, to treat  $p$  and  $q$  as incompatible claims is to take it that commitment to one precludes entitlement to the other. Practices properly articulated to be interpretable as instituting the normative statuses of commitment and entitlement required for incompatibility relations are thereby interpretable as conferring semantic content on the states, attitudes, and performances that stand in incompatibility relations. The content of a claim can be represented by the set of claims that are incompatible with it. For instance, a relation of entailment, required for an inferential semantics, can be derived according to the principle that  $p$  entails  $q$  just in case everything incompatible with  $q$  is incompatible with  $p$ . The formal semantics generated by such incompatibility interpretations is quite rich. It has been shown, for instance, how to represent classical logic, relevance logic, and various systems of orthologic (or quantum logic) by constraints on incompatibility relations.<sup>73</sup>

*Negation*, as a logical connective supporting formally valid inferences, plays the same explicating role with respect to material *incompatibility* relations among judgeable (that is propositional) contents that the *conditional* plays with respect to material *inferential* relations among such contents. The formal negation of a claim is constructed as its minimal incompatible, the claim that is entailed by each one of the claims incompatible with the claim of which it is the negation. Thus in the context of a conditional that makes entailment relations explicit, the introduction of a locution playing the inferential role of negation makes it possible to make explicit the relation of material incompatibility between claims. To assert that  $p$  is incompatible with  $q$ , one asserts the conditional whose antecedent is  $p$  and whose consequent is the negation of  $q$ . Conjunction and disjunction can be handled straightforwardly as corresponding to Boolean operations on the sets of incompatibles that represent conceptual contents according to this sort of semantic model.

Chapter 7 below discusses Frege's treatment of identity locutions as making explicit the substitution-inferential commitments that are implicit in the use of singular terms. It also shows how that idea can be extended to an account of the use of quantifiers as making explicit the different sort of substitution-inferential commitment that is implicit in the use of predicates. The job of the next chapter is to offer an account of the normative practices of claiming and judging, and of the propositional contents conferred on states, attitudes, performances, and expressions by their playing appropriate roles in those practices. This account gives a definite sense to the notion of explicit *sayings*, in terms of norms implicit in *doings*. What is explicit is then the propositional content that is said or believed. In this fundamental sense, H. L. Mencken makes the content of his thought explicit, and expresses it fully, by asserting the declarative sentence: "Natives of Appalachia are clay-eating

sub-humans." But it is also possible to use logical vocabulary to make explicit expressively essential inferential involvements that remain implicit in the concepts employed in making this claim.

In subsequent chapters various other locutions are introduced as being used so as to make explicit, in this sense, some feature of the practices that originally confer propositional content (so that having such contents can be understood to consist in how it is correct for those locutions to be used, according to the practices in question). Not only the standard logical vocabulary, but also traditional semantic vocabulary such as 'true', 'refers', and the 'of' of intentional aboutness, should be understood as semantically explicating. The point of using these sorts of expressions is to make explicit as the contents of claims (whose consequences can be explored and which can be justified and disputed) some critical element of the practices of talking and believing in virtue of which it is possible to interpret anything as propositionally contentful in the first place. Furthermore, another range of expressions, including such locutions as 'claims that', 'believes that', 'intends that', and normative talk of commitments and entitlements, is interpreted as *pragmatically* explicating. The point of using these sorts of expressions is to make explicit as the contents of claims some of the pragmatic elements of the practices of talking, believing, and acting that confer propositional contents. One thread running through the later chapters of this work is the attempt to achieve an analog of the expressive equilibrium Frege achieves in the propositional fragment of the *Begriffsschrift*. The challenge is to show how not only the semantics, but the pragmatics outlined in the first four chapters can be made explicit, in terms of vocabulary that is introduced by specifying practices of using it that are sufficient to confer on it the content that is then employed in making explicit precisely those practices and that content. The ideal is that the theory should specify practices sufficient to confer on the various locutions considered all the kinds of content required to state the theory itself.

## V. CIRCUMSTANCES AND CONSEQUENCES OF APPLICATION

### 1. Dummett's Model

The previous section of this chapter introduced three themes: that conceptual content is to be understood in terms of role in reasoning rather than exclusively in terms of representation, that the capacity for such reasoning is not to be identified exclusively with mastery of a logical calculus, and that besides theoretical and practical reasoning using contents constituted by their role in material inferences, there is a kind of *expressive* rationality that consists in making implicit, content-conferring inferential commitments explicit as the contents of assertional commitments. Being

rational in the primary sense is having states and attitudes and producing performances that have propositional contents. The next chapter discusses how the inferential articulation essential to such contents is conferred on them by the way in which the states, attitudes, and performances exhibiting those contents are caught up in the game of giving and asking for reasons. Rationality consists in mastery of those practices. It is not to be understood as a logical capacity. Rather, specifically logical capacities presuppose and are built upon underlying rational capacities. The fundamental characteristic role of logical vocabulary is to make it possible to talk and think *explicitly* about the inferentially articulated semantic contents *implicitly* conferred on expressions (among other things) by their role in rational practice. The optional introduction of sophisticated logical explicating vocabulary has an expressive point and payoff. By its means the material inferential practices, which govern and make possible the game of giving and asking for reasons, are brought into that game (and so into consciousness) as explicit topics of discussion and justification. In this way, in the context of the three basic themes mentioned above, an *expressive* understanding of logic was introduced—according to which formal validity of inferences is substitutionally defined in terms of material correctness of inferences together with the discrimination of some privileged vocabulary; that privileged vocabulary is identified as logical vocabulary; and what it is for something to be a bit of logical vocabulary is explained in terms of its semantically expressive role.

These ideas, to be found in the early works of Frege and Sellars, provide the beginnings of the structure within which modern inferentialism develops. The approach they suggest can be made more definite by considering a general model of conceptual contents as inferential roles that has been recommended (in somewhat different terms) by Dummett. According to that model, the use of any linguistic expression or concept has two aspects: the circumstances under which it is correctly applied, uttered, or used, and the appropriate consequences of its application, utterance, or use. Though Dummett does not put the point this way, this model connects to inferentialism of the Sellarsian sort via the principle that part of the content to which one is committed by using the concept or expression may be represented by the material inference one implicitly endorses by such use: the inference from the circumstances of appropriate employment to the appropriate consequences of such employment.

The original source for the model lies in a treatment of the grammatical category of sentential connectives. Dummett's two-aspect model is a generalization of a standard way of specifying the inferential roles of logical connectives, due ultimately to Gentzen. Gentzen defined connectives by specifying *introduction rules*, or inferentially sufficient conditions for the employment of the connective, and *elimination rules*, or inferentially necessary consequences of the employment of the connective. So, 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 thereby to count also as committed to  $p \& q$ , and that anyone who is committed to  $p \& q$  is thereby committed both to  $p$  and to  $q$ . The first schema specifies, by means of expressions that do not contain the connective, the *circumstances* under which one is committed to claims expressed by sentences that contain (as principal connective)<sup>74</sup> the connective whose inferential role is being defined, that is, the sets of premises that entail them. The second schema specifies, by means of expressions that do not contain the connective, the *consequences* of being committed to claims expressed by sentences that contain (as principal connective) the connective whose inferential role is being defined, that is, the sets of consequences that they entail.

Dummett makes a remarkable contribution to inferentialist approaches to conceptual content by showing how this model can be generalized from the case of logical connectives to provide a uniform treatment of the meanings of expressions of other important grammatical categories: sentences, predicates and common nouns, and singular terms. The application to the *propositional* contents expressed by whole declarative sentences is straightforward. What corresponds to an introduction rule for a propositional content is the set of inferentially sufficient conditions for asserting it, and what corresponds to an elimination rule is the set of inferentially necessary consequences of asserting it, that is, what follows from doing so. Dummett says: “Learning to use a statement of a given form involves, then, learning two things: the conditions under which one is justified in making the statement, and what constitutes acceptance of it, i.e., the consequences of accepting it.”<sup>75</sup> Dummett presents his model as specifying two fundamental features of the use of linguistic expressions. In what follows, it is applied in the context of the previous ideas, to bring into relief the implicit material inferential content a concept or expression acquires in virtue of being used in the ways specified by these two “aspects.” The link between *pragmatic significance* and *inferential content* is supplied by the fact that asserting a sentence is (among other things) implicitly undertaking a commitment to the correctness of the material inference from its circumstances to its consequences of application.

Dummett applies his model exclusively to conceptually contentful *linguistic* expressions. But it is clear that the model has a wider application—to intentional states and attitudes generally. Dummett’s model just provides some structure to the representation of the functional roles of intentional states. For instance, one could think about the functional role played by a belief with a particular propositional content in terms of the circumstances in which it is appropriate to acquire a belief with that content, and the appropriate consequences of such acquisition. In the next two chapters this basic Dummettian structure is further articulated, to allow a richer representation of the functional roles of states, performances, and expressions.

## 2. Inferential Connection between Even Noninferential Circumstances and Consequences of Application

The concepts least easily assimilated to an inferential model are the empirical concepts whose core employment is in perception and the formulation of observation reports. For such reports are essentially *noninferential*—in the sense that they are elicited as responses to features of the largely nonlinguistic environment, rather than as conclusions drawn from other claims. Their content accordingly derives (at least in large part) from the *reliable differential responsive dispositions* that those who have mastered the concepts exhibit with respect to their application. Such concepts can be assimilated to the inferentialist understanding of conceptual contents by adapting Dummett's idea of distinguishing two crucial features of the use of linguistic expressions: their *circumstances* of appropriate application, and the appropriate *consequences* of such application. In terms of this model, it is possible to understand the use of *any* expression as implicitly involving an inferential commitment. In particular, by using the expression, one is (among other things) committed to the propriety of the inference from its circumstances to its consequences of application. The consequences of application are always themselves inferentially related to the concept in question (although the inference involved may include practical inferences, whose conclusions are commitments to act). The circumstances of application need not themselves be linguistic. For the concept *red*, for instance, they include the presence of visibly red things. Nonetheless, even the use of concepts of this sort can be seen to embody inferential commitments, to the propriety of applying inferential consequences of *red*—for instance, *colored*—to anything that *red* is properly applied to.<sup>76</sup>

It is in this way that a broadly inferential approach can incorporate into its conception of the contents of empirical concepts the nonlinguistic circumstances in which they are correctly noninferentially applied. Thus the concepts *water* and its twin-earth analog *twater*, which are by hypothesis alike except that one is appropriately applied to  $H_2O$  and the other to *XYZ*, count as involving different *inferential* contents. This is so even though they fund inferential moves involving the same sorts of *noises*, from saying "That's water" to "That's liquid," for instance. For they involve different circumstances of appropriate application, and hence different inferential transitions from those circumstances to their consequences. So even though it is the practices of those whose concepts they are that confer on them their contents, the earthlings and twin-earthlings need not be able to tell that they have different concepts, if water and *twater* are indistinguishable to them. They are not omniscient about the inferential commitments implicit in their own concepts. For the interpreter who is making sense of their practices, and who is able (not perceptually, but conceptually) to distinguish  $H_2O$  and *XYZ*, can understand transported earthlings as *mistaking* the *XYZ* they look at for

water, as *inappropriately* applying the concept they express with their word ‘water’ to that unearthly stuff. The circumstances of appropriate noninferential application of the concept expressed by the English word ‘water’ require that it be applied in response to a sample of  $H_2O$ .

In this way the circumstances of appropriate application of a claim can include not only other claims (from which the one in question could be inferred) but also perceptual circumstances (to which one has been trained to respond *noninferentially* by endorsing the target claim). The appropriate consequences of application of a claim can include not only the inferential acquisition of further beliefs whose contents follow from the contents of the belief in question but also, in the context of further contentful intentional states, the noninferential responsive performance of *actions*, under the descriptions by which they can be exhibited as the conclusions of practical inferences. This is explicitly acknowledged in the continuation of the passage from Dummett quoted above: “Here ‘consequences’ must be taken to include both the inferential powers of the statement and anything that counts as acting on the truth of the statement.” So the circumstances and consequences of application Dummett is talking about should not be identified with inferentially necessary and sufficient conditions, where this means identifying them with sets of claims or beliefs that are conclusions or premises of theoretical inferences involving the content in question. It turns out, however, that the circumstances and consequences model can be understood as an inferential model, regardless of whether the circumstances and consequences are themselves already thought of in inferential terms. The inferential element enters this picture in the commitment undertaken by one who employs a given content to the propriety of the transition from the circumstances of appropriate application to the appropriate consequences of application of a conceptual content. This will be construed as a broadly inferential commitment, though the detailed justification for this characterization will not emerge until the next chapter.<sup>77</sup>

One advantage of thinking about conceptual content as determining functional role specified in terms of proprieties governing circumstances and consequences of application is the room it makes for a *pragmatic* picture of understanding or grasping such a content. Understanding can be understood, not as the turning on of a Cartesian light, but as practical mastery of a certain kind of inferentially articulated *doing*: responding differentially according to the circumstances of proper application of a concept, and distinguishing the proper inferential consequences of such application. This is not an all-or-none affair; metallurgists understand the concept *tellurium* better than most of us do, for training has made them master of the inferential intricacies of its employment in a way that we can only crudely approximate. On this inferentialist rendering, thinking clearly is a matter of knowing what one is committing oneself to by a certain claim, and what would entitle one to that commitment. Writing clearly is providing enough clues for a reader to infer

what one intends to be committed to by each claim, and what one takes it would entitle one to that commitment. Failure to grasp either of these components is failure to grasp the inferential commitment that use of the concept involves, and so failure to grasp its conceptual content.

### 3. One-Sided Theories of Meaning

Verificationists, assertibilists, and reliabilists make the mistake of treating the first aspect as exhausting content. Understanding or grasping a content is taken to consist in practically mastering the circumstances under which one becomes entitled or committed to endorse a claim or acquire a belief, quite apart from any grasp of what one becomes entitled or committed to by such endorsement or acquisition. But claims can have the same circumstances of application and different consequences of application, as for instance "I foresee (or predict) that I will write a book about Hegel" and "I will write a book about Hegel" do. Any circumstances under which one is entitled to one of these claims (or to acquire the belief it expresses) are circumstances under which one is entitled to the other. (If this does not seem right for the actual concepts expressed by 'foresee' and 'predict', artificial variants clearly can be constructed for which it is.) Yet what follows from the claims is quite different. If I will write a book about Hegel, then I will write a book about Hegel. Yet that I will write a book about Hegel does not follow from my foreseeing or predicting that I will (as the sad history of orphaned first volumes of ambitious projects attests). The consequences of these claims are quite different. Examples meeting the conditions required for this point are forthcoming in any idiom expressively rich enough to contain *pragmatically explicitating locutions*, which permit one to *say* what one is *doing* in performing a certain speech act or acquiring a certain state or attitude—for instance "I claim that *p*" or "I believe that *p*." (These locutions are discussed in Chapter 8.)

In any idiom expressively rich enough to contain *semantically explicitating locutions*, whose paradigm is the conditional, the difference in inferential consequences of application between the sentence whose utterance performs a speech act (a doing in which the force is left implicit) and the sentence whose utterance explicitly *says* that that is what one is doing (so that force becomes part of the content) itself becomes explicit in the use of conditionals with those sentences as antecedents. The circumstances of appropriate application or assertibility conditions of the conditionals "If I will write a book about Hegel, then I will write a book about Hegel" and "If I foresee (or predict) that I will write a book about Hegel, then I will write a book about Hegel" are quite different. The assertibility of the second conditional, but not the first, depends on auxiliary hypotheses about how good at foreseeing or predicting I am. So for idioms that contain both the pragmatically explicitating locutions that permit the construction of pairs of sentences with identical

circumstances of application and different consequences of application, and the semantically explicating locutions that permit the construction of conditionals whose circumstances of application differ depending on the consequences of application of their antecedents, it is possible to show the inadequacy of a semantics that avails itself only of assertibility conditions or circumstances of appropriate application. For examples of the sort just considered show that substituting another sentence with the same assertibility conditions for a sentence that is the antecedent of a conditional can alter the assertibility conditions of the compound. In such an expressively rich environment, then, assertibility conditions cannot provide an adequate model of what Dummett calls "*ingredient content*," the contribution the occurrence of a sentence makes to the use of sentences in which it appears as a component. But this fact simply reflects the inadequacy of the model for the expression of conceptual content as inferential role, even in the more expressively impoverished idioms in which the pragmatically and semantically explicating locutions are not available.<sup>78</sup>

The inadequacy of a notion of semantic content that is restricted to circumstances of application to the exclusion of consequences of application has already appeared in another guise above. The point of the discussion of Sellars's application of inferentialist ideas to the understanding of noninferential reports, in Section III, was that parrots and photocells and so on might reliably discriminate the circumstances in which the concept *red* should be applied, without thereby grasping that concept. This would happen precisely in the case where they have no mastery of the consequences of such application—when they cannot tell that it follows from something being red that it is colored, that it is not a prime number, and so on. You do not convey to me the content of the concept *gleeb* by supplying me with an infallible gleeblness tester that lights up when and only when exposed to gleebl things. I would in that case know what things were gleebl, without knowing what I was saying about them when I called them that, what I had found out about them or committed myself to. Dummett offers two examples of philosophically important concepts where it is useful to be reminded of this point: "An account, however accurate, of the conditions under which some predicate is rightly applied may thus miss important intuitive features of its meaning; in particular, it may leave out what we take to be the point of our use of the predicate. A philosophical account of the notion of truth can thus not necessarily be attained by a definition of the predicate 'true', even if one is possible, since such a definition may be correct only in the sense that it specifies correctly the application of the predicate, while leaving the connections between this predicate and other notions quite obscure."<sup>79</sup> Even more clearly: "A good example would be the word 'valid' as applied to various forms of argument. We might reckon the syntactic characterization of validity as giving the criterion for applying the predicate 'valid' to an argument, and the semantic characterization of validity as giving the consequences of such an application . . . If [a student] is taught in a very unimaginative way,

he may see the classification of arguments into valid and invalid ones as resembling the classification of poems into sonnets and non-sonnets, and so fail to grasp that the fact that an argument is valid provides any grounds for accepting the conclusion if one accepts the premises. We should naturally say that he had missed the point of the distinction."

Pragmatists of the classical sort, in contrast, make the converse mistake of identifying propositional contents exclusively with the consequences of endorsing a claim: looking downstream to the claim's role as a premise in practical reasoning and ignoring its proper antecedents upstream. The fact that the pragmatist's emphasis is on practical consequences is not relevant to this complaint. The problem is that one can know what follows from the claim that someone is responsible for a particular action, that an action is immoral or sinful, that a remark is true or in bad taste, without for that reason counting as understanding the claims involved, if one has no idea when it is appropriate to make those claims or apply those concepts. Being classified as AWOL does have the consequence that one is liable to be arrested, but the specific circumstances under which one acquires that liability are equally essential to the concept.

It was pointed out that Frege's practice in the *Begriffsschrift* is to specify both the circumstances and the consequences of application of his claims, which in the context of that project (excluding as it does concepts with empirical and practical content deriving from their relation to perception and action) can be identified with the inferentially sufficient premises from which they follow and the inferentially necessary conclusions they lead to. Yet his official definition of conceptual content refers only to consequences, and Carnap follows him in this regard. For the special sort of concepts they are concerned with, where only inferential circumstances and consequences are in play, this restriction does not amount to ignoring circumstances of application. Restricting consideration for the sake of an example to one-premise inferences, associating with each sentence the set of sentences that follow from it determines for each sentence which sentences it follows from. So at the global level, nothing is lost by officially defining content in terms of inferential consequences alone. As will emerge below, it is quite otherwise when one is concerned locally with the content associated with each sentence—for instance in asking what it is to understand the content expressed by one sentence (but perhaps not another), or to alter the content expressed by one sentence, or to introduce a new content. Nor will the technical dodge of restriction of content to consequences be adequate when attention is turned to the sort of empirical and practical content concepts get from their involvement in perception and action.

#### 4. Conservativeness and the Coherence of Logical Concepts

Of course, such one-component theories do not simply ignore the aspects of content they do not treat as central. Dummett says:

Most philosophical observations about meaning embody a claim to perceive . . . a simple pattern: the meaning of a sentence consists in the conditions for its truth and falsity, or in the method of its verification, or in the practical consequences of accepting it. Such dicta cannot be taken to be so naive as to involve overlooking the fact that there are many other features of the use of a sentence than the one singled out as being that in which its meaning consists: rather, the hope is that we shall be able to give an account of the connection that exists between the different aspects of meaning. One particular aspect will be taken as central, as constitutive of the meaning of any given sentence . . . ; all other features of the use of the sentence will then be explained by a uniform account of their derivation from that feature taken as central.<sup>80</sup>

Pursuing this notion of derivation provides a helpful perspective on the idea of conceptual contents articulated according to material inferences, and on the role of explicit inference licenses such as conditional statements in expressing and elucidating such inferences and so such contents.

The strategy of attempting to *derive* one aspect of the use of an expression (or the significance of an intentional state) from another—in particular to derive appropriate consequences of application from circumstances of appropriate application, or vice versa—expresses Dummett's appreciation of the need for the semantic theorist to be able to explain two crucial features of our practices regarding conceptual contents. Concept-users are often confronted with decisions regarding *alternative* concepts and so are obliged to decide not only that certain uses of a given concept should be rejected as incorrect but also that certain concepts should themselves be rejected as inadequate or incorrect. We criticize our concepts and sometimes reject them. Furthermore, doing so is not simply a matter of free or arbitrary stipulation. Criticism of our concepts is constrained and sometimes compelled. These are important phenomena—an attempt to take proper account of them guides the discussion below. Dummett acknowledges them as motivating the theoretical acknowledgment of a need for *harmony* between the circumstances and consequences of application: “A naive view of language regards assertibility-conditions for a statement as exhausting its meaning: the result is to make it impossible to see how a meaning can ever be criticized, revised, or rejected; it was just such a naive view which led to the use of the notorious ‘paradigm-case argument’. An almost equally naive view is that which distinguishes the assertibility-conditions of a statement as its ‘descriptive meaning’ and its consequences as its ‘evaluative meaning’, dispensing with any requirement of harmony between them, but holding that we have the right to attach whatever evaluative meaning we choose to a form of statement irrespective of its descriptive meaning.”<sup>81</sup>

For the special case of defining the inferential roles of logical connectives by pairs of sets of rules for their introduction and for their elimination, which

motivates Dummett's broader model, there is a special condition it is appropriate to impose on the relation between the two sorts of rules. "In the case of a logical constant, we may regard the introduction rules governing it as giving conditions for the assertion of a statement of which it is the main operator, and the elimination rules as giving the consequences of such a statement: the demand for harmony between them is then expressible as the requirement that the addition of the constant to a language produces a conservative extension of that language."<sup>82</sup> Recognition of the appropriateness of such a requirement arises from consideration of connectives with "inconsistent" contents. As Prior pointed out, if a connective, which after Belnap may be called 'tonk', is defined as having the introduction rule proper to disjunction and the elimination rule proper to conjunction, then the first rule licenses the transition from  $p$  to  $p$  tonk  $q$ , for arbitrary  $q$ , and the second licenses the transition from  $p$  tonk  $q$  to  $q$ . The result is what he called a "runabout inference ticket," which permits any arbitrary inference.

Prior thought that this possibility shows the bankruptcy of Gentzen-style definitions of inferential roles. Belnap shows rather that when logical vocabulary is being introduced, one must constrain such definitions by the condition that the rule not license any inferences involving only old vocabulary that were not already licensed before the logical vocabulary was introduced.<sup>83</sup> That is, it must be ensured that the new rules provide an *inferentially conservative extension* of the original field of inferences. From the point of view of the joint commitments to understanding conceptual content in terms of material inference and conceiving the distinctive role of logical vocabulary as making those content-conferring inferential connections explicit in the form of claims, this constraint on the definition of logical particles by introduction and elimination rules makes perfect sense. For if those rules are not inferentially conservative, the introduction of the new vocabulary licenses new material inferences and so alters the contents associated with the old vocabulary. The expressive approach to logic motivates a criterion of adequacy for introducing logical vocabulary to the effect that no *new* inferences involving only the *old* vocabulary be made appropriate thereby. Only in this way can logical vocabulary play the expressive role of making explicit the original material inferences and so nonlogical conceptual contents.

##### 5. Nonlogical Concepts Can Incorporate Materially Bad Inferences

The problem of what Dummett calls a lack of "harmony" between the circumstances and the consequences of application of a concept can arise, however, not only for logical vocabulary but also for concepts with material contents. Seeing how it does provides further help in understanding the notion of expressive rationality and the way in which the explicating role of logical vocabulary contributes to the clarification of concepts. For conceptual change can be:

motivated by the desire to attain or preserve a harmony between the two aspects of an expression's meaning. A simple case would be that of a pejorative term, e.g. 'Boche'. The condition for applying the term to someone is that he is of German nationality; the consequences of its application are that he is barbarous and more prone to cruelty than other Europeans. We should envisage the connections in both directions as sufficiently tight as to be involved in the very meaning of the word: neither could be severed without altering its meaning. Someone who rejects the word does so because he does not want to permit a transition from the grounds for applying the term to the consequences of doing so. The addition of the term 'Boche' to a language which did not previously contain it would produce a non-conservative extension, i.e. one in which certain other statements which did not contain the term were inferable from other statements not containing it which were not previously inferable.<sup>84</sup>

This crucial passage makes a number of points that are worth untangling.

First of all, it shows how concepts can be criticized on the basis of substantive beliefs. If one does not believe that the inference from German nationality to cruelty is a good one, one must eschew the concept *Boche*. For one cannot deny that there are any Boche—that is just denying that anyone is German, which is patently false. One cannot admit that there are Boche and deny that they are cruel—that is just attempting to take back with one claim what one has committed oneself to with another. One can only refuse to employ the concept, on the grounds that it embodies an inference one does not endorse. (When the prosecutor at Oscar Wilde's trial asked him to say under oath whether a particular passage in one of his works did or did not constitute blasphemy, Wilde replied, "Blasphemy is not one of my words."<sup>85</sup>) Highly charged words like 'nigger', 'whore', 'Republican', and 'Christian' have seemed a special case to some because they couple "descriptive" circumstances of application to "evaluative" consequences. But this is not the only sort of expression embodying inferences that requires close scrutiny. The use of *any* concept or expression involves commitment to an inference from its grounds to its consequences of application. Critical thinkers, or merely fastidious ones, must examine their idioms to be sure that they are prepared to endorse and so defend the appropriateness of the material inferential transitions implicit in the concepts they employ. In Reason's fight against thought debased by prejudice and propaganda, the first rule is that material inferential commitments that are potentially controversial should be made explicit as claims, exposing them both as vulnerable to reasoned challenge and as in need of reasoned defense.

It is in this process that formal logical vocabulary such as the conditional plays its explicating role. It permits the formulation, as explicit claims, of the inferential commitments that otherwise remain implicit and unexam-

ined in the contents of material concepts. Logical locutions make it possible to display the relevant grounds and consequences and to assert their inferential relation. Formulating as an explicit claim the inferential commitment implicit in the content brings it out into the open as liable to challenges and demands for justification, just as with any assertion. In this way explicit expression plays an elucidating role, functioning to groom and improve our inferential commitments and so our conceptual contents—a role, in short, in the practices of reflective rationality that Sellars talks about under the heading of “Socratic method.”

But if Dummett is suggesting that what is wrong with the concept *Boche* is that its addition represents a nonconservative extension of the rest of the language, he is mistaken. Its nonconservativeness just shows that it has a substantive content, in that it implicitly involves a material inference that is not already implicit in the contents of other concepts being employed. This is no bad thing. Conceptual progress in science often consists in introducing just such novel contents. The concept *temperature* was introduced with certain criteria or circumstances of appropriate application and with certain consequences of application. As new ways of measuring temperature are introduced, and new consequences of temperature measurements adopted, the complex inferential commitment that determines the significance of using the concept of temperature evolves.

The proper question to ask in evaluating the introduction and evolution of a concept is not whether the inference embodied is one that is already endorsed (so that no new content is really involved) but whether that inference is one that ought to be endorsed. The problem with ‘Boche’ or ‘nigger’ is not that once we explicitly confront the material inferential commitment that gives them their content, it turns out to be *novel*, but that it can then be seen to be indefensible and inappropriate. We want to be aware of the inferential commitments our concepts involve, to be able to make them explicit, and to be able to justify them. But there are other ways of justifying them than showing that we were already implicitly committed to them, before introducing or altering the concept in question. Making implicit commitments explicit is only a necessary condition of justifying them.

Even in the cases where it does make sense to identify harmony of circumstances and consequences with inferential conservativeness, the attribution of conservativeness is always relative to a background set of material inferential practices, the ones that are conservatively extended by the vocabulary in question. Conservativeness is a property of the conceptual content only in the context of other contents, not something it has by itself. There can be pairs of logical connectives, either of which is all right by itself, but both of which cannot be included in a consistent system. It is a peculiar ideal of harmony that would be realized by a system of conceptual contents such that the material inferences implicit in every subset of concepts represented a conservative extension of the remaining concepts, in that no infer-

ences involving only the remaining ones are licensed that are not licensed already by the contents associated just with those remaining concepts. Such a system is an idealization because all of its concepts would already be out in the open, with none remaining hidden, to be revealed only by drawing conclusions from premises that have never been conjoined before, following out unexplored lines of reasoning, drawing consequences one was not previously aware one would be entitled or committed to by some set of premises. In short, this would be a case where Socratic reflection—making implicit commitments explicit and examining their consequences and possible justifications—would never motivate one to alter contents or commitments. Such complete transparency of commitment and entitlement is in some sense an ideal projected by the sort of Socratic practice that finds current contents and commitments wanting by confronting them with each other, pointing out inferential features of each of which we were unaware. But as Wittgenstein teaches in general, it should not be assumed that our scheme is like this, or depends upon an underlying set of contents like this, just because we are obliged to remove any particular ways in which we discover it to fall short.

These are reasons to part company with the suggestion, forwarded in the passage above, that inferential conservatism is a *necessary* condition of a "harmonious" concept—one that won't "tonk up" a conceptual scheme. In a footnote, Dummett explicitly denies that conservativeness can in general be treated as a *sufficient* condition of harmony: "This is not to say that the character of the harmony demanded is always easy to explain, or that it can always be accounted for in terms of the notion of a conservative extension . . . The most difficult case is probably the vexed problem of personal identity."<sup>86</sup> In another place, this remark about personal identity is laid out in more detail:

We have reasonably sharp criteria which we apply in ordinary cases for deciding questions of personal identity: and there are also fairly clear consequences attaching to the settlement of such a question one way or the other, namely those relating to ascriptions of responsibility, both moral and legal, to the rights and obligations which a person has . . . What is much harder is to give an account of the connection between the criteria for the truth of a statement of personal identity and the consequences of accepting it. We can easily imagine people who use different criteria from ours . . . Precisely what would make the criteria they used criteria for *personal identity* would lie in their attaching the same consequence, in regard to responsibility, motivation, etc., to their statements of personal identity as we do to ours. If there existed a clear method for deriving, as it were, the consequences of a statement from the criteria for its truth, then the difference between such people and ourselves would have the character of a factual disagreement, and one

side would be able to show the other to be wrong. If there were no connection between truth-grounds and consequences, then the disagreement between us would lie merely in a preference for different concepts, and there would be no right or wrong in the matter at all.<sup>87</sup>

Dummett thinks that there is a general problem concerning the way in which the circumstances and consequences of application of expressions or concepts ought to fit together. Some sort of "harmony" seems to be required between these two aspects of the use. The puzzling thing, he seems to be saying, is that the harmony required cannot happily be assimilated either to compulsion by facts or to the dictates of freely chosen meanings. But the options—matter of fact or relation of ideas, expression of commitment as belief or expression of commitment as meaning—are not ones that readers of "Two Dogmas of Empiricism" and its heirs ought to be tempted to treat as exhaustive.<sup>88</sup>

As already pointed out, talk of derivability is strictly stronger than talk of conservativeness. On the other side of the divide, the notion of a completely factual issue that Dummett appeals to in this passage is one in which the applicability of a concept is settled straightforwardly by the application of other concepts: the concepts that specify the necessary and sufficient conditions that determine the truth conditions of claims involving the original concept.<sup>89</sup> This conception, envisaged by a model of conceptual content as necessary and sufficient conditions, seems to require a conceptual scheme that is ideally transparent in the way mentioned above, in that it is immune to Socratic criticism. For that conception insists that these *coincide*—in that the individually sufficient conditions *already* entail the jointly necessary ones. Only then is it attractive to talk about content as truth conditions, rather than focusing on the substantive inferential commitments that relate the sufficient to the distinct necessary conditions, as recommended here. By contrast to the either/or that Dummett presents, in a picture according to which conceptual contents are conferred by being caught up in a social practical structure of inferentially articulated commitments and entitlements, material inferential commitments are a necessary part of any package of practices that includes material assertional or doxastic commitments. From this point of view, rendering conceptual content as truth conditions and thinking of them as necessary and sufficient conditions leaves out precisely the material content of concepts.

For the circumstances and consequences of application of a concept may stand in a substantive material-inferential relation. To ask what sort of "harmony" they should exhibit is to ask what material inferences we ought to endorse, and so what conceptual contents we ought to employ. This is not the sort of a question to which one ought to expect or even welcome a general or wholesale answer. Grooming our concepts and material inferential commitments in the light of our assertional commitments (including those we

find ourselves with noninferentially through observation) and the latter in the light of the former is a messy, retail business. Dummett thinks that a theory of meaning should take the form of an account of the nature of the "harmony" that ought to obtain between the circumstances and the consequences of application of the concepts we ought to employ. The present point is that one should not expect a theory of that sort to take the form of a specification of necessary and sufficient conditions for the circumstances and consequences of application of a concept to be harmonious. Rather, insofar as the idea of such a theory makes sense at all, it must take the form of an investigation of the ongoing elucidative process, of the "Socratic method" of discovering and repairing discordant concepts, which alone gives the notion of harmony any content. It is given content only by the process of *harmonizing* commitments, from which it is abstracted.

In Sellars's characterization of induction, introduced above, inductive inference is assigned an expressive role insofar as its conclusion is understood as being an inference license making explicit a commitment that is implicit in the use of conceptual contents antecedently in play. Rules of this sort assert an authority over future practice and for their entitlement answer both to the prior practice being codified and to concomitant inferential and doxastic commitments. In this way they may be likened to the principles formulated by judges at common law, intended both to codify prior practice, as represented by precedent, expressing explicitly as a rule what was implicit therein, and to have regulative authority for subsequent practice. The expressive task of making material inferential commitments explicit plays an essential role in the reflectively rational Socratic practice of harmonizing our commitments. For a commitment to become explicit is for it to be thrown into the game of giving and asking for reasons as something whose justification, in terms of other commitments and entitlements, is liable to question. Any theory of the sort of inferential *harmony* of commitments we are aiming at by engaging in this reflective, rational process must derive its credentials from its expressive adequacy to that practice, before it should be accorded any authority over it.

## 6. Varieties of Inferentialism

Section IV of this chapter introduced three related ideas:

1. the inferential understanding of conceptual content,
2. the idea of materially good inferences, and
3. the idea of expressive rationality.

These contrast, respectively, with

- 1'. an understanding of content exclusively according to the model of the representation of states of affairs,

- 2'. an understanding of the goodness of inference exclusively on the model of formal validity, and
- 3'. an understanding of rationality exclusively on the model of instrumental or means-end reasoning.

In this section these ideas were considered in relation to the representation of inferential role suggested by Dummett, in terms of the circumstances of appropriate application of an expression or concept and the appropriate consequences of such application. Both sections sought to introduce an *expressive* view of the characteristic role played by logical vocabulary and to indicate its relation to the practices constitutive of rationality.

One of the important benefits afforded by the emphasis in this section on understanding the Dummettian model of the use of linguistic expressions in terms of appropriate circumstances and consequences of application linked by an inferential commitment is the clarification it offers concerning the options that are open in working out an inferentialist approach to semantics. There are three different ways in which one might take inference to be of particular significance for understanding conceptual content. The weak inferentialist thesis is that inferential articulation is *necessary* for specifically *conceptual* contentfulness. The strong inferentialist thesis is that broadly inferential articulation is *sufficient* for specifically conceptual contentfulness—that is, that there is nothing more to conceptual content than its broadly inferential articulation. Dummett's model is particularly helpful for focusing attention on how important the qualification 'broadly' is in this formulation. For strong inferentialism as it is worked out in the rest of this project is *not* committed to the hyperinferentialist thesis, which maintains that *narrowly* inferential articulation is sufficient for conceptual contentfulness of all sorts.

The difference between the broad and the narrow conception of inferential articulation has three dimensions. First, and most important, the broad conception includes the possibility of noninferential circumstances and consequences of application. In this way (discussed in Chapter 4) the specifically *empirical* conceptual content that concepts exhibit in virtue of their connection to language entries in *perception* and the specifically *practical* conceptual content that concepts exhibit in virtue of their connection to language exits in *action* are incorporated into the inferentialist picture. The use of concepts with contents of these sorts can still be understood in terms of the material inferential commitment one who uses them undertakes: the commitment to the propriety or correctness of the inference from their circumstances to their consequences of application. Conceiving such inferences broadly means conceiving them as involving those circumstances and consequences, as well as the connection between them. The hyperinferentialist about conceptual content (adopting a position not endorsed here) would allow only *inferential* circumstances and consequences of application. Under

such a restriction, it is impossible to reconstruct the contents of actual concepts, except perhaps in some regions of mathematics.

Second, relations of *incompatibility* among claims and (so) concepts are considered broadly inferential relations, on grounds of their antecedents and their consequences. On the side of consequences, incompatibility relations underwrite the modal inferences codified by strict implication. For  $p$  entails  $q$  in this sense just in case everything incompatible with  $q$  is incompatible with  $p$ . So being a square entails being a rectangle, because everything incompatible with being a rectangle is incompatible with being a square. On the side of antecedents, the semantic relation of incompatibility will be understood (in the next chapter) in terms of the very same normative statuses of doxastic commitment and entitlement to such commitments, in terms of which inferences are construed (with commitment-preserving inferences corresponding roughly to deductive inferences, and entitlement-preserving inferences corresponding roughly to inductive inferences).

Finally, the notion of broadly inferential articulation is extended in subsequent chapters to include the crucial inferential substructures of *substitution* and *anaphora*. Substitutional commitments are defined as a species of inferential commitments (in Chapter 6) by distinguishing a class of substitution inferences. In this way the inferentialist paradigm can be extended so as to apply to the conceptual contents of subsentential expressions such as singular terms and predicates. Then anaphoric commitments are defined in terms of the *inheritance* of substitution-inferential commitments (in Chapter 7). In this way the inferentialist paradigm can be extended so as to apply to unrepeatable or token-reflexive expressions, such as demonstratives, indexicals, and pronouns.

It is important to keep in mind in reading what follows that the inferentialist project pursued here is a defense of the strong, not only the weak, inferentialist thesis. But it is not a form of hyperinferentialism. And while it eschews representational semantic primitives in favor of others more easily grounded in pragmatics, this is not because of a denial of the importance of the representational dimension of discursive practice. On the contrary, that choice serves rather an aspiration to make intelligible in a new way just what that representational dimension consists in.

## VI. CONCLUSION

### 1. *Grounding an Inferential Semantics on a Normative Pragmatics*

Inferentialism about conceptual content is not an explanatory strategy that can be pursued in complete abstraction from pragmatism about the norms implicit in the practical application of concepts. The considerations assembled here to motivate and recommend an inferentialist order of

semantic explanation appeal to a notion of *materially correct inferences*. In this chapter material proprieties of inference have been treated as primitives, playing the role of unexplained explainers. A critical criterion of adequacy by which such an approach should be assessed is clearly the extent to which a philosophically satisfying story can be told about these primitive proprieties of nonformal inference. The semantic theorist's entitlement to explanatory use of such primitives must be vindicated by situating the project of semantic theory in a broader context. Conceptual contents, paradigmatically propositional ones, are associated with linguistic expressions as part of an attempt to specify, systematically and explicitly, the correct use of those expressions. Such contents are associated with intentional states such as belief as part of a corresponding attempt to specify their behavioral significance—the difference those states make to what it is appropriate for the one to whom they are attributed to do.

The study of the practical significance of intentional states, attitudes, and performances (including speech acts) is *pragmatics*, as that term is used here. The projects of semantic theory and of pragmatic theory are intricately interrelated. If the semantic content and pragmatic context of a linguistic performance of a particular kind (paradigmatically assertion) are specified, a general theory of speech acts seeks to determine in a systematic way the pragmatic significance of that contentful performance in that context. But besides the direction of explanation involved in the local determination of pragmatic significance by semantic content, there is also a converse direction of explanation involved in the global conferral of semantic content by pragmatic significance. It must be explained how expressions can be used so as to confer on them the contents they have—what functional role the states they manifest must play in practice for them to be correctly interpreted as having certain intentional contents. Such an explanation amounts to an account of what it is for a state, attitude, performance, or expression to be *propositionally* contentful. Once a general notion of content has been made sense of in this way, particular attributions of contentfulness can then be offered as part of explanations or explicit specifications of the pragmatic significance of a state, attitude, performance, or expression.

The discussion of the next chapter should begin to make clearer just how a story about the conferral of content by practice is envisaged as relating to the use of attributions of content in the determination of pragmatic significances. One aspect of the situation of the semantic concept of content in a wider pragmatic context, however, is of particular relevance to the issue of entitlement to appeal to material proprieties of inference as semantic primitives. For the inferential proprieties that from the point of view of semantic theory are treated as primitive can be explained in the pragmatic theory as implicit in discursive practice (which includes intentional agency). An inferential move's normative status as correct or incorrect can be construed as instituted in the first instance by practical attitudes of taking or treating it

as correct or incorrect. The inferential norms that govern the use of expressions (or the significance of states, attitudes, and performances) are then understood as constituted by practical attitudes toward what the content is attributed to; they in turn confer that content on what it is attributed to.

Expressions come to mean what they mean by being used as they are in practice, and intentional states and attitudes have the contents they do in virtue of the role they play in the behavioral economy of those to whom they are attributed. Content is understood in terms of proprieties of inference, and those are understood in terms of the norm-instituting attitudes of taking or treating moves as appropriate or inappropriate in practice. A theoretical route is accordingly made available from what people *do* to what they *mean*, from their *practice* to the *contents* of their states and expressions. In this way a suitable pragmatic theory can ground an inferentialist semantic theory; its explanations of what it is in practice to treat inferences as correct are what ultimately license appeal to material proprieties of inference, which can then function as semantic primitives.

Sketching the possibility of such an explanatory path from attributions of practical attitudes to attributions of semantic content should help alleviate one sort of worry that might be elicited by the inferentialist invocation of materially correct inferences in explaining conceptual contentfulness. For otherwise the employment of a notion of *material* proprieties of inference in explaining *content* might seem blatantly circular. After all, are not materially good inferences just those that are good in virtue of the contents of the nonlogical concepts applied in their premises and conclusions, by contrast to the logically valid inferences, which are good in virtue of the logical form of those premises and conclusions? Presystematically, this is indeed how they should be thought of. But officially, the strategy is to start with proprieties of inference and to elucidate the notion of conceptual content in terms of those proprieties.

Talk of materially correct inferences is indeed intended to enforce a contrast with those that are formally correct (in the sense of logically valid). But the force of this contrast is just that the validity of inferences in virtue of their logical form is to be understood as a sophisticated, late-coming sort of propriety of inference, founded and conceptually parasitic on a more primitive sort of propriety of inference. This is the repudiation of the formalist approach to inference, for which the correctness of inference is intelligible only as formal logical validity, correctness in virtue of logical form. Calling the more primitive sort of propriety of inference *materially* correct simply registers the rejection of this order of explanation. It does not involve commitment to a prior notion of nonlogical content. If what it means to call an inference correct in the relevant sense can be explained without appeal to the use of logical concepts—for instance in terms of conduct interpretable as a practical taking or treating of an inference as correct—then there need be no circularity in appeal to such inferential proprieties in elaborating a notion of conceptual content.

2. *Knowing-That in Terms of Knowing-How, Formal Proprieties of Inference in Terms of Material Ones, Representational Content in Terms of Inferential Content*

A story that begins with inferring as a kind of practical *doing* and that leads to an account of the specifically *propositional* contentfulness of speech acts and intentional states holds out the promise of yielding an account of propositionally *explicit* saying, judging, or knowing-*that*, in terms of practically *implicit* capacities, abilities, or knowing-*how*. This would discharge one of the primary explanatory obligations of the pragmatist foe of the intellectualist understanding of norms. For if practical knowing-*how* is taken as prior in the order of explanation to theoretical knowing-*that*, one must not only offer an independent account of the practically implicit grasp or mastery of norms. One must also explain how the propositionally explicit grasp of norms expressed in the form of rules, principles, or claims can be understood as arising out of those practical capacities.

In the same way, the inferentialist approach to content treats material proprieties of inferences as prior in the order of explanation to formal logical proprieties of inference. It is accordingly obliged not only to offer an independent account of those material proprieties but also to offer an account of how logical goodness of inference can be explained in terms of that primitive sort of goodness of inference. One who denies that logic is to be understood as underlying (and so presupposed by) rationality in the sense involved in the inferential articulation of conceptual contents (and so in any exercise of the capacity to give and ask for reasons) is obliged to offer another account of logic. This obligation is discharged by the combination of two moves. The first is offering a criterion of demarcation for logical vocabulary that is couched in terms of the semantically expressive role played by such vocabulary in making implicitly content-conferring inferential commitments explicit in the form of judgments. This move depends on having a view about what it is for something to be explicit in the form of a judgeable, that is propositional, content. Such a view is precisely what the account of propositional contents in terms of material proprieties implicit in inferential practice, mentioned above, is intended to supply. The second element required to discharge the obligation to show how the notion of logically good inferences grows out of that of materially good inferences is the substitutional account of formal logical validity of inference—according to which an inference is valid or good in virtue of its logical form if it is primitively good and cannot be turned into one that is not primitively good by any (grammatical) substitution of nonlogical for nonlogical vocabulary.

An explanatory demand exhibiting the same structure as that just rehearsed for the anti-intellectualist about norms and the antiformalist about logic is incumbent on the inferentialist account of conceptual content in virtue of its commitment to invert the representationalist order of semantic explanation.<sup>90</sup> A viable working-out of the inferentialist order of explanation

must, to begin with, offer an account of correctness of inference that is not parasitic on correctness of representation. This demand is addressed in Chapters 3 and 4, which specify sufficient conditions for an attribution of implicitly normative social practices to a community to count as interpreting them as engaging in practices of giving and asking for reasons—as practically assessing inferences as correct or incorrect, and so as instituting material inferential proprieties that confer propositional conceptual content on their states and performances. It is not enough, however, for the inferentialist explanatory strategy to produce an account of the pragmatic basis of its own semantic primitives that does not rely on the prior intelligibility of representational concepts. It must also show how representational relations and the sorts of representational content they underwrite can be made intelligible in terms of those inferential primitives. That is, another critical criterion of adequacy of inferentialism is the extent to which, if this approach is granted its preferred starting point, it can develop it into an account of the sort of objective representational content other approaches begin with.

### *3. Objective Representational Content*

Meeting this demand involves offering accounts of three important dimensions along which the notion of objective representational content is articulated. First is the referential dimension. The representationalist tradition has, beginning with Frege, developed rich accounts of *inference* in terms of *reference*. How is it possible conversely to make sense of reference in terms of inference? In the absence of such an account, the inferentialist's attempt to turn the explanatory tables on the representationalist tradition must be deemed desperate and unsuccessful.

The second dimension is categorial. An account must be offered not just of reference and representation but of reference to and representation of particular *objects* and general *properties*. That is, the peculiar kind of representational content expressed by subsentential expressions, paradigmatically singular terms and predicates, must be explained. For reasons already indicated, inferential approaches to conceptual content apply directly only to what is expressed by declarative sentences, which can play the role of premises and conclusions of inferences. Somehow the inferential approach to conceptual content must be extended to apply to subsentential parts of speech as well. The discussion of Dummett's model of circumstances and consequences of application provides some suggestive hints. But these must be developed far beyond the remarks already offered in order to put the inferentialist in a position to claim to have shown that the nominalist order of explanation standard prior to Kant, beginning with a doctrine of terms or concepts and moving from there to a doctrine of judgments, can successfully be stood on its head.

Finally, there is the *objective* dimension of representational content. It

must be shown how on inferentialist grounds it is possible to fund *objective* proprieties of inferring and judging—to make intelligible the way in which what it is correct to conclude or to say depends on how the objects referred to, talked about, or represented actually are. Even if, to begin with, attention is restricted to inferential proprieties, it is clear that not just any notion of correctness of inference will do as a rendering of the sort of content we take our claims and beliefs to have. A semantically adequate notion of correct inference must generate an acceptable notion of conceptual content. But such a notion must fund the idea of *objective* truth conditions and so of *objectively* correct inferences. Such proprieties of judgment and inference outrun actual attitudes of taking or treating judgments and inferences as correct. They are determined by how things actually are, independently of how they are taken to be. Our cognitive attitudes must ultimately answer to these attitude-transcendent facts.

This means that although the inferentialist order of explanation may start with inferences that are correct in the sense that they are accepted in the practice of a community, it cannot end there. It must somehow move beyond this sense of correctness if it is to reach a notion of propositional conceptual content recognizable as that expressed by our ordinary empirical claims and possessed by our ordinary empirical beliefs. Pursuing the inferentialist order of explanation as outlined above accordingly requires explaining how—if actual practical attitudes of taking or treating as correct institute the normative statuses of materially correct inferences, and these material proprieties of inference in turn confer conceptual content—that content nonetheless involves objective proprieties to which the practical attitudes underlying the meanings themselves answer.<sup>91</sup> How is it possible for our use of an expression to confer on it a content that settles that we might all be wrong about how it is correctly used, at least in some cases? How can normative attitudes of taking or treating applications of concepts as correct or incorrect institute normative statuses that transcend those attitudes in the sense that the instituting attitudes can be assessed according to those instituted norms and found wanting? This issue of objectivity is perhaps the most serious conceptual challenge facing any attempt to ground the proprieties governing concept use in social practice—and the pragmatist version of inferentialism being pursued here is a view of this stripe.

In the terms set up in Section I of this chapter, the referential, categorial, and objective can be thought of as three interlocking dimensions of the project of explaining *object-representing* contentfulness in terms of *propositional* contentfulness, according to a semantic rendering of propositional contentfulness in terms of *material proprieties of inference* and a *pragmatic* rendering of those basic inferential proprieties.<sup>92</sup> The relation between inference and reference is discussed in an introductory way in Chapter 5, which examines the use of the semantic vocabulary (paradigmatically ‘refers’ and ‘true’) by whose means the implicit referential dimension of conceptual

contents is made explicit. The anaphoric relations that are invoked in the explanation offered there are then explained in more primitive pragmatic terms in Chapter 7, which relates them to the sort of substitution-inferential commitments discussed in Chapter 6. The categorial issue is addressed by further development of Frege's substitutional methodology, in Chapter 6, which shows how the contents expressed by the use of singular terms and predicates can be understood in terms of substitution inferences.

The objectivity issue, which concerns the relation between what is properly *said* and what is talked *about*, intimately involves both the referential or representational dimension and the categorial. As with all substantive semantic and pragmatic concepts officially employed in this work in describing the contents and significances of discursive commitments, the concept of states, attitudes, and performances that are objectively contentful in purporting to represent how things are independently of anyone's states, attitudes, and performances is discussed at two levels. (This two-leveled account is part of the effort to secure for the use of such vocabulary in this project an analog of the kind of *expressive equilibrium* already adverted to as achieved by Frege's treatment of sentential logical vocabulary in the *Begriffsschrift*.) The first is a story about what it is for such purport and its uptake to be *implicit* in the practices of those whose states, attitudes, and performances are properly interpreted as having such content.

The second is a story about what it is for such purport and its uptake to be made *explicit* in the specification of contents of ascribed states, attitudes, and performances. Although technical philosophical vocabulary such as 'refers' or 'denotes' (discussed in Chapter 5) can play this explicating role, the fundamental locutions used in ordinary talk to express representational commitments are those used to form *de re* specifications of the contents of ascribed intentional states, attitudes, and performances—paradigmatically 'of' and 'about'. The use of *de re* ascriptions makes it possible to specify explicitly what is said in terms of what is talked about. What such ascriptions express and how those objective content-specifications are made explicit by their use is the topic of Chapter 8. The account of what we are doing when we interpret ourselves and each other as making claims with objective representational content that is offered in that chapter requires the expressive resources of all of the sorts of locutions whose use is introduced in prior chapters. It is accordingly only in the last substantive chapter of this work that this critical explanatory obligation of an inferential approach to semantics is finally discharged.

The foundation of that account is laid in the next chapter. It consists in the *social* structure of the inferential norms that confer propositional content. (Government by such norms is what such contentfullness is.) The development into a more full-blooded notion of conceptual content of the abstract notion of inferential role introduced in this chapter proceeds by taking account of the social dimension of inferential practice—which is

implicit in the idea that abstract talk about inferential relations must be rooted in consideration of what Sellars calls “the game of giving and asking for reasons.” The pragmatic significance of making a claim or acquiring a commitment whose content could be expressed by the use of a declarative sentence cannot be determined by associating with that sentence a set of sentences that entail it and a set of sentences that it entails—not even if these are enriched by throwing in nonlinguistic circumstances and consequences of application as well. This is because of the interaction of two features of inferentially articulated commitments.

First, as Frege acknowledges in his original definition of *begriffliche Inhalt*, specification of the inferential role of a sentence requires looking at *multipremise* inferences.<sup>93</sup> Many of the important “consequences of application” of a sentence are not consequences it has all on its own; they consist rather in the differential contribution its inclusion makes to the consequences of a set of collateral premises or auxiliary hypotheses. Similarly, its purely inferential antecedents must be thought of not as individual sentences but as sets of them.

Second, the collateral concomitant commitments available as auxiliary hypotheses in multipremise inferences vary from individual to individual (and from occasion to occasion or context to context). If they did not, not only the notion of communication but even that of empirical information would find no application. The significance of acquiring a commitment or making a claim whose content could be expressed by the use of a particular sentence, when it would be appropriate to do so and what the appropriate consequences of doing so would be, depends on what other commitments are available as further premises in assessing grounds and consequences. What is an appropriate ground or consequence of that commitment from the point of view of one set of background beliefs may not be from the point of view of another. In view of the difference in their other attitudes, a single commitment typically has a different significance for the one undertaking it, a speaker or believer, from that which it would have for those attributing it, an audience or intentional interpreter. Of course this does not make communication or interpretation impossible—on the contrary. As was just mentioned, it is only the prevalence of situations in which background commitments do differ that give communication and interpretation their point.

The fact that the implicitly normative inferential significance of a commitment may be different from the point of view of one undertaking the commitment and one attributing it means that the inferential articulation of conceptual contents has a fundamental *social* dimension. It introduces a relativity to social perspective into the specification of such contents. The practical attitudes of taking or treating as committed, which ultimately institute the normative status of commitment, come in the two socially distinct flavors of undertaking or acknowledging a commitment (oneself) and

attributing a commitment (to another). Inferentially articulated contents are conferred on states, attitudes, and performances by the norms instituted by *social* practices: those that essentially involve the interaction of attitudes corresponding to both social perspectives. Investigation of the use of locutions that make explicit various aspects of the social perspectival character of conceptual contents will reveal what they express as the source of objective representational content. So, it will be claimed, what must be added to the *normative* approach to pragmatics and the *inferential* approach to semantics in order to make intelligible the *representational* dimension of conceptual contents is a *social* account of the interaction between them.

# Linguistic Practice and Discursive Commitment

---

Language most shows a man: Speak, that I may see thee.

BEN JONSON, *Timber or Discoveries*

Language is called the Garment of Thought: however, it should rather be,  
Language is the Flesh-Garment, the Body, of thought.

THOMAS CARLYLE, *Sartor Resartus*

Clearly human beings could dispense with all discourse, though only at  
the expense of having nothing to say.

WILFRID SELLARS, "A Semantical Solution to  
the Mind-Body Problem"

---

## I. INTENTIONAL STATES AND LINGUISTIC PRACTICES

### 1. *Discursive Practice as Deontic Scorekeeping*

This chapter introduces a particular model of language use: the deontic scorekeeping model of discursive practice. The implicitly normative social practices it describes are inferentially articulated in such a way as to confer specifically *propositional* contents on expressions and performances that play suitable roles in those practices. The basic idea is the one motivated by the discussion in Chapter 2, namely that propositional contentfulness must be understood in terms of practices of giving and asking for *reasons*. A central contention is that such practices must be understood as *social* practices—indeed, as *linguistic* practices. The fundamental sort of move in the game of giving and asking for reasons is making a *claim*—producing a performance that is propositionally contentful in that it can be the offering of a reason, and reasons can be demanded for it. Other theoretically important concepts are defined in terms of this one: *linguistic* practice is distinguished by its according some performances the significance of claimings, and (declarative) *sentences* are distinguished as expressions whose utterances, inscriptions, or other tokenings have the default significance of claimings. The basic explanatory challenge faced by the model is to say what structure a set of social practices must exhibit in order properly to be understood as includ-

ing practical attitudes of taking or treating performances as having the significance of claims or assertions.

According to the model, to treat a performance as an assertion is to treat it as the undertaking or acknowledging of a certain kind of *commitment*—what will be called a ‘doxastic’, or ‘assertional’, commitment. To be doxastically committed is to have a certain social status. Doxastic commitments are normative, more specifically *deontic*, statuses. Such statuses are creatures of the practical attitudes of the members of a linguistic community—they are instituted by practices governing the taking and treating of individuals *as committed*. Doxastic commitments are essentially a kind of deontic status for which the question of *entitlement* can arise. Their inferential articulation, in virtue of which they deserve to be understood as propositionally contentful, consists in consequential relations among the particular doxastic commitments and entitlements—the ways in which one claim can commit or entitle one to others (for which it accordingly can serve as a reason).

Competent linguistic practitioners keep track of their own and each other’s commitments and entitlements. They are (we are) *deontic scorekeepers*. Speech acts, paradigmatically assertions, alter the deontic score; they change what commitments and entitlements it is appropriate to attribute, not only to the one producing the speech act, but also to those to whom it is addressed. The job of pragmatic theory is to explain the *significance* of various sorts of speech acts in terms of practical proprieties governing the keeping of deontic score—what moves are appropriate given a certain score, and what difference those moves make to that score. The job of semantic theory is to develop a notion of the *contents* of discursive commitments (and the performances that express them) that combines with the account of the significance of different kinds of speech act to determine a scorekeeping kinematics.

The basic elements of this deontic scorekeeping model of discursive practice are presented in this chapter. The next chapter develops it further by attending in particular to the inferential articulation of *perception* and *action*. These are the entries to and exits from the realm of discursive commitments and entitlements—the source respectively of the empirical and practical dimensions of conceptual content, which are usually (and in one sense correctly) thought of as *noninferential*. Before plunging into a description of the details of the features of a system of social practices in virtue of which they should be understood as incorporating practical scorekeeping attitudes that institute deontic statuses and confer propositional contents, however, some methodological preliminaries are in order. The rest of this work presents not only an account of linguistic intentionality (thought of as one sophisticated species in a genus comprising other, more primitive sorts) but a linguistic account of intentionality generally. It is claimed that the propositional contentfulness even of the beliefs and other states intentional

interpreters attribute to nonlinguistic animals cannot properly be understood without reference to the specifically linguistic practice of the interpreters, from which it is derived. Original, independent, or nonderivative intentionality is an exclusively linguistic affair. The reasons for insisting on the conceptual primacy of linguistic intentionality cannot be presented until all the materials needed for the analysis of the representational dimension of propositional content (and of conceptual content generally) have been assembled, in Chapter 8. Nonetheless, the explanatory strategy being pursued will be easier to understand if the picture of the relations between language and belief that it incorporates has been sketched, even if the warrant for that picture cannot emerge until it is more fully developed (in Part 2 of this work).

## 2. *Philosophical Semantics and Formal Semantics*

One of the fundamental methodological commitments governing the account presented here is *pragmatism* about the relations between semantics and pragmatics. Pragmatism in this sense is the view that what attributions of semantic contentfulness are *for* is explaining the normative significance of intentional states such as beliefs and of speech acts such as assertions. Thus the criteria of adequacy to which semantic theory's concept of content must answer are to be set by the pragmatic theory, which deals with contentful intentional states and the sentences used to express them in speech acts. The idea that *philosophical* theories of meaning or content must be concerned with the larger pragmatic context within which attributions of contentfulness play an explanatory role may seem to be brought into question by the autonomy of *formal* semantics. But the independence of formal semantics from pragmatic concerns is only apparent.

The project of formal semantics entitles the theorist at the outset to stipulate an association of semantic interpretants with primitive interpreteds, typically linguistic expressions. Then this interpretation is extended to interpreteds that are derived from those primitives by syntactic operations—which for standard compositional syntactic structures include category-sensitive concatenation and various grammatical transformations of such concatenations. This is achieved by defining, for each syntactic operation on interpreteds, a corresponding operation on their associated interpretants that yields a new interpretant, which is thereby associated with the result of the syntactic operation. So the formal semantic theorist might begin by associating truth-values with sentence-letters, and then for each connective that produces compound sentences introduce a function taking sets of truth-values into truth-values that can then be assigned to the corresponding compound sentences. Or instead of truth-values, the semantic interpretants might be sets of possible worlds, and the operations corresponding to sentential connectives be set-theoretic operations on them (such as intersection for conjunction).<sup>1</sup>

So formal semantics is concerned generically with structure-preserving mappings. But not every mathematical representation theorem, which shows such a correspondence between structures of one kind and those of another, deserves to be called a semantics. What else ought to be required for a set of such mappings to count as presenting a specifically *semantic* interpretation of something? To ask this question is already to begin to move from the domain of purely formal semantics to that of philosophical semantics. When Tarski proved an algebraic representation theorem in which the interpretants assigned to quantificational expressions are topological closure operators, what qualifies that as a formal semantics for the first-order predicate calculus is not anything about the intrinsic properties of those interpretants but just that he is able in those terms to reproduce the relation of logical consequence appropriate to that idiom. From a purely formal or mathematical point of view, the task would be no different if the property to be reproduced were specified simply by randomly partitioning the elements of one grammatical category, placing an asterisk next to some of them and not others (and similarly for the relation in question). From the point of view of the philosophical motivation of calling what one is doing thereby 'semantics', however, it makes all the difference that the elements involved be interpretable as *sentences*, and that the property distinguished be interpretable as theorem-hood, a kind of *truth*, and that the relation distinguished be interpretable as derivation, a kind of *inference*. Indeed, to take the elements as subject to evaluations concerning propriety of *judging* and propriety of *reasoning*, truth and inference, is just what it is to interpret them as sentences.

What gives semantic theory its philosophical point is the contribution that its investigation of the nature of contentfulness can make to the understanding of proprieties of practice, paradigmatically of judging and inferring. That semantic theory is embedded in this way in a larger explanatory matrix is accordingly important for how it is appropriate to conceive the semantic interpretants associated with what is interpreted. It means that it is pointless to attribute semantic structure or content that does no pragmatic explanatory work. It is only insofar as it is appealed to in explaining the circumstances under which judgments and inferences are properly made and the proper consequences of doing so that something associated by the theorist with interpreted states or expressions qualifies as a *semantic* interpretant, or deserves to be called a theoretical concept of *content*. Dummett puts the point this way:

The term 'semantics', at least as commonly applied to formalized languages, usually denotes a systematic account of the truth-conditions of sentences of the language: the purpose of thus assigning a value, true or false, to every well-formed sentence of the language is taken as already understood, and receives no explanation within the semantic theory itself . . . The classification of the sentences of a formalized

language into true ones and false ones relates to the purposes for which we want to use the language. But in the case of natural language, it is already in use: the only point of constructing a semantics for the language can be as an instrument for the systematic description of that use, that is, as part of a whole theory of meaning for the language, which as a whole constitutes an account of its working. If the semantic part of the theory is taken as issuing an assignment of conditions under which each sentence of the language, as uttered on a particular occasion, has this or that truth-value, the rest of the theory must connect the truth conditions of the sentences with the use to which they are put, that is, with the actual practice of speakers of the language . . . a semantic theory which determines the truth-conditions of sentences of a language gets its point from a systematic connection between the notions of truth and falsity and the practice of using those sentences.<sup>2</sup>

The essential point is that philosophical semantic theory incorporates an obligation to make the semantic notions it appeals to intelligible in terms of their pragmatic significance. Formal semantics qualifies as *semantics* only insofar as it is implicitly presupposed that this obligation can be satisfied by conjoining the semantics with some suitable pragmatics.<sup>3</sup>

Philosophical semantics is distinguished from formal semantics by its explicit concern with the relation between the use of semantic concepts, on the one hand, and pragmatic accounts of the proprieties of practice governing the employment of what those concepts apply to, on the other. Philosophical semantics is committed to explaining the content of concepts such as *content*, *truth*, *inference*, *reference*, and *representation*, while formal semantics is content to use such concepts, assuming them (and so the pragmatic significance of applying them) already to be implicitly intelligible. The difference between doing either sort of semantics for artificial languages and for natural languages is that in the former case there are no antecedent proprieties governing the use of the expressions, to which the semantic theorist is responsible. Since the language is not already in use, the theorist is free to *stipulate* an association of contents with expressions, in order to determine how they are to be understood to be correctly used. In the case of natural languages, however, the theorist's use of semantic concepts is not *synthetic* (to settle the proper employment of expressions that antecedently are subject to no such proprieties) but *analytic* (to codify and express antecedently existing proprieties of employment).

### 3. Associating Content Explicitly by Stipulation and Implicitly Conferring It by Practice

Philosophical semantic theories of expressions and states that already play normatively articulated roles in linguistic practice or in the

practical reasoning of rational agents accordingly cannot afford the luxury (enjoyed by formal semantics of all sorts and even by philosophical semantics of artificial languages) of employing a *stipulative* method. Such theories are obliged to explain what the association of content with expression or state *consists in*: what one is saying or doing in attributing content to them. At this point it has seemed to many that the cases of contentful *sentences* and of contentful *beliefs* diverge. It makes sense to think of the contents of linguistic expressions as *conferred* on them by the way they are used. Noises and marks mean nothing all on their own. No one thinks they are *intrinsically* contentful. The sign-design 'dog' could as well be used to express the concept expressed by the sign-design 'horse', or to express none at all, like 'gleeb'. It is only by being caught up in linguistic practice that they come to express propositions, make claims, have or express conceptual or intentional contents. Apart from their role in human activity, apart from the norms thereby imposed on their employment—which make it the case that some uses are correct and others incorrect—these linguistic vehicles are semantically mute, inert, dead.

The philosophical semantics of natural languages must begin, then, with the observation that it is the practice of those who use the language that confers content on the utterances and inscriptions that are the overt, explicitly expressive performances whose propriety is governed by that practice. Is something similar true of intentional states? There are some important asymmetries between the two cases. There is a familiar line of thought, already adverted to, according to which quite a different story must be told about the association of content with the states and attitudes, paradigmatically beliefs, that are expressed by such linguistic performances. The critical question is how to understand the *use* of language in which the pragmatic significance of speech acts consists and which accordingly confers semantic content on those speech acts and so indirectly on the expressions they involve. One way to think about such use is *instrumentally*. This line of thought may be traced back to Locke, who thought of speech as an instrument for communicating thoughts or ideas. It is successful when the noise emitted by the speaker arouses in the audience an idea with the same content as that prompting the speech act.

Contemporary elaborations of this approach see "nonnatural" meaning as rooted in the capacity of individuals deliberately to imbue signals with significance by producing them with the intention that they be understood in a certain way by their auditors. According to Grice's picture,<sup>4</sup> linguistic practitioners *make* their expressions have a certain content by producing them with the intention that others *take* them to have that content. In particular, assertion is understood as the expression of belief in the sense that a sentence is produced with the intention that those who hear it will acquire a certain belief in virtue of their recognition that the speaker uttered the sentence intending those who hear it to acquire that belief, in virtue of their

recognition of that very intention. The notion of a linguistic expression's meaning something is in this way derived from the capacity of language users deliberately to mean something by their utterances. In somewhat different ways, Lewis, Bennett, and Searle develop this instrumental model by showing how a layer of *conventions* can be built on such communicative *intensions*, in such a way that members of a linguistic community are for the most part relieved of the necessity for elaborate deliberation about each other's beliefs and intentions in choosing and interpreting each other's remarks.<sup>5</sup> The foundation on which the conventional meaningfulness of linguistic acts and expressions rests remains their intentional employment as means to an explicitly envisaged communicational end.<sup>6</sup>

Rosenberg calls this explanatory strategy "agent semantics," because linguistic meaning is explained in terms of a prior capacity to engage in practical reasoning.<sup>7</sup> If the pragmatic use of language that confers semantic content on utterances and expressions is understood in these terms, it is clear that the contentfulness of intentional states such as beliefs, desires, and intentions must be understood antecedently, and hence according to some other model. Agent semantics treats the contentfulness of utterances as derivative from that of intentional states. The content of an assertion derives from the content of the belief it is the expression of, and from the content of the intention that it be understood as expressing that belief. It follows that it must be possible to make sense of the contents of beliefs and intentions prior to and independently of telling this sort of story about the use of linguistic expressions. Their content cannot be taken to be conferred on them by the way they are used or employed, according to this model of use or employment.<sup>8</sup>

That the content of intentional states cannot be understood as conferred on them by proprieties governing their significance—when it is appropriate to acquire them and what the appropriate consequences of acquiring them are—follows only if the only candidate for content-conferring use is deliberate, instrumental employment in order to secure the *explicitly* envisaged purpose of being understood as having a certain content. It would not follow that semantic content could not be conferred on intentional states by proprieties *implicit* in the way those states are treated in practice. According to such a conception, the conferral of content might be a side effect of the way they are treated, not requiring that anyone explicitly intend to confer it by their behavior. Broadly functionalist approaches to content are of this sort. They understand intentional states to be contentful in virtue of the role they play in the proper functioning of some system of which they are a part. Going into a certain state is something that is done appropriately under some circumstances, according to the functional interpretation of the system, and it has certain appropriate consequences. Together these proprieties of input and output, antecedents and consequences, determine the functional role of the state in the system. According to the functionalist explanatory strategy,

it is in virtue of playing this role, being subject to these transitional properties, that intentional states have the content they do.

#### 4. Intentionality: Linguistic Practice versus Rational Agency

The first question that needs to be addressed in working out such an approach is how the relevant functional system should be understood. Is it possible to understand propositional and other genuinely conceptual contents as conferred on states and performances by their role in a functional system that comprises only a single individual? Or is discursive practice essentially *social* practice, so that the functional system must be taken to comprise the activities of an entire community? The most popular and promising way of developing the first answer looks to the role belief plays in the practical instrumental reasoning of intelligent agents. The most popular and promising way of developing the second answer looks to the role assertion—the explicit expression of belief—plays in linguistic practice.

The considerations assembled in the first two chapters suggest the motivation that these two approaches have in common: states, attitudes, and performances are intentionally contentful in virtue of the role they play in *inferentially* articulated, implicitly *normative* practices. It is by looking at the practices in which the status of some states, attitudes, and performances as providing *reasons* for others is implicitly (and constitutively) acknowledged that the pragmatic significance of associating them with intentional contents is to be understood. There are two different sorts of context in which the specifically inferential significance of intentional states such as belief is to the fore: rational agency and linguistic practice. On the one hand, beliefs and other intentional states are expressed in *actions*, nonlinguistic performances that are intelligible in virtue of the beliefs and desires that are reasons for them. On the other hand, beliefs are expressed in *claims*. Overt assertions are the fundamental counters in the game of giving and asking for reasons—they can both be offered as reasons and themselves stand in need of such reasons.

Stalnaker points out that these two sorts of context in which intentional states are significant for practice generate two basically different ways of looking at those states: what he calls the *pragmatic* picture of intentionality and the *linguistic* picture of intentionality. The pragmatic picture is one according to which “rational creatures are essentially agents . . . According to this picture, our conceptions of belief and of attitudes pro and con are conceptions of states which explain why a rational agent does what he does . . . Linguistic action, according to this picture, has no special status. Speech is just one kind of action which is to be explained and evaluated according to the same pattern. Linguistic action may be a particularly rich source of evidence about the speaker’s attitudes, but it has no special conceptual connection with them.”<sup>9</sup> This picture amounts to a generalization of the ap-

proach of agent semantics. It shares the emphasis on rational agency as the conceptually and explanatorily fundamental context in which to understand the significance of the contentfulness of intentional states. It is more general in that it does not necessarily involve a commitment to understanding the contentfulness of speech acts as deriving from their deliberate instrumental employment to secure antecedently envisageable goals. Thus this picture leaves room for a picture in which the way speech acts inherit the contents of the intentional states they express might be less intellectualized, the conferral of content being implicit in the practice of expression, rather than explicit as its instrumentally conceived motive. The contrasting linguistic picture is one according to which "rational creatures are essentially speakers . . . Representational mental states represent the world because of their resemblance to or relation with, the most basic kind of representations: linguistic expressions."<sup>10</sup>

This way of dividing up the fundamental orientations of various approaches to intentionality, accordingly as rational agency or linguistic capacity is taken as primary, evidently cuts at important joints. It is a measure of the robustness of this botanization that it is serviceable even across large differences in collateral theoretical commitments. Here is how Stalnaker puts the setting in which he sees the pictures as competing: "The linguistic and pragmatic pictures each suggest strategies for giving a naturalistic explanation of representation—both mental and linguistic representation—but the two strategies differ in what kind of representation they take to be more fundamental. The pragmatic picture suggests that we explain the intentionality of language in terms of the intentionality of mental states, while the linguistic strategy suggests that we explain the intentionality of mental states in terms of, or by analogy with, the intentionality of linguistic expressions."<sup>11</sup> On the side of semantic content, Stalnaker follows the tradition in seeing the issue as fundamentally one of representation, with inferential relations presumably to be explained further along in ultimately representational terms. And on the side of the pragmatic significance of intentional contentfulness, Stalnaker begins with a commitment to a naturalistic approach, with the normative character of the practice in which intentional states are significant (whether that practice is conceived in the first instance as rational agency or as essentially linguistic) presumably to be explained further along in ultimately naturalistic terms.

The order of explanation that frames his discussion is the reverse of that pursued in this work. For the semantic explanatory strategy being developed here looks first to inference, on the semantic side, and aspires to making the representational dimension of intentional content intelligible ultimately in inferential terms. And on the pragmatic side, the strategy is to begin with an account of norms implicit in practice and work out toward an understanding of their relation to their naturalistic setting, which the normative practices in their most sophisticated form make it possible to describe objectively. It

is noteworthy that in spite of these major differences in approach, the large division of options for explaining intentionality into those focusing on rational agency and those focusing on language seems compelling from both points of view.

### *5. Analogical and Relational Versions of Linguistic Approaches to Intentionality*

Broadly linguistic approaches comprise many importantly different variants, however, and these correspond to importantly different motivations. Stalnaker implicitly acknowledges one significant subdivision within the linguistic approach in his general characterization of the linguistic picture as seeking to explain the contentfulness of intentional states by appealing to their "resemblance to or relation with" contentful linguistic performances. The disjunction links two very different ways in which it might be thought that taking account of specifically linguistic practice is essential to making the contentfulness of intentional states explicit (in the sense of theoretically intelligible). The resemblance limb, according to which the contentfulness of states is modeled on that of expressions, involves a commitment to the claim that the theorist's or interpreter's capacity to attribute (and understand attributions of) contentful intentional states is parasitic on the theorist's or interpreter's capacity to attribute (and understand attributions of) contentful speech acts. It need not entail, for instance, that the intentional states attributed to nonlinguistic creatures are somehow second class. It requires only that what one is doing in attributing contentful states to nonlinguistic creatures cannot be understood apart from the capacity to attribute them to linguistic ones. By contrast, the relational limb—according to which the contentfulness of intentional states consists in or essentially involves the contentfulness of the speech acts that express them—involves a commitment to the claim that the theorist's or interpreter's capacity to attribute (and understand attributions of) contentful intentional states is in the first instance parasitic on the theorist's or interpreter's capacity to attribute (and understand attributions of) contentful speech acts to the *same* individuals who are taken to have the intentional states. It does entail that the intentional states attributed to nonlinguistic creatures are in important regards second-class statuses.<sup>12</sup>

One reason it is important to distinguish the claim that the intentionality of speech is conceptually prior to the intentionality of belief (as analogical theories have it) from the claim that as believers in the full sense, we are essentially rather than only accidentally speakers (as relational theories have it) is that only theories committed to the former thesis are obliged to offer accounts of linguistic practice that do not make reference to intentional states. It is open to one who subscribes to the second view to hold, as Davidson does, that attributions of contentful intentional states and content-

ful speech acts go hand in hand, that neither sort of attribution is intelligible apart from its relation to the other. So a relational account can understand the possibility of speech as essential to intentionality (in the paradigm cases from which we derive our grip on what intentional interpretation is) without thereby becoming obliged simply to invert the order of explanation characteristic of agent semantics—though such an account evidently cannot appeal to the sort of independently intelligible role of contentful states in rational agency presupposed by agent semantics.

In contrast, analogical linguistic theories of intentionality are committed to that converse order of explanation. Agent semantics employs an antecedent and independent notion of the contentfulness of intentional states to explain the derivative contentfulness of speech acts and linguistic expressions. A theory insisting that the contentfulness of intentional states is intelligible only by analogy to the contentfulness of speech acts and linguistic expressions would be obliged correspondingly to appeal to an antecedent and independent notion of the contentfulness of speech in order to explain the derivative contentfulness of intentional states. A relational linguistic theory of intentionality maintains rather that the understanding of intentionally contentful states that permits us to stretch the application of that notion and apply it in a second-class way to nonlinguistic animals (simple intentional systems) derives from and essentially depends on an understanding of the relation between the intentional states and the linguistic performances of language-using animals (communicating or interpreting intentional systems).

According to this sort of approach, understanding intentionality requires looking at practices that essentially involve *both* intentional states and linguistic performances. Neither sort of intentionality need be understood as conceptually prior to the other, and linguistic practice and rational agency can be presented as two aspects of one complex of jointly content-conferring practices. Davidson puts the characteristic contention of relational linguistic views of intentionality this way: “Neither language nor thinking can be fully explained in terms of the other, and neither has conceptual priority. The two are, indeed, linked, in the sense that each requires the other in order to be understood, but the linkage is not so complete that either suffices, even when reasonably reinforced, to explicate the other.”<sup>13</sup> The account of intentionality introduced in the rest of this chapter is a linguistic theory in this relational sense. The key to motivating a theory of this sort is to show what it is about the contents of intentional states that can be explained only by appealing to the relation between such states and specifically *linguistic* performances.

Davidson suggests that an argument for a relational theory can be provided in two pieces. He claims first that “someone cannot have a belief unless he understands the possibility of being mistaken, and this requires grasping the contrast between truth and error—true belief and false belief.”<sup>14</sup> What a

creature has does not function as a belief for that creature unless it has a certain kind of significance for that creature. It must be able to adopt a certain kind of practical attitude toward that state, to treat it in its practice or behavior as contentful in a special sense. In particular, Davidson is claiming, what it has is not recognizable as belief unless the creature whose state it is somehow in its practice acknowledges the applicability of a distinction between beliefs that are correct and those that are incorrect, in the sense of being true and false. We are not permitted to attribute the belief that *p* (a propositionally contentful intentional state) unless somehow the putative believer acknowledges in practice the *objective representational* dimension of its content—that its being held is one thing, but its being *correct* is another, something to be settled by how it is with what it is *about*. The second piece of Davidson's argument is the claim that a grasp of the contrast between correct and incorrect belief, true and false belief, "can emerge only in the context of interpretation, which alone forces us to the idea of an objective, public truth." The key claim is that "the concepts of objective truth and of error necessarily emerge in the context of interpretation."

The rest of this work focuses on the development of an account of linguistic social practices within which states, attitudes, and performances have, and are acknowledged by the practitioners to have, pragmatic significances sufficient to confer on them objective representational propositional contents. The view propounded is like Davidson's in seeing intentional states and speech acts as fundamentally of coeval conceptual status, neither being explicable except in an account that includes the other. It deserves nonetheless to be called a *linguistic* view of intentionality (of the relational rather than the analogical variety) because linguistic practice is nonetheless accorded a certain kind of explanatory priority over rational agency. The intentionality of nonlinguistic creatures is presented as dependent on, and in a specific sense derivative from, that of their linguistically qualified interpreters, who as a community exhibit a nonderivative, original intentionality. The sort of derivation in question is explicated in terms of the context that must be appealed to in making intelligible the sort of contents (*conceptual*—paradigmatically *propositional*—contents) that are associated with the intentional states attributed by interpreters. The contents of the intentional states attributed to nonlinguistic creatures can be understood only in a way that involves the activities of the language users who attribute them, and not entirely in terms of the activities of those who exhibit them. By contrast, the contents of the intentional states attributed to a community of language users can be understood as conferred on their states, attitudes, and performances entirely by the practices of those community members.

The argument that provides the ultimate justification for treating specifically *linguistic* practice as central in this way to intentionality has just the two-part form outlined by Davidson and rehearsed above. For what he

has really given us is not so much an argument as the form of one. Turning it into an actual argument requires filling in various notions of content, of objective representational correctness of content, of practical acknowledgement of the significance of assessments of correctness of content, and so on. That is the task of the rest of this work; the final justification for giving pride of place to language will thus not be complete until the end (in fact, in Chapter 8). At that point it will be possible to return again to the beginning, and know it for the first time—the warrant for this fundamental theoretical commitment will then be explicit.

### 6. *Believing and Claiming*

Claiming and believing are linked by the principle that assertions are one way of expressing beliefs. A fundamental question of explanatory strategy is then whether this principle can be exploited so as to account for one of these notions in terms of the other. Since there can be beliefs that are not avowed, the temptation is to start with belief and explain assertion as a speech act by which belief is expressed. But when the representational dimension of propositional and other conceptual contents is examined (in Chapter 8), it turns out to be intelligible only in the context of *linguistic* social practices of communicating by giving and asking for reasons in the form of claims. So if assertion were to be explained in terms of a prior notion of belief, the propositional contents of beliefs would have to be taken for granted. Their association with beliefs would have to be stipulated, rather than made intelligible as established by the functional role of beliefs in the behavioral economies of believers. The only sort of inferential practice that is socially articulated in the way that turns out to be required for the conferral of *propositional* content, in the form of objective truth conditions, is *assertional*, and therefore *linguistic* practice.

The idea pursued here is that the state or status of *believing* is essentially, and not merely accidentally, related to the linguistic performance of *claiming*. Beliefs are essentially the sort of thing that can be expressed by making an assertion. Dummett offers a crisp formulation of a view along these lines: "We have opposed throughout the view of assertion as the expression of an interior act of judgment; judgment, rather, is the interiorization of the external act of assertion."<sup>15</sup> Although the satisfyingly symmetric phrasing of this remark can obscure the point, it is important to realize that Dummett is not committed by it to the possibility of making sense of the speech act of asserting without mentioning anything but speech acts. For instance, this stance does not preclude an account of asserting that incorporates an account of the particular sort of *commitment* (a deontic status) one undertakes in making an assertion. What is precluded is only explaining assertion as the expression of a kind of intentional state or deontic status that is supposed to

be intelligible apart from the possibility of expressing it by asserting something.

The claim is, then, that speech acts having the pragmatic significance of assertions play an essential role in (social) functional systems within which states or statuses can be understood as propositionally contentful in the way beliefs are. A good way to see how explanatory progress might be made in thinking about beliefs by insisting on their linguistic expression in claims is provided by combining Dummett's thought with a suggestion of Hartry Field's concerning how an appeal to language might function as part of a divide-and-conquer strategy for explaining intentional states. His idea is that having a belief with a certain propositional content should be understood as standing in a certain relation to a sentence that expresses that content.<sup>16</sup> In particular, according to what he calls the "two-stage" explanatory strategy, A believes that *p* if and only if there is a sentence  $\sigma$  such that:

1. A believes\*  $\sigma$ , and
2.  $\sigma$  means that *p*.

In the first stage of an account with this structure, the relevant relation between a believer and a sentence—what Field calls “belief\*”—must be explained. In the second stage, what it is for a sentence to express a propositional content must be explained. Field's own way of pursuing this strategy takes the sentences involved to belong to a “language of thought,” which is conceived by analogy to public languages. This is an additional theoretical commitment on his part; nothing about the two-stage decomposition dictates that the sentences that play the role of middle terms should not be sentences in a public language such as English. Filling in the appeal to sentences by invoking a language used in interpersonal communication makes the two-stage strategy available for duty in what were called above relational linguistic theories of intentionality.

In this form, Dummett's claim about the fundamental importance of the speech act of assertion can be pressed into service in addressing the first subproblem of the two-stage strategy. One way in which beliefs are manifested or expressed is by the utterance of sentences. Sentence-utterings can have many sorts of force or pragmatic significance, but when such performances have the significance of *assertions*, they express or purport to express beliefs. As Dummett suggests, this fact can be exploited by two different orders of explanation. If the theorist can get an independent grip on the notion of belief, typically from the consideration of its functional role in the sort of practical reasoning implicitly attributed by intentional interpretations that use the model of rational agency to make nonlinguistic performances instrumentally intelligible, then asserting might be explained in terms of it, as a further way in which beliefs can be manifested. Or, if the theorist can get an independent grip on the practices within which performances are accorded the significance characteristic of assertions, belief might be ex-

plained as what is thereby expressed. If this path is followed, then only parties to linguistic practices, those that institute assertional significances, will qualify as believers. The decision as to which direction of explanation to adopt is in part an issue concerning just how important language is to us.

Like all fundamental demarcational matters, however, it is only in part a factual issue. There clearly is a sense in which nonlinguistic animals can be said to have beliefs. But the sense of belief that Sellars, Dummett, and Davidson are interested in (and that is the subject of this work) is one in which beliefs can be attributed only to language users. The best reason for adopting the contrary order of explanation, for treating the sort of nonlinguistic belief that is manifested in behavior that can be construed instrumentally as fundamental—seeking thereby to explain the sort of belief that is essentially and not merely accidentally expressible in speech acts—is that it is clear that there were nonlinguistic animals before there were linguistic ones, and the latter did not arise by magic. If the instrumental sense of belief could be made sense of antecedently, and the linguistic sense explained in terms of it, the prospects for explaining how linguistic practice could come into the world in the first place would be bright. This is a laudable aspiration, and it may seem perverse to spurn it. Yet it is a consequence of the account of propositional contents to be offered here that they can be made sense of only in the context of *linguistic* social practices, which have as their core the interpersonal communication of information by assertions. Likewise, rational agency, on which instrumental behavior is modeled, depends essentially on specifically linguistic practices, including asserting. It follows that simple, nonlinguistic, instrumental intentionality can *not* be made fully intelligible apart from consideration of the linguistic practices that make available to the interpreter (but *not* to the interpreted animal) a grasp of the propositional contents attributed in such intentional interpretations.

A theory with such a consequence obviously involves a collateral commitment to show that the conceptual priority of the linguistic sense of belief need not make mysterious the advent of linguistic practices from the capacities of hitherto nonlinguistic creatures. The story to be told here assumes only that suitable social creatures can learn to distinguish in their practice between performances that are treated as correct by their fellows (itself a responsive discrimination) and those that are not. In accord with the pride of place being granted to the linguistic sense of belief, no appeal will be made to instrumental rationality on the part of fledgling linguistic practitioners. The primary explanatory target is what it is to grasp a propositional content and to have and to attribute to others states and performances with such contents—in the sense of explaining what doing the trick consists in, what would count as doing it, rather than how it is done by creatures wired up as we are. Thus no attempt will be made to show how the linguistic enterprise might have gotten off the ground in the first place. But it should be clear at each stage in the account that the abilities attributed to linguistic practition-

ers are not magical, mysterious, or extraordinary. They are compounded out of reliable dispositions to respond differentially to linguistic and nonlinguistic stimuli. Nothing more is required to get into the game of giving and asking for reasons—though to say this is not to say that an interpretation of a community as engaged in such practices can be paraphrased in a vocabulary that is limited to descriptions of such dispositions. Norms are not just regularities, though to be properly understood as subject to them, and even as instituting them by one's conduct (along with that of one's fellows), no more need be required than a capacity to conform to regularities.

If the strategy Field proposes is pursued by looking at the use of the sentences of a public language to perform communicative speech acts, the two subproblems into which he divides the problem of how to understand the attribution of intentionally contentful states correspond to a fundamental pragmatic question and a fundamental semantic question. The first concerns what it is for the utterance of a linguistic expression to have the pragmatic significance of an assertion. This can be rephrased as the question of what it is to use a sentence to make an assertion, provided it is remembered both (1) that it cannot be assumed that *sentences* can be distinguished from other linguistic expressions in advance of saying what it is to use an expression to make an assertion and (2) that *use* should not be assumed to involve a deliberate instrumental exercise of rational agency. Chapter 1 recommended a broadly *normative* approach to the pragmatic question; Chapter 2 recommended a broadly *inferential* approach to the semantic question. It is the task of the rest of this chapter to weave these approaches together into an account of discursive practice—comprising implicitly normative, inferentially articulated statuses, attitudes, and performances. It is the role they play in discursive practice that confers on them objectively representational content, in the most basic case objectively representational *propositional* content. The capacity of practice to confer such contents depends essentially, it will be argued, on its being not only *social* practice but *linguistic* social practice, in that at its core is *communication*, specifically by practitioners' interpretation of each others' *assertions*.

Addressing the pragmatic limb of the two-stage explanatory strategy by appealing to the speech act of assertion yields a further subdivision of issues. As just indicated, it requires an account of what it is for a performance to have the force or pragmatic significance of an assertion, for it to function as an assertion in the game of giving and asking for reasons. Such an account, however, would not by itself yield a suitable reading of Field's *belief\** relation between a sentence and a potential subject of intentional interpretation or attribution of intentional states. For however tightly the two might be linked, there is a substantial difference between believing that *p* and claiming that *p*. (Commitment to a suitable resolution of the semantic subproblem entails a commitment to the eventual appropriateness of this sort of paraphrase in terms of propositional contents rather than sentences.) No sort of intentional

state (or normative status) that might be reconstructed in terms of assertion will provide a suitable analog for belief unless it preserves this contrast by leaving room for the possibility of being in the relevant intentional state without producing the corresponding assertional performance, and for producing the assertional performance without being in the corresponding intentional state. So besides an account of what the assertional significance of a speech act consists in, an account is required also of how attribution of such a speech act is to be understood to be related to attribution of the intentional state it expresses. Such an account might appeal to dispositions—for instance treating being in the state as being *disposed*, under appropriate conditions, to perform the speech act. Or it might appeal to norms—for instance treating the performance of the speech act as involving a *commitment* (which might or might not be fulfilled) involving the state. Or the account might involve both.

## II. DEONTIC STATUS AND DEONTIC ATTITUDES

### 1. *Doxastic Commitments*

The leading idea of the account to be presented here is that belief can be modeled on the kind of inferentially articulated *commitment* that is undertaken or acknowledged by making an assertion. These may be called *doxastic* or *assertional* commitments. This is the basic kind of *discursive commitment*. The strategy is to describe a simplified system of social practices in which something can be taken or treated as (having the significance of) an assertion—the acknowledging of commitment to an assertible content. “Assertible content” is what Frege’s “judgeable content”<sup>17</sup> becomes from the point of view of an explanatory commitment to understand judgments in the first instance as what is expressed by assertions. Specifically *propositional* contents (believables) are accordingly to be picked out by the pragmatic property of being assertible. Likewise, what is uttered or inscribed in producing an assertional performance is thereby recognizable as a declarative *sentence*. The role of *propositional* contents marks off *discursive* practice, and the role of *sentential* expressions of such contents is distinctive of *linguistic* social practice. In this way, everything comes down to being able to say what it is for what practitioners are doing to deserve to count as adopting a practical attitude of acknowledging the *assertional* significance of a performance: taking or treating it as an assertion. It is in terms of such attitudes that the pragmatic significance of assertional speech acts, the normative status of assertional commitments, and the possession or expression of propositional semantic contents are to be understood.

That the contents conferred by those practices are recognizable as *discursive* or *conceptual* contents (the genus of which propositional contents form the most basic species) depends on their *inferential* articulation and relation

to each other. The practices that constitute the significance characteristic of assertional performances and the status characteristic of assertional commitments must be inferential practices. Asserting cannot be understood apart from inferring. So one fundamental question is, What makes something that is done according to a practice—for instance the production of a performance or the acquisition of a status—deserve to count as *inferring*? The answer developed here is that inferring is to be distinguished as a certain kind of move in the game of giving and asking for reasons. To say this is to say that inferring should be understood in an interpersonal context, as an aspect of an essentially *social* practice of communication.

The contentfulness of the states attributed as part of a simple intentional interpretation of an individual consists in a sort of inferential articulation that is not intelligible solely in terms of the role those states play in practical reasoning—if practical reasoning is conceived of as restricted to the sort of intrapersonal instrumental deliberation implicitly imputed by such interpretation. The explanation of behavior according to the model of rational agency depends on treating attributed intentional states as having propositional contents, which involve objective truth conditions. But there is, it will be argued, no pattern of moves a single individual might make that would qualify that individual's states as inferentially articulated in this sense. The inferential practice (including practical reasoning) that confers contents of this kind comprises not only first-person reasoning but also third-person attributions and assessments of it—and both aspects are essential to it. Deliberation is the internalization of the interpersonal, communicative practice of giving reasons to and asking reasons of others, just as judgment is the internalization of a public process of assertion. Inferring cannot be understood apart from asserting. To say this is to say that inferring should be understood as an aspect of an essentially *linguistic* practice. The practice of giving and asking for reasons must be conceived as including assertion because, although there are other kinds of performances besides assertings that can stand in need of reasons—indeed for a performance to be an action just is for it to be something it is appropriate to demand a reason for—to offer a reason is always to make an assertion.

The rest of this chapter is given over to developing a model of assertional and inferential practice. What is described is not our actual practice but an artificial idealization of it. Simplified and schematic though the model may be, it should nonetheless be recognizable as a version of what we do. The model is intended to serve as the core of a layered account of linguistic practice. Where our practice diverges from that specified in the model, those divergences should be explicable as late-coming additions to or modifications of the underlying practice. For instance, the model appeals only to semantic inferences, that is, inferences involving what is claimed. Pragmatic inferences such as Gricean implicatures have to do rather with the antecedents and consequents of the performance of claiming it. These pragmatic inferen-

tial practices form a shell around the more basic semantic ones, which they presuppose. The critical criterion of adequacy the model answers to is that the core linguistic practices it specifies be *sufficient* to confer propositional and other conceptual contents on the expressions, performances, and deontic statuses that play appropriate roles in those practices. It is also claimed, however, that the fundamental structural features of the model provide *necessary* conditions for the conferral of such contents. So there is a subsidiary commitment to the effect that sophisticated linguistic practices of the sort not addressed by the model are ultimately intelligible only by showing how they could develop out of the sort of practices the model does specify.

## 2. *Commitment and Entitlement*

At the core of discursive practice is the game of giving and asking for reasons. Chapter 1 sought to motivate the claim that discursive practice is implicitly *normative*; it essentially includes assessments of moves as correct or incorrect, appropriate or inappropriate. The institution of these proprieties by practical assessments on the part of the practitioners is the ultimate source of the meanings of the noises and marks they make, and of the other things they do.<sup>18</sup> As the term is used here, to talk of practices is to talk of *proprieties* of performance, rather than of regularities; it is to prescribe rather than describe. The general notion of proprieties of practice in terms of which the discussion of implicit norms has been conducted up to this point, however, does not cut fine enough to pick out what is distinctive of discursive norms. For that purpose the pragmatics Dummett suggests—which specifies the significance of linguistic expressions (and implicitly of speech acts and alterations of intentional states) in terms of circumstances of *appropriate* application and *appropriate* consequences of such application—must be further refined. Different sorts of propriety must be acknowledged.

The fundamental normative concept required is the notion of *commitment*. Being committed is a normative status—more specifically a *deontic* status. The project of the central sections of this chapter is to introduce a notion of *discursive commitment* as a species of deontic status that can do much of the explanatory theoretical work that is normally assigned to the notion of intentional state. But deontic statuses come in two flavors. Coordinate with the notion of *commitment* is that of *entitlement*. Doing what one is committed to do is appropriate in one sense, while doing what one is entitled to do is appropriate in another. The model of linguistic practice described here elaborates on the Dummettian bipartite pragmatics by distinguishing on the side of consequences, for instance, what a particular speech act commits one to from what it entitles one to. This permits a finer-grained specification of functional roles in linguistic practice than does using a single-sorted notion of propriety of circumstances and consequences of performance.

Commitment and entitlement correspond to the traditional deontic primitives of obligation and permission. Those traditional terms are avoided here because of the stigmata they contain betraying their origin in a picture of norms as resulting exclusively from the commands or edicts of a superior, who lays an obligation on or offers permission to a subordinate. Framed this way, the question of what one is obliged or permitted to do can slip insensibly into the question of who has a right to impose those statuses (as it does explicitly for Pufendorf, for instance). The picture presented here does not depend on a hierarchy of authority. The concepts of obligation and permission, as of duties and rights, can be reconstructed in terms of commitment and entitlement as they will be construed here.

Another way in which the treatment here of the deontic primitives of commitment and entitlement differs from that usually accorded to obligation and permission concerns the relation between them. It has been traditional to acknowledge the relations between these deontic modalities by defining one in terms of another: being permitted to do something is to be rendered as not being obliged not to do it, or being obliged as not being permitted not to. It does make sense to think of being committed to do something as not being entitled not to do it, but within the order of explanation pursued here it would be a fundamental mistake to try to exploit this relation to define one deontic status in terms of the other. Doing so requires taking a formal notion of negation for granted. The strategy employed here is rather to use the relation between commitment and entitlement (which are not defined in terms of this relation) to get a grip on a *material* notion of negation, or better, incompatibility. Two claims are incompatible with each other if commitment to one precludes entitlement to the other. One of the prime advantages normative-functional analyses of the notion of intentional states have over causal-functional analyses is that rendering the phenomenon of "inconsistent" beliefs<sup>19</sup> as incompatible commitments makes it intelligible in a way not available to causal accounts.

It may also be remarked, in a preliminary fashion, that supposing that sense can be made of the underlying deontic statuses of commitment and entitlement, the notion of material incompatibility of commitments they give rise to leads in a straightforward way to a notion of the *contents* of such commitments. For the content of a commitment can for many purposes be represented by the theorist as the set of commitments that are incompatible with it. For instance, a kind of entailment relation is induced on commitments by inclusion relations among such sets of incompatibles. The commitment  $p$  incompatibility-entails the commitment  $q$  just in case everything incompatible with  $q$  is incompatible with  $p$ . Thus "Wulf is a dog" incompatibility-entails "Wulf is a mammal," since everything incompatible with his being a mammal is incompatible with his being a dog. The notion of material incompatibility that is made available by *not* defining commitment and entitlement in terms of one another accordingly provides a route from

the pragmatics that deals with these deontic statuses to the semantics that deals with their contents. How this hint might be exploited will become clearer presently.

### *3. Attributing and Undertaking Commitments*

Deontic statuses of the sort to be considered here are creatures of practical attitudes. There were no commitments before people started treating each other as committed; they are not part of the natural furniture of the world. Rather they are social statuses, instituted by individuals attributing such statuses to each other, recognizing or acknowledging those statuses. Considered purely as a natural occurrence, the signing of a contract is just the motion of a hand and the deposition of ink on paper. It is the undertaking of a commitment only because of the significance that performance is taken to have by those who attribute or acknowledge such a commitment, by those who take or treat that performance as committing the signatories to further performances of various kinds. Similarly for entitlements. A license, such as a ticket, entitles one to do something. Apart from practices of treating people as entitled or not, though, there is just what is actually done. The natural world we consider when bracketing the influence of such social practices contains no distinction between performances one was entitled to and those one was not.

It will be useful to see how this basic vocabulary can be used to discuss the authority and responsibility involved in familiar sorts of deontic statuses. The way in which such statuses can be instituted by practical attitudes can be illustrated by artificially simplified versions of some fundamental normative practices. Authority may be considered first, apart from responsibility. A license, invitation, or entrance ticket entitles or authorizes one to do something one was otherwise not entitled to do. It is always a license in the eyes of someone, for example a ticket-taker or doorman. These "consumers" of licenses (along with the others whose attitudes make the practice into a going concern) constitute them by attributing the authority they thereby come to possess. They do so by treating the authorized one as entitled to a performance. It is not appropriate to enter unless one is authorized by a ticket acquired in the appropriate way. Being given a ticket by the ticket-giver is being authorized or entitled to enter, because and insofar as the ticket-taker will not treat entry as appropriate unless so authorized. In the simplest case, the ticket-taker is the attributor of authority, the one who recognizes or acknowledges it and who by taking the ticket as authorizing, makes it authorizing, so instituting the entitlement.

Practices of this sort can be described in purely responsive terms for prelinguistic communities. The entitlement given and recognized in these practices has a content for an attributor insofar as that attributor practically partitions the space of possible performances into those that have been

authorized and those that have not, by being disposed to respond differently in the two cases. These sanctioning responses (for instance admitting versus ejecting) and the performances they discriminate (enterings of the theater) can be characterized apart from and antecedent to specification of the practice of conferring and recognizing entitlement defined by their means. For this reason the entitling authority will be said to be *externally* defined. The sanctions applied in taking or treating someone as entitled can be specified in nonnormative terms.

The basic structure just considered involves entitlement without commitment, authority without responsibility. A corresponding way into the basic structure of commitments or responsibilities is provided by describing a similarly simplified and artificial version of an actual eighteenth-century British practice. According to this practice, taking "the queen's shilling" from a recruiting officer counts as committing the recipient to military service. A performance of this kind has the same significance that signing a contract would have—in either case one has joined the army and undertaken all the commitments entailed by that change of status. (The official rationale was that some such overt irrevocable nonlinguistic performance was required to do duty for signing a contract, given that those enlisting were largely illiterate. The actual function of the practice was to enable "recruiting" by disguised officers, who frequented taverns and offered what was, unbeknownst to their victims, the queen's shilling, as a gesture of goodwill to those who had drunk up all of their own money. Those who accepted found out the significance of what they had done—the commitment they had undertaken, and so the alteration of their status—only upon awakening from the resulting stupor.) The significance of a commitment is to be understood in terms of the practical attitude of those attributing it, that is in terms of what taking or treating someone *as* committed consists in. In this case, it consists in being liable for punishment by a court-martial if one fails to discharge or fulfill the commitment. The content of the commitment attributed corresponds to the subsequent behavior that would or would not elicit a sanction. Or more precisely, that content can be thought of according to Dummett's model of circumstances and consequences; the particular sanctions (court-martial) are the consequences, and the various failures to perform as a soldier that elicit them are the circumstances.

Two features of this simple commitment structure are worthy of note. First, for someone to *undertake* a commitment, according to this story, is to do something that makes it appropriate to *attribute* the commitment to that individual. That the performance of accepting the coin has the significance of altering the *status* of the one whose performance it is consists in the change it brings about in what *attitudes* are in order. It is by reference to the attitudes of others toward the deontic status (attributing a commitment) that the attitude of the one whose status is in question (acknowledging or undertaking a commitment) is to be understood. So all that is required to make

sense of the normative significance of the performance as an undertaking of commitment is an account of what it is to take or treat someone as committed to do something. The possibility of sanctioning failure to perform appropriately—that is, as one is (thereby) taken to be committed to do—offers a way of construing this fundamental practical deontic attitude.

Second, the basic notion of responsibility or commitment that is introduced by consideration of this simple practice can be understood in terms of the notion of authority or entitlement already discussed. For undertaking a commitment can be understood as authorizing, licensing, or entitling those who attribute that commitment to sanction nonperformance. Such sanctioning would be inappropriate (and so itself subject to sanction) unless it had been authorized by the undertaking of a commitment. One may not court-martial someone who has never joined the service. Thought of this way, the effect of undertaking a commitment is not a matter of in fact eliciting punishment if one does not fulfill the commitment but rather of making such punishment *appropriate*. It is not a matter of the actual conditional dispositions to sanction of those who attribute the commitment but a matter of the conditional normative status of such sanctions. What is being considered is a slightly more sophisticated practice, in which the significance of taking the queen's shilling (the consequences of undertaking a commitment and thereby acquiring a new deontic status) is itself defined in terms of deontic statuses—in particular, of *entitlement to punish*.

The significance of that deontic status (*entitlement to punish*) might itself be defined in normatively *external* terms; those who attempt to court-martial someone who has not committed himself to the service (and so entitled their superiors to hold such a court) are taken out and shot. Or the difference that entitlement makes might itself be cashed out only in normative terms; those who attempt to court-martial someone who has not committed himself to the service (and so entitled their superiors to hold such a court) thereby *license* or *entitle* their superiors to have them taken out and shot, make it *appropriate* or *correct* for them to be taken out and shot. Normatively *internal* definitions of the significance of changes of deontic status, which specify the consequences of such changes in terms of further changes of deontic status, link various statuses and attitudes into *systems* of interdefined practices.

#### 4. The Example of Promising

These points can be illustrated by considering an idealized version of a more familiar sort of practice of undertaking commitments, namely *promising*. Promising is another way of undertaking a task responsibility: committing oneself to perform in a certain way. A special defining feature of promising performances is that they involve specifying what one is committed to do by explicitly *saying* it—*describing* the performances that would

count as fulfilling the commitment. This feature of promising will not be officially intelligible in terms of the theory presented here until the practice of assertion has been introduced. For the promise offers a linguistic characterization of a performance, and that characterization (or a grammatical transform of it) must turn out to be assertible, on pain of the promisor's having failed to fulfill the commitment undertaken. Only one who claims that MacArthur returned will take it that MacArthur's promise to return was fulfilled. Furthermore, promises are typically made *to* someone. The promisees are the ones who are entitled to hold the promiser responsible.

A social-practical description of promising displays how the deontic attitudes of undertaking and attributing the commitment that is a promise are two sides of one coin, aspects of a single social practice. Promising (like taking the queen's shilling) can be understood as consisting in a dimension of responsibility and a coordinate dimension of authority. The authority of the licensing, in the eyes of the attributors of the promise-commitment, is an entitlement on the part of others to rely (as perhaps they were not previously entitled to do) upon the promised performance. The responsibility consists in conditionally entitling others to sanction the committed one. A commitment to perform includes a license to those who attribute the commitment to hold the committed one responsible for nonperformance. To be entitled to hold responsible is to be conditionally entitled to sanction, in case of nonperformance. In the simplest sort of promising practice, both the promised performances and the conditionally authorized sanctions can be specified antecedently (that is, defined externally to the definition of the practice that then ties them together). The promised performance might be washing the dinner dishes, and the conventionally authorized sanction for nonperformance might be being beaten with sticks.

Commitments to perform that are externally defined by practical attitudes manifested in fulfilling performances, and sanctions that are antecedently specifiable in nonnormative terms, share some features with assertional commitments. As already indicated, however, this category must be developed in the direction of a system of practices governing interlocking internally specified significances, defined in terms of fulfilling performances and sanctions. The toy promising practice just mentioned can be developed further by introducing an internally defined sanction for failure to perform as promised. Instead of responding to the failed promiser by beatings or refusals of entry to the theater, attributors of commitments to perform might rather withhold their recognition of that individual's entitlement to undertake such commitments. If the commitment is not fulfilled, the cost is that those who attribute both the commitment and the failure will not (or will become less disposed to) recognize the authorization (making entitled) of reliance on performances, that is, will not recognize the promises of the failed commitment undertaker as counting for anything. Undertaking the commitment is still doing something that conventionally has the significance of entitling others to attribute the commitment, that is both to attribute the commit-

ment and to hold the undertaking performer to it. One who succeeds in making a promise still authorizes others (makes them entitled) to rely on one's future performance, to hold one responsible for a failure to perform according to one's commitment.

These are both authority and responsibility, adding up to commitment. The commitment is undertaken by the one who authorizes or licenses, and undertaking is always licensing the attribution of—that is the holding of the undertaker responsible for—the commitment. Responsibility is of authorizer-undertaker to attributor. The difference is that in the sort of promising being considered, uttering certain words is not always sufficient to undertake the commitment that defines making a promise. For as a sanction consequential on previous failures to keep commitments undertaken by promises, attributors may withhold their recognition of the undertaker's entitlement to issue licenses of that sort, refuse to recognize as successful attempts to undertake such commitments. Unlike the queen's shilling case, no performance compels attribution of authority and responsibility—that is, commitment. Such a sanction is defined internally to the practice being considered, since apart from the practice of promising, one cannot specify what the sanction is.<sup>20</sup>

Part of the definition of what it is to undertake a commitment to perform by promising is the significance of this status for the deontic attitudes of others—the practical interpretation that attributors are entitled to place on the undertaking—namely their right to rely on future performance. The second sort of responsibility undertaken is not a task responsibility (as the commitment to perform could be called). It is a becoming liable to be held responsible (taken to be responsible) for failure to perform as one promised. What the promiser entitles others to do, in this simple practice, is to withdraw their recognition (taking) of one's entitlement to issue a license, of one's authority. When responsibility of this sort is added to authority of the invitation or ticket-giving variety, besides the conditional task responsibility, there must also be the special sort of authority involved in undertaking responsibility. This is authorizing others to hold the undertaker responsible. It is a license to do something, conditional upon the undertaker's failure to fulfill his or her task responsibility once its conditions are satisfied. A distinct and important species of such authority to hold responsible arises when the licensed consequences of failure to perform consist in withholding recognition of entitlement to undertake further responsibilities of the same form. Both promising, as reconstructed here, and asserting, as discussed below, exhibit this special commitment-entitlement structure.

### 5. Social Practice: Deontic Statuses and Deontic Attitudes

The discussion of these simplified examples introduced a number of points. The sort of implicitly normative practice of which language use is a paradigm is to be discussed in terms of two sorts of *deontic status*, namely

commitment and entitlement. The notion of normative status, and of the significance of performances that alter normative status, is in turn to be understood in terms of the practical deontic *attitude* of taking or treating someone as committed or entitled. This is in the first instance *attributing* a commitment or entitlement. Adopting this practical attitude can be explained, to begin with, as consisting in the disposition or willingness to impose *sanctions*. (Later, in more sophisticated practices, *entitlement* to such a response, or its *propriety*, is at issue.) Attributors of these statuses may punish those who act in ways they are not (taken to be) entitled to act, and those who do not act in ways they are (taken to be) committed to act. What counts as punishment may (according to the one who interprets a community as exhibiting practices of this sort) be specifiable in nonnormative terms, such as causing pain or otherwise negatively reinforcing the punished behavior. Or what counts as punishment with respect to a particular practice may be specifiable only in normative terms, by appeal to alterations in deontic status or attitude. A performance expresses the practical attitude or has the significance of an *undertaking* of a commitment in case it *entitles* others to attribute that commitment. So there are two sorts of practical deontic attitudes that can be adopted toward commitments: attributing them (to others) and acknowledging or undertaking them (oneself). Of these, attributing is fundamental.

Here, then, is a way of thinking about implicitly normative social practices. Social practices are games in which each participant exhibits various deontic statuses—that is, commitments and entitlements—and each practically significant performance alters those statuses in some way. The significance of the performance is how it alters the deontic statuses of the practitioners. Looking at the practices a little more closely involves cashing out the talk of deontic statuses by translating it into talk of deontic *attitudes*. Practitioners take or treat themselves and others as having various commitments and entitlements. They *keep score* on deontic statuses by attributing those statuses to others and undertaking them themselves. The significance of a performance is the difference it makes in the deontic score—that is, the way in which it changes what commitments and entitlements the practitioners, including the performer, attribute to each other and acquire, acknowledge, or undertake themselves. The significance of taking the queen's shilling lies in its being an undertaking of a commitment on the part of the recipient, altering the attributions of commitment by those who appreciate the significance of the performance. It entitles other authorities—those who according to the antecedent score already had undertaken various commitments or duties and entitlements or sorts of authority, those who therefore play a certain role or hold a certain office in the system of practices in question—to punish the performer in particular ways under particular circumstances. The normative significances of performances and the deontic states of performers are instituted by the practice that consists in keeping score by adopting attitudes of attributing and acknowledging them.<sup>21</sup>

### III. ASSERTING AND INFERRING

#### 1. *Linguistic Practice: Assertion and Inference*

The discussion of the significance of performances as altering the deontic attitudes that keep track of normative statuses has so far addressed implicitly normative social practices—whether or not they are specifically *linguistic* (and so more generally *discursive*) practices. What distinguishes the latter sort is the *inferential* articulation of the normative significances they involve—and so their conferral of specifically *conceptual* content on the states, attitudes, performances, and expressions they govern. The challenge is to show how these two approaches (normative pragmatics modeled on deontic scorekeeping, and inferential semantics) can be combined into a single story about social practices of treating speech acts as having the significance of assertions.

Describing practices sufficient to institute such a significance is the way to fill in the notion of assertional commitment. Such an account provides an answer to the question, What is it that we are *doing* when we assert, claim, or declare something? The general answer is that we are undertaking a certain kind of *commitment*. Saying specifically *what* kind is explaining what structure must be exhibited by the practices a community is interpreted as engaging in for that interpretation to be recognizable as taking the practitioners to be keeping score for themselves and each other in virtue of the alterations of their practical deontic attitudes of attributing and undertaking assertional commitments and their corresponding entitlements.

The key to seeing how the scorekeeping model of deontic social practices can be used to make sense of asserting is Sellars's notion of a "game of giving and asking for reasons." The idea is that assertings (performances that are overt undertakings of assertional commitments) are in the fundamental case what reasons are asked for, and what giving a reason always consists in. The kind of commitment that a claim of the assertional sort is an expression of is something that can stand in need of (and so be liable to the demand for) a reason; and it is something that can be offered as a reason. This is the principle motivating the present strategy for discriminating assertional commitments from other species of commitment. Other things besides assertional commitments involve liability to demands for justification or other demonstration of entitlement—for instance, the practical commitments involved in *actions*. Other things besides assertional commitments can entitle interlocutors to assertional commitments—for instance *reliability* in the responsive acquisition of assertional commitments of a certain kind. For being a reliable reporter of currently visible red things who responsively acquires a disposition to claim that there is something red in the vicinity may entitle someone to that commitment.<sup>22</sup> But only assertional commitments stand in *both* these relations.

That claims play the dual role of justifier and subject of demand for justification is a necessary condition of their kind properly being called

*assertional* commitments. It is here employed as well as part of a sufficient condition, in an idealized artificial practice constructed to model this central aspect of the use of natural language. Specifically *linguistic* practices are distinguished as just the social practices according to which some performances have the significance of undertakings of assertional commitment (in virtue of their role in giving and asking for reasons); *declarative sentences* are picked out as the expressions uttered or inscribed in such assertional performances. What is expressed by such performances and determine the particular features of their significance within the assertional genus count thereby as *propositional* contents.

The idea exploited here, then, is that *assertions* are fundamentally fodder for *inferences*. Uttering a sentence with assertional force or significance is putting it forward as a potential reason. Asserting is giving reasons—not necessarily reasons addressed to some particular question or issue, or to a particular individual, but making claims whose availability as reasons for others is essential to their assertional force. Assertions are essentially *fit* to be reasons. The function of assertion is making sentences available for use as premises in inferences. For performances to play this role or have this significance requires that assertional endorsement of or commitment to something entitles or obliges one to other endorsements. The pragmatic significance of assertional commitments and entitlements to such commitments consists in the ways in which they are *heritable*; their heritability is the form taken by the *inferential* articulation in virtue of which they count as semantically contentful.

## 2. Three Dimensions of Inferential Articulation

The basic model of the inferential practices that institute assertional significance—and thereby confer propositional contents on states, attitudes, and performances playing suitable roles in those practices—is defined by a structure that must be understood in terms of the interaction of three different dimensions. First, there are two different sorts of deontic status involved: *commitments*, and *entitlements* to commitments. Inheritance of commitment (being committed to one claim as a consequence of commitment to another) is what will be called a *committive*, or commitment-preserving, inferential relation. Deductive, logically good inferences exploit relations of this genus. But so do materially good inferences, such as inferences of the form: *A* is to the West of *B*, so *B* is to the East of *A*; This monochromatic patch is green, so it is not red; Thunder now, so lightning earlier. Anyone committed to the premises of such inferences is committed thereby to the conclusions.

Inheritance of entitlement (being entitled to one claim as a consequence of entitlement to another) is what will be called a *permissive*, or entitlement-preserving inferential relation. Inductive empirical inferences exploit rela-

tions of this genus. The premises of these inferences entitle one to commitment to their conclusions (in the absence of countervailing evidence) but do not compel such commitment. For the possibility of entitlement to commitments incompatible with the conclusion is left open. In this way the claim that this is a dry, well-made match can serve as a justification entitling someone to the claim that it will light if struck. But the premise does not commit one to the conclusion, for it is compatible with that premise that the match is at such a low temperature that friction will not succeed in igniting it. The interplay between the two sorts of deontic status is at the center of the model of assertional and inferential practices presented here.

The broadly inferential roles that are identified with propositional contents involve not only commitment- and entitlement-preserving inferential connections among such contents but also relations of *incompatibility*. To say that two claims have (materially) incompatible contents is to say that commitment to one precludes entitlement to the other.<sup>23</sup> No candidate notion can count as a construal of the sort of propositional content we take assertions, judgments, and beliefs to have unless it underwrites incompatibility relations among them. (That possible worlds, for instance, must be understood as corresponding to maximal sets of *compatible* propositions is acknowledged both by those who want to exploit that principle to define propositions and their material compatibility in terms of possible worlds and by those who would reverse that order of explanation.) The explanatory strategy adopted here is to begin with practices that institute deontic statuses of commitment and entitlement and then to show how those practices thereby confer specifically propositional conceptual contents on what is assertible—contents recognizable as such in virtue of the deontic inferential and incompatibility relations they stand in.

The second dimension of broadly inferential articulation that is crucial to understanding assertional practice turns on the distinction between the *concomitant* and the *communicative* inheritance of deontic statuses. This is the *social* difference between *intrapersonal* and *interpersonal* uses of a claim as a premise. Undertaking a commitment or acquiring an entitlement has consequences for the one whose statuses those are. One commitment carries with it other concomitant commitments as consequences. Its consequences are those that it entails according to the commitment-preserving inferential relations that its content stands in to other contents of possible commitments. Similarly, being entitled to a commitment can entitle one to others, which stand to it in suitable permissive or entitlement-preserving inferential relations. Again, the definition of incompatibility of contents in terms of commitment and entitlement means that acquiring a commitment may have as a consequence the loss of entitlement to concomitant commitments one was heretofore entitled to.

But these intrapersonal inferential consequences of changes in deontic status do not exhaust the significance of assertional performances. Such

performances also have a significance for interpersonal *communication*. Putting a sentence forward in the public arena *as true* is something *one* interlocutor can do to make that sentence available for *others* to use in making further assertions. Acknowledging the undertaking of an assertional commitment has the *social* consequence of licensing or entitling others to *attribute* that commitment. The adoption of that deontic attitude on the part of the audience in turn has consequences for the commitments the audience is entitled to undertake. Putting a claim forward *as true* is putting it forward as one that it is appropriate for others to *take true*, that is to endorse themselves. Assertion that is communicatively successful in the sense that what is put forward as true by a speaker is taken as true by the audience consists in the interpersonal inheritance of commitment.

The third dimension of broadly inferential articulation that is crucial to understanding assertional practice is that in which discursive *authority* is linked to and dependent upon a corresponding *responsibility*.<sup>24</sup> In uttering a sentence assertively, the claim one makes involves an *endorsement*. One aspect of this sort of endorsement was indicated above in a preliminary fashion in terms of the function of an asserting as licensing or *authorizing* further assertions (and eventually, actions—but consideration of practical rather than doxastic commitments is postponed until the next section). But unless some independent grasp is offered of the status or significance that must be bestowed on a performance for it to count as an asserting, invoking the inferential warranting of further assertions merely traces out a rather small circle. It is the second aspect of endorsement, of the sort of *responsibility* involved in assertional commitment, that permits a larger horizon. Understanding that aspect requires putting together the distinction between the deontic statuses of commitment and entitlement, on the one hand, and of intra- and interpersonal inference, on the other. Putting forward a sentence as true or as information—that is, asserting it—has been glossed as putting it forward as fit to be a reason for other assertions, making it available as a premise from which others can be inferred. This means that others can inherit *entitlement* to an assertional commitment from the one who makes an assertion and thereby licenses or warrants its reassertion and the assertion of what follows from it. To understand this warranting function, the heritability of entitlement, requires understanding the social significance of the distinction between warranted and unwarranted assertional commitments.

Ordinarily the relation of an authorizing event to the performances it licenses requires at least that in the context of that event, performances become socially appropriate that otherwise would not be. For example purchasing a ticket entitles one to take a seat in the theater, which it would be inappropriate to do without the ticket. This observation presents a dilemma. If asserting a sentence is not a performance requiring prior authorization, then it seems one cannot understand the function of assertion as inferentially licensing further assertions. If, however, asserting is a performance requiring

authorization, how does one become entitled to the original licensing assertion? Talk of *inheritance* of entitlement makes sense only in an explanatory context that includes a story about the significance of *possession* of entitlement. It is this question that is addressed by an account of the dimension of *responsibility* characteristic of asserting. In asserting a claim, one not only authorizes further assertions (for oneself and for others) but undertakes a responsibility, for one commits oneself to being able to vindicate the original claim by showing that one is entitled to make it. Others cannot inherit an entitlement that the asserter does not possess. Overtly acknowledging or undertaking a doxastic commitment by issuing an assertional performance can warrant further commitments, whether by the asserter or by the audience, only if that warranting commitment itself is one the asserter is entitled to. Only assertions one is entitled to make can serve to entitle anyone to their inferential consequences.

The function of asserting as the giving of reasons is intelligible only as part of a practice in which reasons can be asked for or required. That some performances admit or stand in need of reasons is presupposed by the practice of offering them. Many kinds of performances are subject to demands for or explanation according to reasons. The two fundamental sorts discussed here are intentional actions that are not speech acts and assertions themselves. Both actions and assertions—overt performances corresponding to practical and to doxastic commitments respectively—are essential and ineliminable aspects of discursive practice as here construed. Nonetheless, the significance of assertional performances can be filled in to a considerable extent before it is necessary to look at the role of assertions as reasons for anything other than more assertions.

The converse is not the case. Actions just are performances for which it is appropriate to offer reasons, and offering a reason is making an assertion. So actions are not intelligible as such except in a context that includes assertional giving of reasons. Where intentional explanations are offered of the behavior of nonlinguistic creatures (those that are not understood as interpreters of others), the reasons are offered, the assertions are made, by the interpreter of the simple intentional system, who seeks to make its behavior intelligible by treating it *as if* it could act according to reasons it offers itself. That is why what is attributed in such interpretations is *derivative* intentionality. Assertions play both roles; reasons can be offered for them, and they can be offered as reasons. Actions play only the first role in the game of giving and asking for reasons. Thus a description of that game that has a certain autonomy can be offered if, to begin with, attention is restricted to assertions alone, and as a result asserting has a certain explanatory priority over action. Although for this reason the discussion of actions and the practical commitments they express is postponed until the next chapter, it remains that asserting is a *doing*, and the responsibility it involves should also be understood as a responsibility to *do something*.

Besides its specifically linguistic use, 'assert' has a broader normative use according to which one can assert one's authority or one's rights. An important component of this sense of asserting is defending, championing, or justifying. Milton uses the word this way in the famous expression of his intent in *Paradise Lost*: "That to the hight of this great Argument I may assert Eternal Providence, and Justifie the wayes of God to men." This use points to the fact that the sort of commitment involved in linguistic asserting involves the undertaking of a specifically *justificatory* responsibility for what is claimed. In asserting a sentence, one not only licenses further assertions (for others and for oneself) but commits oneself to *justifying* the original claim. The responsibility in question is of the sort Baier calls "task-responsibility"; it requires the performance of a task of some kind for its fulfillment.<sup>25</sup> Specifically, in making a claim, one undertakes the conditional task responsibility to demonstrate one's entitlement to the claim, if that entitlement is brought into question. Justifying the claim when it is queried, giving reasons for it when reasons are asked for, is one way to discharge this obligation. If the commitment can be defended, entitlement to it demonstrated by justifying the claim, then endorsement of it can have genuine authority, an entitlement that can be inherited.

### 3. Assertion as a Doing: Authority and Responsibility

The position maintained here is that *discursive* (in the Kantian sense of concept-mongering) practice can only be *linguistic* practice, and that what distinguishes a practice as specifically linguistic is that within it some performances are accorded the significance of *assertions*. It is only because some performances function as assertions that others deserve to be distinguished as *speech acts*. The class of questions, for instance, is recognizable in virtue of its relation to possible answers, and offering an answer is making an assertion—not in every individual case, but the exceptions (for example, questions answered by orders or by other questions) are themselves intelligible only in terms of assertions. Orders or commands are not just performances that alter the boundaries of what is permissible or obligatory. They are performances that do so specifically by *saying* or *describing* what is and is not appropriate, and this sort of making explicit is parasitic on claiming. Saying "Shut the door!" counts as an order only in the context of a practice that includes judgments, and therefore assertions, that the door is shut or that it is not shut. (The "slab" *Sprachspiel* Wittgenstein describes in the opening paragraphs of the *Investigations* is not in this sense a *language game*—it is a set of practices that include only vocal, but not yet verbal, signals.) In the same way, promises are not just undertakings of responsibility to perform in a certain way. They are performances that undertake such responsibility by *saying* or *describing* explicitly what one undertakes to do. One promises in effect to make a proposition true, and the propositional

contents appealed to can be understood only in connection with practices of saying or describing, of taking-true—in short, of asserting what are, in virtue of the role they play in such assertions, declarative sentences. As it is with these examples, so it is with other speech acts. Asserting is the fundamental speech act, defining the specific difference between linguistic practice and social practices more generally.

A crucial measure according to which a theory of speech acts ought to be assessed, then, is its treatment of what one is *doing* in producing an assertion. This challenge is not always accepted. One prominent theorist defines the assertion of the declarative sentence *p* as “an undertaking to the effect that *p*.<sup>26</sup> One does not have to subscribe to the pragmatist project of explaining the propositional contents that are asserted in terms of the practices of asserting them in order to find this disappointing. What sort of an undertaking is this? What, exactly, is the effect? The theory being presented here aims to answer just these questions.

In producing assertions, performers are doing two sorts of things. They are first *authorizing* further assertions (and the commitments they express), both concomitant commitments on their part (inferential consequences) and claims on the part of their audience (communicational consequences). In doing so, they become responsible in the sense of answerable for their claims. That is, they are also *undertaking* a specific task *responsibility*, namely the responsibility to show that they are *entitled* to the commitment expressed by their assertions, should that entitlement be brought into question. This is the responsibility to *do* something, and it may be fulfilled for instance by issuing other assertions that *justify* the original claim. The semantic content of the commitment expressed by the performance—that the authority it claims and the justificatory responsibility it undertakes are specifically “to the effect that *p*” (rather than some other *q*)—consists in its specific inferential articulation: what else it commits the asserter to, what commitments it entitles its audience to, what would count as a justification of it, and so on. On this account, then, the pragmatic force or significance characteristic of asserting (and therefore ultimately also the concepts of declarative *sentence*, *propositional content*, and specifically *linguistic practice*) is to be understood in terms of performances with the dual function in the game of giving and asking for reasons of being givings of reasons, and themselves also performances for which reasons can be asked. The conceptual contents expressed by assertional performances are to be explained by appeal to the inferential roles they play in that reason-mongering practice. What is *done* in asserting—the pragmatic *significance* or *effect* of producing an assertional performance—consists in the way in which, by authorizing particular further inferentially related performances and undertaking responsibility to produce yet other inferentially related performances, asserters alter the score interlocutors keep of the deontic statuses (commitments and entitlements) of their fellow practitioners.

The analysis being suggested divides this significance into a component having to do with *authority* and one having to do with *responsibility*. The particular way these components are intertwined defines the sort of pragmatic significance that is being identified with assertional force. The constellation of authority and responsibility characteristic of the assertional significance of speech acts is *socially* articulated. In producing an assertion, one undertakes a responsibility oneself. The authority of that performance (which is conditioned on the responsibility) in turn consists in opening up a new avenue along which those in the audience can fulfill the responsibilities associated with *their* assertions. At the core of assertional practice lie three fundamental ways in which one can demonstrate one's entitlement to a claim and thereby fulfill the responsibility associated with making that claim. Two of these—justifying the content of an assertion and deferring to the authority of an asserter—can be discussed here. The third—invoking one's own authority as a reliable noninferential reporter—is best left for later.

First, as already mentioned, one can demonstrate one's entitlement to a claim by *justifying* it, that is, by giving reasons for it. Giving reasons for a claim always consists in making more claims: asserting premises from which the original claim follows as a conclusion.<sup>27</sup> Interlocutors who accept such a vindication of the commitment—who accept the reasons offered as a justification demonstrating entitlement to the conclusion—thereby implicitly endorse a certain inference. These practical attitudes of taking or treating inferences as correct (distinguishing them from the incorrect ones by responses to attempted justifications) institute inferential proprieties relating the performances of asserters (and the commitments adopted thereby) and so confer contents on them. For it is the practical inferential proprieties acknowledged by such attitudes that make noises and marks mean what they mean. Assertions play a dual role in justification: as justifiers and as justifieds, premises and conclusions. That it plays this dual role, that it is caught up in justificatory inferences both as premise and as conclusion, is what makes it a specifically *propositional* (= assertible, therefore believable) content at all. That it exhibits the particular inferential grounds and consequences that it does is what makes it the particular determinate content that it is—settling, for instance, what information it conveys, the significance that undertaking a commitment with that content would have for what else one is committed and entitled to. Thus the inferential articulation of speech acts depends on this practice of demonstrating entitlement to the commitment acknowledged by the performance of a speech act.

The second way of vindicating a commitment by demonstrating entitlement to it is to appeal to the authority of another asserter. The *communicational* function of assertions is to license others who hear the claim to reassert it. The significance of this license is that it makes available to those who rely on it and reassert the original claim a special way of discharging their responsibility to demonstrate their entitlement to it. They can invoke

the license or authority of the asserter, thus deferring to the interlocutor who communicated the claim and passing along to that other individual any demands for demonstration of entitlement. The authority of an assertion includes an offer to pick up the justificatory check for the reassessments of others. That *A*'s assertion has the social significance of authorizing *B*'s reassertion consists in the appropriateness of *B*'s deferring to *A* the responsibility to demonstrate entitlement to the claim. *B*'s responsibility can be discharged by the invocation of *A*'s authority, upon which *B* exercised the right to rely. The buck is passed to *A*.

So communication does not involve only the sharing of *commitments*—their spreading from one individual to another as the speaker who produces an assertion communicates to, and possibly infects, an audience. It involves also the way that *entitlements* to claims can be inherited by the consumers of an assertion from its producer. In this way the *authority* of an assertional performance consists in part in making available a new way in which those to whom it is communicated can discharge their *responsibility* for demonstrating entitlement to commitments they undertake. Furthermore, assertions can be seen to play a dual role on the side of communication, just as they do on the side of justification. For assertions are on the one hand what is communicated (made available to others), and on the other hand they are what communication is *for*: one interlocutor's claim is fodder for inferences by others to further claims. The audience not only attributes to the one producing an assertional performance commitment to claims entailed (according to commitment-preserving inferences) by the assertion, but it also may undertake commitments and acquire entitlements that are its consequences.

In such inheritance of entitlement by communication, the content of the commitment is preserved intact and merely transferred from one scorekeeper to another. The communicational mechanism for fulfilling the responsibility to demonstrate entitlement appeals to *interpersonal*, *intracontent* inheritance of entitlement to a propositional commitment. By contrast, the justificatory mechanism appeals to *intrapersonal*, *intercontent*<sup>28</sup> inheritance of entitlement to a propositional commitment—since the contents of premises and conclusions will differ in any inference that is nontrivial in the sense of being available to do justificatory work. This combination of the *person-based* authority (invoked by *deferring* to the claim of another) and *content-based* authority (invoked by *justifying* the claim through assertion of other sentences from which the claim to be vindicated can appropriately be inferred) is characteristic of asserting as a doing. This constellation—of commitment and entitlement, of authority and responsibility, and of an inheritance of entitlement to assertional commitments that is *interpersonal* and *intracontent* as well as *intrapersonal* and *intercontent*—constitutes a fundamental substructure of the model of assertional practices presented here.

#### 4. The Default and Challenge Structure of Entitlement

More clearly needs to be said about the practices that govern the attribution of entitlement to assertional commitments. The two mechanisms considered so far for demonstrating such entitlement are ways in which entitlement to commitments can be *inherited*. Entitlement to commitment to one claim can be extended to entitlement to another either according to the *inferential* pattern appealed to by justification, in which case that entitlement is inherited by another commitment (to a different claim) undertaken by the same interlocutor, or according to the *communicational* pattern appealed to by deferral, in which case that entitlement is inherited by another commitment (to the same claim) by a different interlocutor. Tracing back an entitlement secured by either of these sorts of inheritance potentially sets off a regress.

The justificatory style of vindication, in which one interlocutor offers premises with different contents as reasons for a claim, threatens a regress on claim contents. At each stage vindication of one commitment may involve appeal to commitments that have not previously been invoked, for which the issue of demonstrating entitlement can arise anew, so the issue is merely put off. Or at some point a circle is closed by appeal to a set of premises whose entitlement has already been brought into question (and put off). Then the argument offered for a claim amounts to something that could be made explicit (eliminating the intervening steps) by a stuttering inference of the form “*p*, therefore *p*,” which cannot create entitlement.

The communicational style of vindication, in which one interlocutor appeals to another interlocutor’s assertional avowal of a commitment with the same content, threatens a corresponding regress on interlocutors. At each stage vindication of one interlocutor’s commitment may involve appeal to the commitment of some interlocutor who has not yet been appealed to, for whom the issue of demonstrating entitlement can arise anew—so the issue is merely put off. Or at some point a circle is closed by appeal to the assertion of some interlocutor whose entitlement has already been brought into question (and put off). Then the deferral that seeks to vindicate the claim amounts to something that could be made explicit (eliminating the intervening steps) by a self-citation of the form “I am relying for my entitlement to *p* on the authority of my own claims that *p*,” which cannot create entitlement.

The situation is not fundamentally altered by the fact that tracing back a single entitlement might involve both inferential and communicational appeals—that the chain of inheritance might comprise both justifications and deferrals. These are mechanisms for spreading entitlements, not for originating them; combining the two merely results in more complicated regresses and circles. What gets the process off the ground? What gives these multipli-cative mechanisms something to work with in the first place, so that chains of vindication can come to an end? This question arose above in connection

with the authorizing function of assertions and was pursued through the notion of fulfilling the responsibility to vindicate the authorizing commitment, by demonstrating one's entitlement. But looking at those mechanisms raises the same issue all over again.

The worry about a regress of entitlements is recognizably foundationalist. It can be responded to by appealing to the fundamental pragmatic commitment to seeing normative statuses (in this case entitlement) as implicit in the social practices that govern the giving and asking for reasons. Those practices need not be—and the ones that actually confer content on our utterances are not—such that the default entitlement status of a claim or assertional commitment is to be guilty until proven innocent. Even if all of the methods of demonstrating entitlement to a commitment are regressive (that is, depend on the inheritance of entitlement), a grounding problem arises in general only if entitlement is never attributed until and unless it has been demonstrated. If many claims are treated as innocent until proven guilty—taken to be entitled commitments until and unless someone is in a position to raise a legitimate question about them—the global threat of regress dissolves.

One of the lessons we have learned from thinking about hyperbolic Cartesian doubt is that doubts too sometimes need to be justified in order to have the standing to impugn entitlement to doxastic commitments. Which commitments stand in need of vindication (count as defective in the absence of a demonstration of entitlement to them) is itself a matter of social practice—a matter of the practical attitudes adopted toward them by the practitioners. The different circumstances under which various claims are taken or treated as requiring justification (or vindication by deferral) is part of what confers on the sentences that express them the different meanings that they have. It is part of the inferential role they play, in the broad practical sense of that expression, that includes the conditions under which inferential performances of various sorts are appropriate or obligatory. Claims such as “There have been black dogs” and “I have ten fingers” are ones to which interlocutors are treated as *prima facie* entitled. They are not immune to doubt in the form of questions about entitlement, but such questions themselves stand in need of some sort of warrant or justification. Entitlement is, to begin with, a social status that a performance or commitment has within a community.<sup>29</sup> Practices in which that status is attributed only upon actual vindication by appeal to inheritance from other commitments are simply unworkable; nothing recognizable as a game of giving and asking for reasons results if justifications are not permitted to come to an end.

The model presented here has what might be called a *default and challenge structure* of entitlement. Often when a commitment is attributed to an interlocutor, entitlement to it is attributed as well, by default. The *prima facie* status of the commitment as one the interlocutor is entitled to is not permanent or unshakeable; entitlement to an assertional commitment can

be challenged. When it is *appropriately* challenged (when the challenger is *entitled* to the challenge), the effect is to void the inferential and communicative authority of the corresponding assertions (their capacity to transmit entitlement) unless the asserter can vindicate the commitment by demonstrating entitlement to it.

This is what was meant by saying that the broadly justificatory responsibility to vindicate an assertional commitment by demonstrating entitlement to it is a *conditional* task-responsibility. It is conditional on the commitment's being subject to a challenge that itself has, either by default or by demonstration, the status of an entitled performance. Indeed, the simplest way to implement such a feature of the model of asserting is to require that the performances that have the significance of challenging entitlements to assertional commitments themselves be assertions. One then can challenge an assertion only by making an assertion incompatible with it. (Recall that two claims are incompatible just in case commitment to one precludes entitlement to the other.) Then challenges have no privileged status: their entitlement is on the table along with that of what they challenge. Tracing the provenance of the entitlement of a claim through chains of justification and communication is appropriate only where an actual conflict has arisen, where two *prima facie* entitlements conflict. There is no point fixed in advance where demands for justification or demonstration of entitlement come to an end, but there are enough places where such demands *can* end that there need be no *global* threat of debilitating regress.

This is the sort of picture of the practices of giving and asking for reasons that Wittgenstein suggests, but it is recognizable already in Socratic elenchus. In the present context, the proper question is not whether practices that incorporate such a default-and-challenge structure of entitlements are somehow in principle defective in view of some *a priori* rationalistic criterion of what it is to be *really* entitled to a claim. The proper question is rather, What sort of propositional contents can reason-constituting practices of this sort confer on the scorekeeping attitudes they govern, the deontic statuses they institute, and the performances they acknowledge as having the significance of assertions? The claim eventually to be made is that such practices suffice to confer objectively representational propositional contents on claims, objective truth conditions according to which the correctness of an assertion can depend on how things are with the objects represented by it, to the extent that the entire linguistic community could be wrong in its assessment regarding it.

##### 5. Internal Sanctions: Doxastic Commitments without Entitlements Lack Authority

The picture, then, is one in which giving reasons is obligatory only if they have been appropriately asked for. What has the significance of a

challenge (a demand for reasons) is just more assertions, whose entitlements are subject to the same sort of assessments as any others. All are weighed in the same balance. The fundamental concept in terms of which the default-and-challenge structure is adumbrated is the deontic attitude of *attributing* entitlement, of one interlocutor's taking or treating another as entitled to a commitment or performance. Now that the background presupposed by an interlocutor's conditional task-responsibility to demonstrate entitlement to the commitment undertaken by an assertional performance has been filled in, it is possible to be a bit clearer about this deontic attitude. The question is, What practical difference does it make whether the asserter is entitled to an assertional commitment? That is, What is the pragmatic significance of the distinction between warranted and unwarranted assertional commitments? What is it about a scorekeeper's treatment of an attributed commitment that makes it appropriate to describe that practical attitude as one of taking the commitment to be one the asserter is not *entitled* to?

Answering this question requires considering all three dimensions of assertion mentioned above: relations between commitments and entitlements, relations between intra- and interpersonal inferential significance, and relations between authority and responsibility. Since the *authority* of an assertion consists in its inferentially licensing or warranting further commitments, and this is a matter of *inheritance* of entitlement, an assertion expressing a commitment that is not taken to be one the performer is entitled to will not be taken to have inferential authority. Although making a claim by asserting a sentence is putting it forward as a fit premise for oneself and others to draw conclusions from, it will be accorded that status (its authority recognized) only by those scorekeepers who attribute not only the *commitment* the performance expresses but also an *entitlement* to that commitment. Absent such entitlement, assertion is an attempt to lend what one does not possess. Failure to shoulder the justificatory *responsibility* associated with entitlement to an assertional commitment (supposing it to have been appropriately challenged) renders void its *authority* as an inferential warrant for further commitments. Inferential authority and justificatory responsibility are coordinate and commensurate.

In the ideal Sprachspiel being described, making a claim one is not entitled to (even as a challenge) is a kind of impropriety, the violation of a norm. For a performance to have this sort of status or significance within or according to a set of practices—for this sort of norm to be implicit in or instituted by those practices—requires that the practices include attitudes of taking, treating, or acknowledging performances as incorrect in that particular way. Some sort of sanction must be available, with respect to which it can be specified what a practitioner *does* in adopting those practical attitudes. The practical significance of lack of entitlement consists in liability to *punishment* of some kind.

As has already been pointed out, however, such punishment need not

consist in *external* sanctions—responses such as beating with sticks, which are interpretable as punishments (perhaps in virtue of functioning as negative reinforcement in a behavioral-statistical sense) apart from the normative significance they in turn have within the practices in question. One can coherently interpret a community as engaging in practices in which performances are treated as having the significance of promises (or of the more primitive sort of nonlinguistic undertaking of task-responsibility, of which taking the queen's shilling is an example) even if the *only* sanction for failing to perform as one has committed oneself to do is to disqualify oneself from counting in the future as eligible to undertake such commitments. Something like this is what happened to the boy who cried "Wolf." Having several times committed himself to the claim that a wolf was present (thereby licensing and indeed obliging others to draw various conclusions, both practical and theoretical) under circumstances in which he was not entitled by the evident presence of a wolf to undertake such a commitment and to exercise such authority, the boy was punished—his conduct practically acknowledged as inappropriate—by withdrawal of his franchise to have his performances treated as normatively significant.

Unlike the case of the liar who ceases to be believed or of the irresponsible promiser who ceases to be relied upon, however, the internal sanctions constituting the practical recognition of an assertional performance as one the performer is not entitled to do not, on the present model, deal with the significance accorded to *other* performances of the same sort by that individual. Those sanctions have rather to do with the significance assigned to that very performance. Treating the commitment expressed by an assertional performance as one the asserter is not entitled to is treating it as not entitling that interlocutor and the audience of the performance to commitments whose contents follow inferentially from the asserted content. The practical sanction constitutive of the implicit norm governing entitlement to assertional commitments is internal to the system of scorekeeping attitudes the practice comprises. Taking someone to be (or not to be) entitled to a claim has consequences for what deferring performances one acknowledges as in order—and this in turn affects what deferrals one is oneself disposed to produce under various circumstances. But there need be no social pattern of performances and dispositions describable in nonnormative terms that is either necessary or sufficient for the constitution of such deontic attitudes.

#### IV. SCOREKEEPING: PRAGMATIC SIGNIFICANCE AND SEMANTIC CONTENT

##### 1. Lewis's Version of Scorekeeping in Language Games

The particular way in which the pragmatic significance of speech acts and deontic statuses is related to their semantic contents can be clarified

by looking more closely at the metaphor of *scorekeeping* by linguistic practitioners. The use made of the notion here is an adaptation of an idea introduced by David Lewis, in his paper "Scorekeeping in a Language Game." He suggests thinking about the rule-governedness of conversation by using some of the concepts appropriate to games that evolve according to a score function. This notion is explained in terms of baseball:

At any stage in a well-run baseball game there is a septuple of numbers  $\langle r_v, r_h, h, i, s, b, o \rangle$  which I shall call the *score* of the game at that stage. We can recite the score as follows: the visiting team has  $r_v$  runs, the home team has  $r_h$  runs, it is the  $h^{\text{th}}$  half ( $h$  being 1 or 2) of the  $i^{\text{th}}$  inning; there are  $s$  strikes,  $b$  balls, and  $o$  outs.<sup>30</sup>

The *constitutive rules* of the game are then of two sorts:

*Specifications of the kinematics of score.* Initially the score is  $\langle 0, 0, 1, 1, 0, 0, 0 \rangle$ . Thereafter, if at time  $t$  the score is  $s$ , and if between time  $t$  and  $t'$  the players behave in manner  $m$ , then at time  $t'$  the score is  $s'$ , where  $s'$  is determined in a certain way by  $s$  and  $m$ .

*Specifications of correct play.* If at time  $t$  the score is  $s$ , and between time  $t$  and time  $t'$  the players behave in manner  $m$ , then the players have behaved incorrectly. (Correctness depends on score: what is correct play after two strikes differs from what is correct play after three.) What is not incorrect play according to these rules is correct.<sup>31</sup>

He then points out that it is possible to use specifications of these sorts to define 'score' and 'correct play', by using the notion of *score function*, which is

the function from game-stages to septuples of numbers that gives the score at every stage. The specifications of the kinematics of score, taken together, tell us that the score function evolves in such-and-such way. We may then simply define the score function as the function which evolves in such-and-such way . . . Once we have defined the score function, we have thereby defined the score and all its components at any stage. There are two outs at a certain stage of a game, for instance, if and only if the score function assigns to that game-stage a septuple whose seventh component is the number 2.<sup>32</sup>

Correct play is specified in terms of current score and current behavior. Since the required relation between these is codified in the score function, it also defines correct play.

The idea is, then, that the evolution of a linguistic interchange or conversation can be thought of as governed by implicit norms that can be made explicit (by the theorist) in the form of a score function. Here are some of the analogies Lewis points to:

—Like the components of a baseball score, the components of a conversational score at any given stage are abstract entities. They may not be numbers, but they are other set-theoretic constructs . . .

—What play is correct depends on the score . . .

—Score evolves in a more-or-less rule-governed way. There are rules that specify the kinematics of score:

If at time  $t$  the conversational score is  $s$ , and if between time  $t$  and time  $t'$  the course of conversation is  $c$ , then at time  $t'$  the score is  $s'$ , where  $s'$  is determined in a certain way by  $s$  and  $c$ .

Or at least:

. . . then at time  $t'$  the score is some member of the class  $S$  of possible scores, where  $S$  is determined in a certain way by  $s$  and  $c$  . . .

—To the extent that conversational score is determined, given the history of the conversation and the rules that specify its kinematics, these rules can be regarded as constitutive rules akin to definitions. Again, constitutive rules could be traded in for definitions: the conversational score function could be defined as that function from conversation-stages to  $n$ -tuples of suitable entities that evolves in the specified way.<sup>33</sup>

As Lewis applies this idea, the elements of the conversational score are things such as sets of presupposed propositions and boundaries between permissible and impermissible actions. Then the acceptability of uttering a particular sentence at a given stage can depend on what is being presupposed. Similarly, the saliences established by the current score can determine the extension or even the intension of terms such as 'the pig'.

## *2. Deontic Scores and the Pragmatic Significance of Speech Acts*

This idiom can be adapted to the model of linguistic practice introduced in this chapter by specifying scores in terms of *deontic statuses*. Linguistic practice as here described can be explained in terms of a score function that determines how the *deontic score* at each stage in a conversation constrains both what performances are appropriate and what the consequences of various performances are—that is, the way they alter the score. The concept of the *pragmatic significance* of a speech act is central to the theoretical metalanguage being employed here. It is a generalization of Dummett's idea of specifying the use of an expression in terms of the pair of its *circumstances* of application and *consequences* of application. (Recall that Dummett's idea was adopted in Chapter 2 as a way of connecting a normative

pragmatics with an inferential semantics.) In scorekeeping terms, the significance of a speech act consists in the way it interacts with the deontic score: how the current score affects the propriety of performing the speech act in question, and how performing that speech act in turn affects the score. Deontic scores consist in constellations of commitments and entitlements on the part of various interlocutors. So understanding or grasping the significance of a speech act requires being able to tell in terms of such scores when it would be appropriate (circumstances of application) and how it would transform the score characterizing the stage at which it is performed into the score obtaining at the next stage of the conversation of which it is a part (consequences of application). For at any stage, what one is permitted or obliged to do depends on the score, as do the consequences that doing has for the score. Being rational—understanding, knowing how in the sense of being able to play the game of giving and asking for reasons—is mastering in practice the evolution of the score. Talking and thinking is keeping score in this sort of game.

In baseball the components of the score (for instance the status a performance can have as a *ball* or a *strike*, or an *out*) are defined in formal terms by the role they play in the process of keeping score—that is, their function in determining what counts as correct play according to the kinematics of score, as codified in the score function. So it is with the components of deontic score in terms of which linguistic practice is to be understood. The deontic statuses of commitment and entitlement are defined in formal terms by the way they can be used to keep track of the moves made in the game of giving and asking for reasons—that is, their function in determining what counts as correct play according to the kinematics of score, as codified in the score function. In this way the notion of *commitment* in linguistic practice plays a role like that of *strike* in baseball: each is an artificial, scorekeeping device.

Besides this formal characterization, however, there is also a material aspect to each of the components in the score, in virtue of which a particular performance qualifies as a ball, a strike, or an out. This material aspect is represented in Lewis's formulation of the kinematics of score quoted above by “the manner *m*,” which characterizes the behavior that changes the score from one stage to the next. In games such as baseball, which are not purely formal games (by contrast to chess or tic-tac-toe), the manner in which the score is changed cannot itself be specified entirely in terms of the concepts by means of which the score itself is specified. The complex manner in which a concrete performance qualifies as having the status of a strike or an out invokes such further concepts as the swinging of a bat, the passage of the baseball through a certain region of space specified relative to the position of the batter's body, catching the baseball on the fly, and so on. These further concepts give a material content to the scorekeeping concepts, beyond the formal content they have in virtue of their role in scorekeeping. So it is as well with deontic scorekeeping in linguistic practice. In that case the mate-

rial element concerns such issues as which utterances count as undertaking which commitments, or as deferring to the authority of another asserter, or as invoking noninferential responsive authority.

In baseball the application of scorekeeping vocabulary to particular performances is governed by rules, which are expressed largely in nonscorekeeping vocabulary. The use of this nonscorekeeping vocabulary accordingly answers to norms implicit in the practices of using that vocabulary in contexts other than baseball, where terms like 'inch', 'touch', 'between', and so on already have well-established circumstances of application. Final authority over the application of these rules is vested in the practice of officials. The exact character of this authority is a somewhat complicated matter. There is sometimes an inclination to think of it as *constitutive* of the correct application of the scorekeeping vocabulary, as it is taken to be for instance in the escalating claims of the competing umpires in the familiar tale:

*First Umpire*: I calls 'em as I sees 'em.

*Second Umpire*: I calls 'em as they is.

*Third Umpire*: Until I calls 'em, they ain't.

On such a view, the rules function as something like guides or advisory maxims for the judgment of the umpire, who *makes* a throw into a strike when he *takes* it as a strike. But though the attitude of the umpire does determine the status of a throw as a strike for official scorekeeping purposes (that is, does determine what the score is), the use of nonscorekeeping vocabulary in stating the rules that determine how the scorekeeping vocabulary *ought* to be applied to particular cases establishes a perspective from which the judgment of an umpire can nonetheless be understood to be *mistaken*. Metarules explicitly envisage the possibility of such mistakes and, without obliterating their status as such, set up a default-and-challenge system that leaves the umpire entitled to scorekeeping judgments even in the case where they are in fact mistaken, so long as they survive any appropriate challenges that are actually offered.

Linguistic scorekeeping on assertional commitments and entitlements has analogs to both of these dimensions of authority concerning the score. On the one hand, the actual attitudes of scorekeepers are essential in determining the score. On the other hand, the formation of those attitudes is itself subject to norms; scorekeeping is something that can be done correctly or incorrectly. This is not, of course, because it is in general governed by explicit rules; the regress that Wittgenstein and Sellars point to shows that. It is of the utmost importance to the present project to offer an account of what one is doing in taking a scorekeeper to have gotten things wrong, to have attributed commitments different from what the one to whom they are attributed is *really* committed to. For it is in terms of this practical attitude that the possibility of understanding the application of concepts as subject to objective representational constraint—as subject to assessment for being correct

or incorrect in a sense that involves answering to how the things the concepts are applied to actually are, rather than to anyone's attitudes toward them—is eventually to be explained.<sup>34</sup>

One fundamental difference between a game such as baseball and the game of giving and asking for reasons is the *perspectival* nature of the scorekeeping involved in the latter case. As Lewis sets things up, each stage of a baseball game has a single score. One might instead associate a different score with each of the two teams—though some elements, such as the specification of the inning, would be common to both teams at each stage. Linguistic scorekeeping as here construed is more like that: each interlocutor is assigned a different score. For to each, at each stage, different commitments and different entitlements are assigned. There may be large areas of overlap, since almost everyone is committed and entitled to such claims as that  $2 + 2 = 4$ , that red is a color, and that there have been black dogs. But there will also be large areas of difference, if for no other reason than that everyone has noninferentially acquired commitments and entitlements corresponding to different observable situations. These differences ramify both because of the inferential consequences of such observations and because their public availability for inheritance of attitudes and attribution based on testimony varies with conversational exposure. As a result, no two individuals have exactly the same beliefs or acknowledge exactly the same commitments. As with baseball, instead of thinking of these as different scores associated with different interlocutors (corresponding to different teams), they can be aggregated into one grand score for each stage of the conversation of a linguistic community, so long as it is kept clear (in a way corresponding to Lewis's use of subscripts) which deontic statuses are being attributed to which interlocutors.

But linguistic scorekeeping is also perspectival in a way that has no analog at all in baseball. Not only are scores kept for each interlocutor, scores are also kept by each interlocutor. In baseball there is just one official score, whether it is thought of as the score of the whole game or as the set of scores of each of the teams (and this is, as was pointed out, compatible with there nonetheless being a sense in which the umpire who determines the official score may make a mistake in calling a certain performance a strike). But part of playing the game of giving and asking for reasons is keeping track of the commitments and entitlements of the other players, by *attributing* commitments and entitlements. Just as each interlocutor is typically at each stage attributed a different set of deontic statuses, so each interlocutor typically has at each stage a different set of attitudes or attributions. What *C* is committed to according to *A* may be quite different, not only from what *D* is committed to according to *A*, but also from what *C* is committed to according to *B*. Linguistic scorekeeping practice is *doubly* perspectival.

The idea is that the deontic attitudes of each interlocutor *A* constitute one perspective on the deontic statuses of the whole community. There are, to

begin with, the commitments that *A* acknowledges and the entitlements that *A* claims. Then for each other interlocutor, there are the commitments and entitlements that *A* attributes to that individual. The different sorts of speech acts are to be understood in terms of the different consequences they have for the score that each interlocutor keeps, that is, in terms of how they affect the deontic attitudes of various interlocutors. If *B* asserts that *p*, *B* thereby acknowledges (and so undertakes) a commitment to *p*. So such a commitment ought to be attributed to *B* by anyone in a position to overhear or otherwise find out about that remark.

The pragmatic significance of an assertion goes far beyond this simple shift in deontic attitude on the part of other scorekeepers, however. For the speech act *B* performs has an inferentially articulated *content*, which relates it to other contents. Undertaking commitment to *p* is undertaking commitment as well to its inferential consequences—to those claims *q* that are related to it as conclusions of commitment-preserving inferences having *p* as premise. So if, as a result of *B*'s assertion, *A*'s deontic attitudes change in that *A* comes to attribute to *B* a commitment to *p*, then *A* is obliged also to attribute to *B* commitment to *q*. Or rather, *A*'s treating this as a good inference consists in *A*'s being disposed to keep score in this way, linking the attribution of commitment to *q* consequentially to attribution of commitment to *p*. Again, for *r* to be incompatible with *p* is for commitment to *p* to preclude entitlement to *r*. For *A* to treat these contents as incompatible is for *A* to be disposed to withhold attribution of entitlement to *r* whenever *A* attributes commitment to *p*.

Besides these intercontent, intrapersonal scorekeeping consequences of *B*'s speech act, the assertion may have intracontent interpersonal consequences regarding *A*'s attitudes. For if *A* takes *B* to be entitled to the claim that *p* (either noninferentially or as the conclusion of an inference), then this may result in *A*'s taking *C* (who also overheard the remark) to be entitled to that claim—but on the basis of testimony, to be defended by deference, rather than either noninferentially or inferentially. The effects of a speech act on the practical attitudes by means of which *A* keeps score on the deontic statuses of various interlocutors depends both on the antecedent score—what they were already taken to be committed and entitled to—and on the content expressed.

### *3. Inferentially Articulated Significance: Force and Content*

Specifying the pragmatic significance of a speech act kind such as assertion requires showing how the transformation of the score from one conversational stage to the next effected by such a speech act systematically depends on the semantic content of the commitment undertaken thereby. Starting with a notion of the pragmatic significance of speech acts—understood in terms of transformations of the deontic attitudes by which inter-

locutors keep track of each other's commitments and entitlements—it is possible to understand both what it is for two commitments to have the same content and what it is for two commitments to be undertaken by or attributed to the same interlocutor. Not only can these scorekeeping attitudes and shifts of attitude be used to define both contents and interlocutors, the justificatory and communicational links between them can be used to define the notion of *representation*. This is the burden of the discussion of the hybrid deontic attitudes that are made assertorially explicit in the form of *de re* ascriptions of propositional attitudes, in Chapter 8. So the notion of linguistic scorekeeping is intended to play a more fundamental explanatory role here than Lewis has in mind for it. For he is happy to think of conversational scores as kept track of in "mental scoreboards," consisting of attitudes he calls "mental representations" of the score (representations, presumably, whose content is *that* some component of the score is currently such and such).<sup>35</sup> Clearly he does not envisage a project such as the present one, in which both the nature of mental states such as belief and their representational contents are themselves to be understood in terms of their role in scorekeeping practices, rather than the other way around.

Consider first the notion of the *content* of a speech act or an intentional state. It is motivated first by the idea that speech acts, attitudes, and states of different kinds might share a content—in Fregean terminology, that different sorts of *force* can attach to the same sense. It requires further that the significance of a speech act depends in a systematic way on the content and the sort of force that is attached to it. Dummett's way of putting the point is this: "The implicit assumption underlying the idea that there is some one key concept in terms of which we can give a general characterization of the meaning of a sentence is that there must be some uniform pattern of derivation of all the other features of the use of an arbitrary sentence, given its meaning as characterized in terms of the key concept. It is precisely to subserve such a schema of derivation that the distinction between sense and force was introduced: corresponding to each different kind of force will be a different uniform pattern of derivation of the use of a sentence from its sense."<sup>36</sup> It does not simply go without saying that such a notion of content is to be had. Use of the theoretical concept of content involves a commitment to displaying the "uniform pattern of derivation of the use from the content," which Dummett talks about. As he goes on to indicate, one way of reading some of Wittgenstein's remarks is as "rejecting the whole idea that there is any one key idea in the theory of meaning: the meaning of each sentence is to be explained by a direct characterization of all the different features of its use; there is no uniform means of deriving all the other features from any one of them."<sup>37</sup> This is the point at which semantic theory and pragmatic theory must mesh. It is possible to associate many sorts of things with sentences and other linguistic expressions. What makes the association a *semantic* one is precisely the possibility of appealing to it to explain the

proprieties that govern the use of those expressions. Calling what one associates with expressions 'contents', 'propositions', 'sets of possible worlds', 'truth conditions', 'extensions', or 'referents' is at best issuing a promissory note that hints at how what are put forward as their semantic correlates ought to be taken to be relevant to determining how those expressions are correctly used. In the absence of a pragmatics offering an account of what it is to express a content or proposition, to take the actual world to be contained in a set of possible worlds, to try to utter truths, or to employ expressions with various extensions and referents, the theorist's entitlement to the commitment undertaken in treating these associations as semantic is liable to challenge. Semantics answers to pragmatics, attributions of content to explanations of use.

In the present case the pragmatics (comprising the practical proprieties governing linguistic expressions and intentional states alike) is couched in terms of social deontic scorekeeping. The *force* of an utterance, the *significance* of a speech act, is to be understood in terms of the difference it makes to what commitments and entitlements are attributed and undertaken by various interlocutors—that is, in terms of the alteration of deontic scorekeeping attitudes it underwrites. Indeed deontic statuses are to be understood just as ways of keeping such scores. The paradigmatic speech act kind of asserting is specified as having the significance of an undertaking of a commitment (and so the licensing of attributions of that commitment), the licensing or authorizing of further undertakings of such commitments, and the undertaking of a conditional task-responsibility to demonstrate entitlement to the commitment undertaken, if appropriately challenged. This is the sort of significance that must be determined "according to a uniform pattern" (in Dummett's phrase) by the sort of semantic content that is associated with the expressions that qualify as sentences in virtue of their freestanding utterances having this kind of assertional significance.

To be entitled to an *inferential* conception of the contents that qualify as propositional in virtue of their being assertible, then, requires showing how particular assertional significances result from the general account of the speech act of asserting when particular inferential roles are associated with what is asserted. The model of asserting has been constructed with just this criterion of adequacy in mind. To specify the inferential content associated with a sentence, one must, to begin with,<sup>38</sup> indicate the role it plays (in relation to the contents expressed by other sentences) in three different sorts of broadly inferential structure: committive inferences, permissive inferences, and incompatibilities. Doing so is saying what it follows from, what follows from it, and what it precludes or rules out. These are characterized as "broadly" inferential because all of them involve alterations of deontic status that have other alterations of deontic status as their consequences.

In the same sense, the sort of authority that observation reports exhibit counts as broadly inferential because of the reliability inference it involves on the part of the attributor of such authority (discussed in the next chapter).

Although it sounds paradoxical, for this reason the role of a sentence in noninferential reporting should also be understood as falling under the rubric “(broadly) inferential role.” Two features of such specifications are worth focusing on in this context. First, the informal explanations of these inferential relations are in terms of precisely the deontic statuses of commitment and entitlement that are kept track of by the scorekeeping attitudes of interlocutors—that is, just the terms in which the pragmatics is couched. Second, the account of assertional significance in general requires nothing more than inferential roles articulated along these three dimensions in order to determine the significance for social deontic scorekeeping of an assertional utterance.

On the first point, specifying the committive-inferential role of a sentence is specifying the commitment-preserving inferences in which that sentence serves as a conclusion, and those in which (along with other auxiliary hypotheses) it plays an essential role as a premise—essential in that if it is omitted, one could be committed to the remaining premises without therefore counting as committed to the conclusion. This sort of inference is the material-inferential genus of which deductively valid logical inferences are a formal species. Similarly, specifying the permissive-inferential role of a sentence is specifying the entitlement-preserving inferences in which (along with other auxiliary hypotheses) it plays an essential role—essential in that if it is omitted, one could be committed and entitled to the remaining premises (and to no incompatible defeasors) without therefore counting as being entitled to the conclusion. This sort of inference is the material-inferential genus of which inductively good inferences are a species. Incompatibility relations are thought of as broadly inferential because their construal in terms of the deontic statuses of commitment and entitlement is analogous to the first two. The difference is that where the deontic statuses relevant to the two species of strictly inferential relation are homogeneous—both premises and conclusions being assessed in terms of commitments, or both in terms of entitlements—in the case of incompatibility, they are heterogeneous. For to say that two claims are incompatible is to say that *if* one is committed to the first, *then* one is not entitled to the second.<sup>39</sup>

To anticipate the discussion of the next chapter: the sort of default entitlement characteristic of observation reports (perhaps the most important species of this genus) is thought of as broadly inferential because the one who attributes such authority implicitly endorses the reliability of the reporter (under these circumstances and with regard to such contents). Treating someone as a reliable reporter is taking the reporter’s *commitment* (to this content under these circumstances) to be sufficient for the reporter’s *entitlement* to that commitment. This is endorsing an inference, in the broad sense that corresponds to the consequential relationship between attributing commitment and attributing entitlement to it.<sup>40</sup> Like that involved in incompatibility relations among contents, and in contrast to the two strictly inferential components of content, this inference is heterogeneous with respect to the

deontic statuses involved. Unlike incompatibility, the content of the commitment and the content of the entitlement involved in the reliability inference are the same. The essential role played by endorsement by the attributor of this reliability inference qualifies the role of utterances as elicited observationally as the exercise of reliable differential responsive dispositions for inclusion as contributing to the broadly inferential role of sentences.

The connection between these four kinds of broadly inferential proprieties and the deontic statuses of commitment and entitlement has a double significance. On the one hand, it means that it is possible to understand how social practices of keeping score on commitments and entitlements could confer inferential roles articulated along these four dimensions on the expressions that are caught up in them. That is, social deontic scorekeeping provides an explanation of how expressions must be used in order to have contents of this sort associated with them—associated not by the theorist's stipulation but by the practical attitudes of the practitioners whose linguistic conduct is being interpreted. On the other hand, the connection between broadly inferential proprieties of these four kinds and the two sorts of deontic status they involve (homogeneously or heterogeneously) means that once contents articulated in terms of these proprieties have been associated with expressions, it is possible to derive “according to a uniform pattern” the significance that uttering them assertively has for scorekeeping in terms of commitments and entitlements. *Inferring* is accordingly the key concept linking semantic content and pragmatic significance. For not only can propositional semantic contents be understood as inferential roles, but properties of inference can be made sense of pragmatically, and specifically assertional significance can be understood in terms of them.

#### *4. How Inferential and Incompatibility Relations among Contents Affect the Score*

The significance of an assertion of  $p$  can be thought of as a mapping that associates with one social deontic score—characterizing the stage before that speech act is performed, according to some scorekeeper—the set of scores for the conversational stage that results from the assertion, according to the same scorekeeper.<sup>41</sup> Suppose  $A$  is such a scorekeeper, and  $B$  is such an asserter. The way  $A$ 's score ought to be transformed is settled by the content of  $B$ 's claim, thought of as its tripartite inferential role in commitment- and entitlement-preserving consequence relations and in incompatibility relations connecting commitments and entitlements. To begin with,  $A$  must add  $p$  to the list of commitments attributed to  $B$  (supposing the more interesting case in which  $A$  does not already attribute to  $B$  a commitment with that content).  $A$  should then add also commitment to any claims  $q$  that are committive-inferential consequences of  $p$ , in the context of the other claims attributed to  $B$ . These will vary, depending on the auxiliary hypotheses available, according to what other commitments  $A$  already attributes to

B. This is closing A's attributions to B under commitment-preserving inferences. This closure is determined, in the context of the prior score, by whatever committive-inferential role A associates with *p* as part of its content.<sup>42</sup>

Next, the incompatibility relations that *p* (and so its commitment-inferential consequences) stand in must be consulted to determine which, if any, of the entitlements A previously attributed to B are precluded by the newly attributed commitment. Assertions add new commitments, but they can not only add but also subtract entitlements. Then, in the light of the incompatibility relations associated with all of the commitments attributed to B, A can attribute *entitlements* to any claims that are committive-inferential consequences of commitments to which B is already taken to be entitled, closing the attributed score under commitment-preserving inferences, where the resulting attributions of entitlement are not defeated by incompatibilities.<sup>43</sup>

Next, constrained by the entitlement-precluding incompatibility relations associated with all of the other commitments attributed to B, A can attribute entitlements to any claims that are permissive-inferential consequences of commitments to which B is already taken to be entitled, closing the attributed score under entitlement-preserving inferences not defeated by those incompatibilities.<sup>44</sup> Then A needs to assess B's entitlement to the claim that *p*, by looking at good inferences having it as a conclusion and premises to which B is committed and entitled. This is determined by the first two elements of the broadly inferential role A associates with *p*. Similarly, A must assess the possibility of B's noninferential default entitlement to *p*. Whether B's undertaking of commitment to *p* falls within the scope of any good reliability inferences, according to A, depends on what else A is committed to—concerning the conditions under which the deontic status was acquired, implicitly, whether they are among those (if any) in which B is a reliable reporter with respect to contents such as *p*. Again, A must assess B's entitlement to *p* as testimony, by inheritance of the entitlement A attributes to some other interlocutor (possibly even A) who has asserted it at an earlier stage. If A takes B to be entitled to *p* by any of these mechanisms of inheritance and default, then A will take B to have successfully entitled others (including A) to that claim (in the absence of incompatible defeasors). In this way the broadly inferential content that A associates with B's claim determines the significance B's assertional speech act has from the point of view of A's scorekeeping, the difference it makes to the deontic attitudes of attributing and acknowledging commitments and entitlements by means of which A keeps track of everyone's deontic statuses.

### 5. *Deferrals, Disavowals, Queries, and Challenges*

The model of assertion defined by a scorekeeping function that appeals in this way to broadly inferential assertible contents to determine the significance of assertions of those contents can be enriched by allowing

various auxiliary sorts of speech acts. *Deferrals* have already been mentioned. No new sorts of content need be considered in order to specify the significance of deferring for social deontic scorekeeping. The content associated with a deferral is just the assertible content of the commitment that the deferrer is seeking to vindicate by indicating a testimonial path whereby entitlement to it can be inherited. It is the force associated with that content that is different from the assertional case.

*A* assesses *C*'s deferral to *B* concerning *p* by assessing first *B*'s entitlement to *p* (as considered above), and then *C*'s entitlement to inherit it. This latter is a matter not only of its being the case that *A* does not attribute to *C* commitment to anything incompatible with *p* (an issue in general independent of whether *A* attributes to *B* commitment to anything incompatible with *p*). For even if *A* does not attribute to *C* commitments incompatible with the *claim B* made (the commitment undertaken), it is still possible that *A* attributes to *C* commitments incompatible with inheriting *B*'s *entitlement* to it. This would happen if *B*'s entitlement, according to *A*, depends on justifying the claim *p* by appeal to the claim *q*, where *A* takes it that *C*, but not *B*, is committed to some claim incompatible with *q*. Similarly, it might be that *C* is committed, though *A* is not, to some claim incompatible with one of the conditions (according to *A*) for *B*'s observational authority with respect to *p*. Thus if *C* takes it that *B* is looking through a tinted window, *A* may take this to preclude *C*'s inheritance of entitlement to *B*'s noninferential report of the color of a piece of cloth, even though *A* takes it that *C* is wrong about the conditions of observation.

*Disavowals*, *queries*, and *challenges* are three other speech acts auxiliary to assertion that it is useful—from a scorekeeping point of view—to include in a model of the game of giving and asking for reasons. Disavowals permit one to repudiate or disclaim a commitment one has previously undertaken or to make it clear that one does not acknowledge such a commitment. Again no new sorts of content need to be considered; the force or significance of speech acts of this sort (the difference they make to the score that interlocutors keep on each other's deontic statuses) is determined by associating with them broadly inferential, assertible contents of the sort already discussed. It is the force of the speech act to which such contents attach that is different. For *A* to take *B*'s disavowal of commitment to *p* to be successful is for *A* to cease attributing commitment to *p* to *B*, and to reinstate any attributed entitlements that were withheld because they were defeated by their incompatibility with *p*.

Why might a disavowal not be successful? Because of the two fundamental ways *B* can undertake a commitment (and thereby license *A* to attribute it)—about which more below. For *B* might acquire commitment to *p* directly, by avowing it—that is, by overtly asserting it. Or *B* might acquire that commitment indirectly, as a consequence of a commitment (perhaps itself avowed) to *q*, from which it follows by a good commitment-preserving infer-

ence (according to *A*). In such cases, *B*'s disavowal of *p* can be successful (according to *A*) only if *B* is also prepared to disavow *q*. Indeed, disavowing *p* is indirectly disavowing *q*. But if *B* persists in asserting *q*, that commitment is incompatible with the disavowal, and the disavowal of *p* cannot accordingly reinstate entitlement to claims *A* withholds attribution of entitlement to only because of the flaw in *B*'s title represented by commitment to the incompatible *p*. Again, the sort of significance for *A*'s scorekeeping that disavowals of *p* have requires no further elements of content beyond those involved in its assertional use, even though the significances of the two sorts of speech act are quite different—indeed, in some ways complementary.

It would also be useful to those keeping score if there were some way of eliciting the avowal or disavowal of a particular claim—a way for *A* to find out whether *B* acknowledges commitment to *p*. Such a speech act is a basic query: *p?* By itself, such a speech act would have no effect on the deontic score; only responses to speech acts of this kind would alter the score. In the basic model, there is no reason not to allow anyone to be entitled to such a query at any point in a conversation.

Another sort of speech act that might be distinguished is *challenging* the testimonial authority of an assertion. As was indicated above, this might consist in no more than making an incompatible assertion. But it might be useful from a scorekeeping point of view to have a way of addressing an assertion as a challenge to another assertion. The significance of such a challenge is to bring attributions of entitlement by default into question wherever the challenging assertion is one the challenger is at least *prima facie* entitled to. For *A* to treat *C*'s challenge of *B*'s assertion of *p* as successful is for *A* to respond to it by withholding attribution of entitlement to *B* for that claim, pending *B*'s vindication of it, whether inferentially or deferentially. This has the effect of making that assertion unavailable (according to *A*'s score) to other interlocutors who might otherwise inherit entitlement to commitments to the same content testimonially from *B*. There is no reason in principle that conflicts of this sort need to be resolvable. The public status of competing claims may remain equivocal in that neither the challenged nor the challenging claim can be vindicated successfully, or in that both can be—though of course *A* will not take it that any one interlocutor could inherit entitlements to commitments to both of the incompatible contents. “Let a thousand flowers blossom. Let a hundred schools of thought contend.”

#### 6. Acknowledged and Consequential Commitments

In the next chapter the model of assertion is enriched by adding another variety of discursive commitment. Besides the cognitive commitments undertaken by assertions, *practical* discursive commitments—that is, commitments to *act*—are considered. The speech acts that undertake such commitments, namely expressions of intention, have quite a different sort of

significance from assertions. Attributions of them (the attitudes in terms of which score is kept on these deontic statuses) also behave differently. Yet the significances of these further sorts of performances, statuses, and attitudes can be understood straightforwardly by analogy to the sorts of scorekeeping that have been introduced for the pure assertional case. That extension of the model of the game of giving and asking for reasons provides a way of understanding intentional action and intentional interpretation of agents. It has been suggested that the doxastic commitments undertaken by speech acts having the significance of assertions can serve as analogs of belief—that such deontic statuses can do much of the explanatory theoretical work usually done by the paradigmatic sort of intentional state. That claim clearly cannot be assessed until the model is extended so as to include the possibility of manifesting such commitments not only in what is *said* but also in what is *done*—in action as well as assertion. Before turning to that wider context, however, it is worth pausing briefly to consider what can be said about the relation between belief as an intentional state and the deontic statuses with assertible contents as considered so far, already in the more restricted context of purely assertional practice.

It was indicated above that the assertional practices described so far generate two different senses in which one could be taken to be committed to a claim. Interlocutors undertake some commitments directly, by avowing them overtly: performing speech acts that have the significance of assertions. The commitments one is disposed to avow<sup>45</sup> are *acknowledged* commitments. But in virtue of their inferentially articulated conceptual contents, assertional commitments have consequences. Undertaking a commitment to a claim with one content involves undertaking commitments to claims whose contents are (in the context of one's other commitments) its committive-inferential consequences. Undertaking commitment to the claim that Pittsburgh is to the West of Philadelphia is one way of undertaking commitment to the claim that Philadelphia is to the East of Pittsburgh. These *consequential* commitments may not be acknowledged; we do not always acknowledge commitment to all the consequences of the commitments we do acknowledge. They are commitments nonetheless. For the only way that deontic *statuses* enter into the scorekeeping specification of assertional practices is as the objects of deontic *attitudes*. Indeed, *all* one can do with a commitment (or entitlement), in the model presented here, is take up a deontic attitude toward it—attribute it or undertake it, either directly by acknowledging it, or indirectly and consequentially. The scorekeeping model trades in talk about the status of *being* committed for talk about proprieties of practical attitudes of *taking* to be committed. Deontic statuses are just something to keep score with, as balls and strikes are just statuses that performances can be treated as having for scorekeeping purposes. To understand them, one must look at actual practices of keeping score, that is, at deontic attitudes and changes of attitude.

These ways in which one can come to be committed to a claim—by acknowledgment and consequentially—correspond to two ways in which we talk about *belief*. In one sense, one believes just what one takes oneself to believe, what one is prepared to avow or assert. In another sense, one believes, willy-nilly, the consequences of one's beliefs. Believing that Pittsburgh is to the West of Philadelphia is believing that Philadelphia is to the East of Pittsburgh, whether one knows it or not. This second sense looms particularly large for those who take their role in intentional explanations of behavior as the touchstone for identifying beliefs. For such explanations work only to the extent the individual in question is (taken to be) *rational*—not to have contradictory or incompatible beliefs and to believe the consequences of one's beliefs. From the point of view of the present project, this is because the conclusion of intentional explanations, strictly construed, is always a *normative* one—given intentional states whose contents are thus and so, one *ought* rationally or is rationally *obliged* or *committed* to act in such and such a way. Drawing conclusions about what actually will be or was done requires an additional premise, to the effect that the individual in question has mastered the practices of giving and asking for reasons sufficiently to be disposed to respond to acknowledgment of a (practical) commitment by producing a performance that satisfies it.

The sense of belief in which one is taken actually to believe what one ideally *ought* to believe (at least given what else one believes), call it *ideal* or *rational* belief, can conflict with the sense of belief for which avowal is authoritative. Dennett distinguishes these but thinks of them as competing norms to which a univocal sense of 'belief' must answer: "These two interdependent norms of belief, one favoring the truth and rationality of belief, the other favoring accuracy of avowal, normally complement each other, but on occasion can give rise to conflict."<sup>46</sup> The conflict arises precisely because one can avow incompatible beliefs, and fail to avow even obvious consequences of one's avowals:

What better source could here be of a system's beliefs than its avowals? Conflict arises, however, whenever a person falls short of perfect rationality and avows beliefs that are either strongly disconfirmed by the available empirical evidence or are self-contradictory or contradict other avowals he has made. If we lean on the myth that man is perfectly rational, we must find his avowals less than authoritative: "You *can't* mean—understand—what you are saying!"; if we lean on his 'right' as a speaking intentional system to have his word accepted, we grant him an irrational set of beliefs. Neither position provides a stable resting place; for, as we saw earlier, intentional explanation and prediction cannot be accommodated either to breakdown or to less than optimal design, so there is no coherent intentional description of such an *impasse*.<sup>47</sup>

The notion of incompatible beliefs offers no difficulties for a normative construal of intentional states as deontic statuses. There is nothing incoherent or unintelligible about the idea of undertaking incompatible commitments—*incompatible* or *inconsistent* beliefs just go into a box with incompatible or inconsistent promises. This is one of the benefits of this sort of approach over causal-functional accounts of intentional states. Yet the tension that Dennett identifies is a real one. The decision to treat belief just as what one is prepared to avow “amounts to the decision to lean on the accuracy-of-avowal norm at the expense of the rationality norm . . . If we demand perfect rationality, we have simply flown to the other norm at the expense of the norm of accuracy of avowal.”<sup>48</sup>

Dennett does not offer any way to reconcile the competing demands that the norm of rationality and the authority of avowals place on attributions of belief. The terminology employed here is animated in part by the thought that ‘belief’ may simply be *ambiguous* between a sense in which one believes just what one is prepared to avow and a sense in which one also believes what one *ought* rationally to believe, as a consequence of what one is prepared to avow (as already indicated, failures of rationality due to incompatibility cause no particular trouble once intentional states are construed as deontic statuses). An unambiguous, univocal technical term ‘doxastic commitment’ is introduced, which comprises both commitments one is prepared to avow and commitments that follow from those one acknowledges. But attention to the attitudes in terms of which those deontic statuses are explained makes it possible also to distinguish clearly between these two *kinds* of commitment, as ‘belief’-talk does not. The proposal is accordingly not to *analyze* belief in terms of commitment but to discard that concept as insufficiently precise and replace it with clearer talk about different sorts of commitment.

The fundamental concept of the metalanguage employed in specifying the model of assertional practice is that of the deontic attitude *attributing* a commitment. For the deontic attitude of *undertaking* a commitment is definable in terms of attribution: undertaking a commitment is doing something that licenses or entitles others to attribute it. Assertional performances or avowals are performances that express the deontic attitude of acknowledging doxastic commitments. They license attribution of (and insofar as they are successful, deferral with respect to) both the commitments they express and those whose contents are appropriate inferential consequences of the contents, commitment to which is overtly acknowledged. The attitude of acknowledging a commitment is in effect that of attributing it to oneself.<sup>49</sup> The fact that one thereby undertakes consequential commitments that may reach beyond what one acknowledges just shows that the generic attitude of *undertaking* a commitment is not to be identified with its species *attributing* a commitment to oneself, which is acknowledging it. The *social* dimension (invoking the perspective of *other* attributors) is essential to understanding

undertaking in terms of proprieties of attributing. The way in which the collaboration of attitudes adopted from two socially distinct perspectives—attributions of commitment to oneself and by others—is required to institute discursive commitments is the central theme of this work. It is in terms of the social-perspectival character of discursive deontic statuses that the notion of *objectivity* is to be made intelligible—both the general normative distinction between what one is *really* committed to do (or ought to do) and what one is merely *taken* by someone to be committed to do, and the more specific version that underwrites the notion of objective representational content, of a claim's correctness answering to how things are with what it represents, rather than to what anyone *takes* to be correct.

The roots of this social-perspectival account can already be discerned in the distinction that a scorekeeper can make between the commitments an interlocutor has undertaken and those that interlocutor acknowledges, and so is prepared to assert. For the attributions of the scorekeeper distinguish between the actual deontic *status* of the one for whom score is being kept, what that interlocutor is *really* (consequentially) committed to, and the deontic *attitudes* of that subject, what that interlocutor acknowledges commitment to by self-attribution. In other words, the notion of consequentially undertaking commitments provides the basis for distinguishing (in terms of the attitudes of someone keeping score) between deontic statuses and deontic attitudes. Indeed (as will appear in Chapter 8, where this issue is explored) the very notion of one propositional content being an inferential *consequence* of another essentially involves a crucial relativity to social perspective: are the auxiliary hypotheses (the premises to be conjoined with the claim in question in assessing its consequences) to be those the scorekeeper assessing the propriety of the inference *undertakes* commitment to, or those the scorekeeper *attributes* to the one whose statuses are being assessed?

Neither answer is correct. The fact that proprieties of inference a claim is involved in can be assessed from either of two social perspectives—that of the one *attributing* commitment to the claim or that of the one *undertaking* that commitment—is fundamental to the very notion of a *propriety* of inference. And since propositional and so conceptual contents of all sorts are constituted by the broadly inferential proprieties of practice in which they are caught up, such contents are essentially social and perspectival in nature. The propositional content of a claim or commitment can be specified only from some point of view; that it would be differently specified in definite ways from other particular possible social perspectives (that is, scorekeepers occupying such perspectives) is an essential part of its being the content it is.

At this point the phrase “social-perspectival character of the contents of discursive commitments” can be little more than a label attached to a promissory note—though the discussion of scorekeeping in this section is intended to give it enough resonances to make it at least a suggestive label. Before that promissory note is redeemed, in Chapter 8, it is necessary to look

much more closely at the sorts of contents that can be conferred on expressions playing various roles in assertional practices of the sort described here. In the intervening chapters the notion of *inferential* articulation is deepened and extended by adding *substitutional* machinery. The notion of substitution inferences permits the extension of the notion of conceptual content introduced here to essentially *subsentential* expressions such as singular terms and predicates, which can play only indirectly inferential roles—not serving themselves as premises and conclusions of inferences but only occurring in the sentences that can serve in those capacities. The conceptual content expressed by the use of singular terms and predicates is articulated by substitution-inferential commitments (in Chapter 6). Extending this account to the sort of content expressed by the token-reflexive or indexical use of unrepeatable expressions, paradigmatically the sort of *deictic* tokenings that play such an important role in empirical knowledge claims, requires looking still further at *anaphoric* connections among tokenings. (The importance of anaphoric relations for understanding what is expressed by traditional semantic vocabulary, paradigmatically ‘true’ and ‘refers’, is argued in Chapter 5.) Anaphora is explained as a certain structure of *inheritance* of substitution-inferential commitments (in Chapter 7). The result is a layered account of the semantic contents that can be conferred by assertional practice as here described—an account whose key concepts are those of *inference*, *substitution*, and *anaphora*, each permitting a finer-grained analysis of the structures that precede it and which it presupposes.

# Perception and Action: The Conferral of Empirical and Practical Conceptual Content

---

Nature to be commanded must be obeyed; and that which in contemplation is as the cause is in operation as the rule.

BACON, Novum Organum

In the beginning was the deed.

GOETHE, Faust

The true being of man is his deed; in this the individual is actual . . . What the deed is can be said of it. It is this, and its being is not merely a sign, but the fact itself. It is this, and the individual human being is what the deed is. Action simply translates an initially implicit being into a being that is made explicit.

HEGEL, Phenomenology

---

## I. ASSERTIONS AS KNOWLEDGE CLAIMS

### 1. *Five Strategic Explanatory Commitments*

Assertions are the sort of *claims* made in the standard case by uttering freestanding declarative sentences—that is, sentences whose occurrence is not embedded in the occurrence of a compound sentence. A commitment has been undertaken here to an order of explanation dictating that this principle be exploited by defining declarative sentences in terms of an account of assertion, which evidently then is required to be made available independently. This contrasts with the procedure common in formal semantics, in which the theorist leaves until later the task of getting a grip on the activity, force, or significance of assertion but provides an antecedently defined construal of sentences. The strategic commitment to treating what is expressed by the use of *sentences* (rather than what is expressed by the use of singular terms or predicates) as the fundamental sort of semantic content is an element of the present account that has been taken over from Kant.

The pragmatist strategic commitment to understanding semantics in terms of pragmatics (the *contents* associated with expressions in terms of the practices governing their *use*) is an element of the present account that has been taken over from Wittgenstein. The strategic commitment to specifying such a pragmatics in the first instance in *normative* terms is an element of

the present account that has been taken over from Kant, Frege, Wittgenstein, and Sellars. The inferentialist strategic commitment to treating the public linguistic practice of *asserting* as the fundamental activity involving such contents, rather than the private mental practice of judgment, is an element of the present account that has been taken over from Dummett. The strategic commitment to understanding asserting a sentence as a significance a performance acquires in virtue of its role in a practice of giving and asking for *reasons*, of justifying and communicating justifications, is an element of the present account that has been taken over from Sellars.

This constellation of commitments combines the *normative* articulation of the pragmatic significances of assertional performances with the *inferential* articulation of the propositional contents they express, in part by putting the issue of whether an asserter is *entitled* to the *commitment* undertaken by making an assertion at the center of the practice that institutes those significances and confers those contents. To do so is to treat the sort of claim involved in asserting as an implicit *knowledge* claim. From the point of view of the concerns that motivate the present project, this is as it should be. For the aim all along has been to elaborate a criterion of demarcation that sets us off by our peculiar susceptibility to *reasons*. It is this susceptibility that makes it appropriate to think of ourselves in terms of the categories of *knowledge* and *action*. That is why the story really begins with Kant's observation that knowings and actions are to be distinguished from other things we do by the characteristic way in which we are *responsible* for them. The notion of *discursive commitment* arises in the domain of social practice when one focuses specifically on the norms that are articulated in the form of reasons.

Absent the inferential dimension, the norms implicit in a set of social practices could be understood neither as conferring *propositional* contents nor as instituting *assertional* significances, hence not as governing genuinely *linguistic* practice. Inferential connections enter into the alterations of attitude (the scorekeeping that defines assertional practice) in three fundamental ways: one corresponding to each of the basic sorts of deontic status, and a third involving the relation between them. First, part of the significance of acknowledging an assertional commitment is that one thereby undertakes commitment as well to all those contents it entails—that is, to consequences that follow from it by commitment-preserving inferences. One who claims that a lion roared is committed thereby to a mammal's having roared. Second, part of the significance of undertaking an assertional commitment is that one thereby undertakes a conditional task-responsibility to demonstrate one's *entitlement* to that commitment, if faced with a warranted challenge. Here justificatory or entitlement-preserving inferences involving the asserted content help determine what deontic statuses are attributed to which asserters, challengers, and deferrers. Finally, part of the deontic scorekeeping practice within which performances can have the significance characteristic of claim-

ings is to withhold attribution of *entitlement* to commitments *incompatible* with a *commitment* that has been undertaken (whether by overt assertion or consequentially). This is the practice that defines incompatibility relations on the contents of deontic states; two claims are incompatible if commitment to one precludes entitlement to the other. Connections of these three sorts consist in proprieties that govern the alterations of deontic attitudes by which interlocutors keep discursive score. Deontic statuses then count as having inferentially articulated contents because of the pragmatic scorekeeping significance of the performances that express the acquisition of those statuses.

## 2. Knowledge as a Complex Hybrid Deontic Status

That assertions have the default status and significance of implicit knowledge claims is to be understood in terms of these inferentially structured interactions between the two *modally* distinct deontic statuses, *commitment* and *entitlement*, and the two *socially* distinct deontic attitudes, *attributing* and *undertaking* deontic statuses. The status one attributes in attributing knowledge is traditionally understood according to the tripartite structure: justified true belief (JTB).<sup>1</sup> One of the leading ideas of the present approach is that the notion of normative status can be made to do much of the theoretical and explanatory work that the notion of intentional state has heretofore been called on to do. In the social practice model, talk of *belief* is replaced by talk of assertional or doxastic *commitment*. According to the JTB approach, attributing knowledge is attributing a special kind of belief. So attributions of knowledge are to be rendered here in terms of the deontic attitude of *attributing commitments*—specifically commitments of the sort that can be undertaken or acknowledged by performing a speech act that has the significance characteristic of assertions.

As Plato had already pointed out, there is more to attributing knowledge than attributing belief. There is also the issue spoken to by the demand for justification, or as Plato has it, for an account. According to the canonical tripartite understanding, knowledge is not just belief but *justified* belief. Clearly what corresponds to this condition in the deontic version is the demand that the one taken to be a knower not only have a commitment but be *entitled* to that commitment. Making an assertion, it has been said, is making a knowledge claim. Assertional performances as modeled here have the significance not only of undertaking commitments but of defeasible claims to entitlement to those commitments. So one is not attributing knowledge to someone unless one not only attributes a commitment but also attributes a corresponding entitlement. Most classical epistemological problems are really problems concerning this deontic status—justification for believing, or more generally entitlement to believe.

Before considering entitlement to believe, however, a few words are in

order about the third limb of the tripartite rendering of what one is doing in taking someone to have the status of a knower. In attributing knowledge one is not just attributing justified belief—a commitment of the sort that can be undertaken by asserting it and an entitlement to that commitment of the sort that can be demonstrated by justifying it. One must also take the belief to be *true*. What is the social-deontic attitude corresponding to the truth condition on attributions of knowledge?

The attitude of taking-true is just that of *acknowledging* an assertional commitment (the attitude that grounds consequential undertakings of such commitments). A theory of asserting and assertional commitment is a theory of taking-true. Evidently this principle can be exploited according to two different orders of explanation: moving from a prior notion of truth to an understanding of asserting (or judging) as taking, treating, or putting forward *as true*, or moving from a notion of asserting to a notion of truth as what one is taking, treating, or putting forward a claim *as*.<sup>2</sup> The latter line of thought accords ‘true’ an expressive role, in permitting us to *say* something about assertion, rather than an explanatory role, as something that can be understood in advance of understanding assertion and used to advance to such an understanding. This approach is the one pursued in Chapter 5.

In taking someone to be a knower, one *attributes a commitment*, *attributes entitlement* to that commitment, and *acknowledges commitment* to the same content oneself. Undertaking the commitment is part of what the asserter authorizes others to do—not only to attribute the commitment but also to undertake it, on the asserter’s authority. That authority depends on the asserter’s entitlement to the commitment, so the asserter is implicitly claiming that entitlement as well. That is why assertions in the basic practices described here have the significance of claims to knowledge. For others to take those claims to be successful is for them to attribute the commitment undertaken, in addition to attribute entitlement to it, and finally to endorse the claim themselves. These correspond, in the model of linguistic practices in terms of scorekeeping with deontic attitudes, to taking to believe, taking to be justified in that belief, and taking the belief to be true.<sup>3</sup>

According to this way of understanding knowledge claims, the significance of the truth condition on attributions of knowledge lies in the fundamental difference in social perspective between *attributing* a commitment (or other deontic status) to another and *acknowledging* it oneself. Knowledge is a *complex deontic status*, in the sense that it involves both commitment and entitlement. But attributions of knowledge (and so claims of knowledge) are also *hybrid deontic attitudes*. So knowledge can be called a hybrid deontic status. Attributions of knowledge (the attitudes in terms of which that status is to be understood) are hybrid deontic attitudes in the sense that they involve both attributing and acknowledging commitments. These attitudes are perfectly intelligible in the context of the model presented here of the social practices that institute assertional force.

But it is also clear in those terms that there is a great danger of misinterpreting what one is doing in calling a claim true. The danger is in misunderstanding taking-true solely in terms of *attributions* of status (rather than as essentially involving also the *undertaking* of one), by assimilating that attitude too closely to the attributions of commitment and entitlement involved in the other dimensions involved in treating a claim as a bit of knowledge. If the hybrid nature of that attitude is overlooked, it will be thought of as consisting only in attributions; attributions of knowledge will be taken as comprising attributions of belief, attributions of justification, and attributions of truth. One then looks for the property or status one is attributing to a claim in taking it as true.<sup>4</sup> The property or status projected by misconstruing undertaking a commitment as attributing some property or status to it is bound to be "queer."

In the deontic scorekeeping model of inferentially articulated linguistic social practices, asserting is making a knowledge claim. The attitudes in terms of which the hybrid deontic status of knowledge is understood are just those in terms of which the significance of assertions is specified. Assertional practice is accordingly a version not only of the game of giving and asking for reasons but also of the game of making and defending claims to knowledge. Practitioners who can produce and consume assertions are *linguistic* beings. Practitioners who can produce and consume reasons are *rational* beings. Practitioners who can produce and consume knowledge claims are *cognitive* beings. On the account presented here, these are three ways of talking about the same practices and the same capacities.<sup>5</sup>

Underlying all of them is the inferentially and socially articulated notion of *discursive commitment*. It is the topic in which philosophy of language, philosophy of mind,<sup>6</sup> and epistemology are alike rooted. What epistemology studies is a deontic status that is implicitly in play in any practices involving propositional contents—whether or not those practices include the expressive resources provided by words like 'knowledge', which can be used to make attitudes toward that status explicit. For making and defending what are implicitly claims to knowledge is an essential feature of discursive practice as such.

On this account, prizing and searching for knowledge are not specialized intellectual virtues, appropriate only to a sophisticated, culturally late-coming elite. They are built into what we fundamentally are. The complex hybrid deontic status of knowledge defines the *success* of assertion. Treating an assertion as expressing knowledge—attributing to the asserter entitlement to the commitment undertaken thereby and endorsing that commitment oneself—is the response that constitutes the practical recognition of the authority implicitly claimed by the assertion. For that is the authority to license undertakings of commitment to that same claim by those in the audience, in virtue of the asserter's entitlement to the commitment. For a scorekeeper fully to accept the authority implicitly claimed in the making of an assertion

is just for that scorekeeper to treat it as having the status of knowledge. So the aspiration not only to *truth* but to *knowledge* is built right into the normative structure of assertional practice. (And it should be noticed that the sense in which the status of knowledge provides the *point* of assertion can be specified in advance of any consideration of the *intentions* of the practitioners.) Knowledge is on this account an ideal projected by the very possibility of saying anything at all.

### 3. Justifying and Being Entitled

In this context it is useful to look a bit more closely at the structure of attributions of entitlement to assertional commitments, from a more traditionally epistemological perspective. According to the tripartite analysis, to take a claim to express knowledge is to take it to express a justified true belief. The justification condition on knowledge will be misunderstood if one does not distinguish between two senses in which a belief can be said to be justified. In one sense, to call a belief justified is to invoke its relation to the process of *justifying* it. To *be justified* in this sense is to *have been justified*—exhibited as the conclusion of an inference of a certain kind. In another sense, to call a belief justified is to attribute to it what might be called *positive justificatory status*. Positive justificatory status is just what has been talked about here in terms of *entitlement* to a claim.

The relation between possession of such status and the activity of justifying may be quite indirect. In particular, justifying a claim is only one way in which it can acquire positive justificatory status. Indeed, as has already been pointed out, to avoid embarking on a foundationalist regress it is necessary to acknowledge that a commitment may have a positive justificatory status without having been justified (indeed, without that entitlement having been defended in any way, whether *intrapersonally* by inference or *interpersonally* by deference). Since any activity of justifying—even if that term is understood broadly, as entitling (so as to include deferring as well as inferring)—is a mechanism making possible the *inheritance* of entitlements, there must be some at least prima facie entitlements available to get the process off the ground. If dogmatism is to be avoided, such entitlements must not be immune to criticism; there must be mechanisms for bringing them into question. The combination of prima facie entitlements and ways of criticizing and undermining them is what was called the *structure of default and challenge*. It characterizes a dynamic process of acquisition and loss of entitlements by various commitments on the part of various interlocutors (kept track of in the attitudes of claiming and attributing entitlements) and of withholding such claims and attributions.

Classical foundationalism considers only justifying in the narrow sense of an inferential activity, not in the broader sense of vindication that includes the communicational dimension appealed to by deferential entitling (the

authority of testimony). This is unfortunate, for if the analysis just offered of what one is doing in calling something knowledge is correct, the hybrid deontic status of knowledge is incomprehensible in abstraction from the social distinction of perspective distinguishing the deontic attitudes of attributing and undertaking commitments. One of the centerpieces of the present account is its attention to the interaction of the two dimensions of the practice of giving and asking for reasons for commitments to inferentially articulated contents: the intracontent, interpersonal communicational dimension and the intercontent, intrapersonal justificatory dimension. Both the individuation of the contents individuals are responsible for and the individuation of the individuals responsible for them are to be understood in terms of this structure. Equally important, as Chapter 8 shows, an inferential understanding of the *representational* dimension of conceptual content depends upon an appreciation of the *social* articulation of inferential practice. None of this is accessible from the point of view of the one-dimensional approach that ignores the significance of communication for justification. Even within that narrower compass afforded by exclusive attention to intrapersonal, intercontent entitling, however, the consequence of insisting that positive justificatory status can be the result only of justifying is a dual regress—one regress on the side of entitlement to premises, and another on the side of entitlement to inferences.

To illuminate the default and challenge response to the threat of a foundationalist regress on the side of premises, consider its twin on the side of inferences. If entitlement to a commitment to  $q$  is at issue, and that commitment is justified by asserting  $p$ , the vindication might be unsuccessful either because the commitment to  $p$  is not one the interlocutor is entitled to or because the inference from  $p$  to  $q$  is not correct (in this case, not entitlement-preserving). A regress on the side of the inferences results if one insists that each inference is, to begin with, in need of support or justification. Endorsing the propriety of an inference is brought into the game of giving and asking for reasons in a new way by making the inference explicit in the form of a conditional, which can be endorsed, challenged, and defended like any other assertible. The demand is then for an explicit rule or principle to warrant the propriety of every inferential transition appealed to in justifying a claim.

But pragmatism maintains that to demand this is to view things the wrong way around. One must start with a notion of taking or treating inferences as correct in practice. Without such a practice, there is no game of giving and asking for reasons to bring inferences into the form of explicit assertions. Once the game is under way, the practical inferential attitudes it involves can then, on suitable occasions, be made explicit in the form of endorsements of conditionals. But what those conditionals express is intelligible only in terms of the underlying inferential practice.

If it is insisted instead that no move be treated as entitled or entitlement-

preserving until its entitlement has been demonstrated or justified, a new premise is introduced corresponding to every inference, and also a new inference employing that premise. Then the regress ramifies, as the entitlements to those new premises and those new inferences must themselves be secured. Within the resulting regress can be discerned Kant's and Wittgenstein's regress of rules—and where in addition the goodness of inference is identified with *formal* goodness of inference, Lewis Carroll's regress of conditionals and detachment from conditionals as well. Carroll's multiplication of premises standing behind inferences should be halted by acknowledging primitive *rules* of inference; the multiplication of conditionals explicating implicit "enthymematically suppressed" premises should be halted by acknowledging primitive *material* rules of inference; and the multiplication of rules should be halted by acknowledging primitive material-inferential *practices*.

These moves have all already been considered. The default-and-challenge structure of assertional entitlement just amounts to extending to the case of assertions the policy that underwrites these ways of thinking about inferences. What is fundamental in each case is the practical attitude of taking or treating as correct moves in the game of giving and asking for reasons. Though such entitlements can be brought into question later, one initially is entitled to whatever one is in practice taken or treated as entitled to; deontic statuses must be understood in terms of practical deontic attitudes. It makes sense that this way of construing the proprieties of inference that articulate the propositional content of assertional speech acts and the commitments and entitlements they involve should extend as well to the proprieties that govern those assertional performances and deontic statuses. For asserting and inferring are two sides of one coin; neither activity is intelligible except in relation to the other. Undertaking an assertional commitment involves a commitment to the propriety of inferences from the circumstances of application to the consequences of application of the concepts in terms of which its content is articulated. If claiming is to be possible at all, some of those content-constitutive implicit inferential proprieties must in practice be taken for granted, treated as *prima facie* in order—not as innocent until *proven guilty*, but at least as innocent until indicted on the basis of reasonable suspicion. In the same way, sometimes a defeasible presumption that the application of those concepts in an assertion or judgment is appropriate must be in order.

## II. RELIABILITY

### 1. *Reliabilism and Entitlement*

When the justification-as-entitlement of a belief is decoupled to this extent from the activity of inferentially (or, for that matter, deferentially) justifying it, the question arises whether the latter notion need be taken to

play any role whatever in the understanding of the status of being a justified belief that is appealed to by the tripartite analysis of knowledge. It is generally agreed that *some* sort of entitlement to a claim is required for it to be a candidate for expressing knowledge. But it is not obvious that inferring in the sense of justifying is at all fundamental to that sort of entitlement. When examples of the sort that motivate the tripartite analysis in the first place are examined more closely, it appears that what is forbidden is that it be merely *accidental* that one has a true belief. Someone who flips a coin to decide which is the correct road to Athens may by accident pick the right one, and may somehow or other come to believe that the one chosen is the correct road. But such a true belief does not qualify as knowledge. One way of showing that the belief is not merely accidental is indeed to provide an account, to offer reasons for the belief.

It has been suggested, however, that this is merely *one* way, and by no means the most basic, in which the belief could be shown to have credentials beyond those provided by happenstance and coincidence. In particular, the correctness of the belief is not merely fortuitous if it is the outcome of a generally reliable belief-forming mechanism. Epistemological *reliabilists* claim that this is the sort of entitlement status that must be attributed (besides the status of being a true belief) for attributions of knowledge.<sup>7</sup> The perceptual mechanisms underlying entitlement to empirical claims provide the most important and persuasive examples. This line of thought is sometimes extended to an analysis of justification as consisting simply in the demonstration of the reliability of a belief-forming mechanism. The version that is of interest here, however, claims only that reliability can do *all* the work that inferential justifying is taken to do in the standard tripartite analysis (and, by extension, in the assertional practices of making and defending knowledge claims described here). The further Procrustean reductive assimilation of justifying to the paradigm of indicating reliability can be put to one side at this point, for it is the difficulties that arise already with the weaker thesis that are most instructive.

*Reliability* accounts of entitlement to assertional commitments and *regularity* accounts of the correctness of such commitments are species of one genus. They share a common strategy for naturalizing the different norms they address. In tracing the relation between them, it is helpful to keep in mind the basic case in which making an assertion consists in noninferentially applying a ground-level empirical concept in a particular situation. Regularity theories are attempts to determine the boundaries of concepts—which determine the difference between correct and incorrect application—by appealing to regularities or patterns in the actual applications of the concept and the dispositions to apply that concept that are exhibited by an individual or a community. The sort of *correctness* of application of concepts that such theories aim to explicate is what is assessed by judgments of the *truth* of the resulting assertion.<sup>8</sup> What it is for *A* to take *B*'s claim that

something is a porcupine to be a correct application of that concept is for *A* to take the claim to be true, that is, for *A* to endorse it, for *A* to undertake or acknowledge commitment to that same content. This is a different deontic attitude from *A*'s attributing to *B* *entitlement* to the commitment undertaken by that claim, and so to that application of the concept; different sorts of normative status are involved.

The concept of *reliability* in making a claim or applying a concept presupposes, rather than analyzes, such a notion of *correct* claiming or concept-application. For a reliable performer is just one who generally produces a correct performance; assessments of reliability are assessments of the probability of correctness. Thus the issue of reliability cannot be raised until the question of correctness that regularity theories address has been answered. Nonetheless, these different theories share an approach. For reliability theorists offer an account of entitlement that appeals only to patterns or regularities of *correct* claiming or application of concepts. The reliabilist idea is that *entitlement* to a particular claim or application of a concept—a derivative sort of correctness of claiming—can be understood entirely in terms of *dispositions* to produce *correct* performances of that kind. The regularist idea is that such correctness of claims or applications of concepts can be understood entirely in terms of *dispositions* (in some variants, those that are in some sense communal dispositions) to produce performances of that kind.

One of the major difficulties raised for the strategy of construing the correctness of discursive performances in terms of regularities or dispositions specified in nonnormative terms was the gerrymandering problem. There is no single pattern or regularity exhibited by any set of actual or virtual performances; where there is one, there are many—indeed an infinite number. No matter what a candidate performance whose correctness is at issue is like, and no matter what the history to which it must answer is like, there is *some* way of specifying the pattern exhibited by those prior performances so as to include the candidate as just what is required to continue that pattern “in the same way.” The attempt to identify the normative distinction between correct and incorrect performances with the naturalistic distinction between regular and irregular performances fails because no performance is simply irregular (even relative to a specified set of performances with respect to which its coregularity is to be assessed), and so none would be counted as incorrect by such a criterion. Appeal to regularity and irregularity can do normative explanatory work only if there is some way of *privileging* some regularities over others—some way, in other words, of saying what the *correct* regularity is. The problem of sorting performances into correct and incorrect is transformed by the regularist strategy into the problem of sorting regularities into the relevant and the irrelevant, the ones that *ought* to be taken account of in assessments of correctness, and those that ought not. From this vantage point, regularity theories appear as merely putting off the normative issue, moving the bump in the carpet around rather than smoothing it out.

## 2. Barn Facades

Reliability theories share with regularity theories the same fundamental strategy for explaining in naturalistic terms the normative statuses involved in discursive practice (although the phenomena they address are at different levels). So it might be expected that the possibility of gerrymandering would raise similar difficulties for a pure reliability strategy for constraining entitlements to claims that it does for a pure regularity strategy for construing the correctness of those claims. This is indeed the case. A striking illustration of how the gerrymandering considerations get a grip on assessments of entitlement in terms of reliability is provided by Goldman's barn-facade example.<sup>9</sup> The example is forwarded to show the inadequacy of an account that seeks to ground the cognitive authority of noninferential reports exclusively in features of the causal chain leading from the reported state of affairs to the perceptual reporting of it. The leading idea of such causal theories is that a true belief, paradigmatically one acquired perceptually, counts as knowledge just in case it is caused in the right way by what it is about. To see that such a particular causal connection is not sufficient to make a true belief qualify as knowledge (and so cannot by itself perform the explanatory job assigned to the entitlement condition by the tripartite analysis), Goldman suggests comparing two cases that are alike as far as the causal chain leading to a claim is concerned, but unlike in the status of that claim as knowledge.

In each case the subject is in ideal circumstances for visual perception and is confronted by what is in fact a barn. In each the subject responds to the visible presence of the barn by confidently reporting the presence of the barn. The causal chains in each case are entirely standard, the barns reflecting light, which travels undisturbed to the subject's retina, and so on. Yet one of the subjects is, and the other is not, without knowing it located in Barn-Facade County. The local hobby in that county is building incredibly lifelike trompe l'oeil facades of barns. In this county, 99 percent of what appear to be barns are actually such facades. Each subject would in fact, if confronted with such a facade (and not alerted to the special practices of the natives), confidently report the presence of a barn. Goldman's plausible claim is that the claim of the subject who is *not* in Barn-Facade County *does* express knowledge (is a claim that subject is in the relevant sense entitled to), while the claim of the subject who *is* in Barn-Facade County *does not* express knowledge.

For the first point: the fact that there are *some* circumstances under which a subject could be fooled does not in general preclude the subject from having knowledge in the case in which that subject is not fooled. As Austin argued, the fact that it is possible to make a replica of a sparrow so cunningly contrived that I cannot tell it from the real thing does not mean that I cannot see a sparrow and know that it is a sparrow.<sup>10</sup> The mere possibility of hyperbolic doubt does not entitle anyone to it and does not undermine

entitlements in ordinary cases. For the second point: even though the subject who has the cognitive bad luck to be in Barn-Facade County is in fact looking at a barn, it is in an important sense just an *accident* that that is the case. The county is rife with perceptual situations in which the subject would with equal confidence and warrant *false*ly report the presence of a barn. Under these circumstances, the belief just *happens* to be true, and the subject should not be taken to *know* that there is a barn present.

Goldman claims first that this sort of example shows that one must look beyond the particular causal antecedents of a belief in order to determine its status as one the believer is entitled to in the sense relevant to assessments of knowledge. For in this case what distinguishes the two subjects is not the causal chains connecting them to the barns but only the incidence of barn facades in the vicinity, which is causally irrelevant to their perceptual transactions with the barns they are in fact looking at. His second claim is that the way in which that difference of causally irrelevant circumstance makes a difference to the assessment of entitlement and hence of knowledge can be understood in terms of the variable *reliability*, in those different circumstances, of the belief-forming mechanism that leads to the true belief in each case. The same differential responsive dispositions, the same noninferential reporting capacity, is in play in both cases.

The difference is that in Barn-Facade County it is not a reliable mechanism, while in the rest of the (largely barn-facadeless) world it is. How reliable a belief-forming mechanism is, how likely it is to yield a true claim, a correct application of a concept, depends on the circumstances in which it is exercised. My inability to tell sparrows from cunning duplicates does not disqualify me from being a reliable reporter of sparrows, so long as my environment is quite unlikely to confront me with a ringer. If such duplicates were to become common, the reliability of my differential responsive dispositions would degrade (and with it my capacity to acquire knowledge thereby in the cases where all goes well), even though the way in which that mechanism would respond to each possible case remained the same throughout. The probability of being correct in a particular case depends on the actual incidence of indistinguishable phonies. Thus the notion of reliability of belief-forming mechanisms provides just what is wanted to explain the barn-facade cases.

### 3. *Gerrymandering and the Problem of Reference Classes*

Goldman's argument is decisive against exclusively causal theories of knowledge, and it shows how assessments of reliability can function in assessments of entitlement—particularly entitlement to commitments acquired as a result of noninferential reporting capacities. But (though he does not make the point) it also underscores the possibility of gerrymandering, and hence the inadequacy of construing cognitive entitlement exclu-

sively in terms of reliability. In the case of regularity theories of the correctness of the application of a concept, it is the boundaries of the concepts that can be gerrymandered in such a way as to preclude assessments of irregularity, and hence of error. In the case of reliability theories of entitlement, it is rather the boundaries of the reference class with respect to which reliability is assessed that can be gerrymandered in such a way as to preclude assessments of unreliability, and hence of lack of entitlement.

Goldman's idea is that reliability is an objective affair, determined by the objective probability of a correct judgment, given one's circumstances. But such probabilities vary with the specification of those circumstances. Given a reference class of relevantly similar cases, frequencies of success define objective probabilities. The question remains how a privileged reference class is to be determined. What is the *correct* reference class with respect to which to assess such probabilities?

If the reference class is restricted to the actual case of the perceptual judgment that a barn is present, even in Barn-Facade County (since in the case being considered by hypothesis one is actually looking at a barn) the frequency of correct judgments is 1. So relative to that quite restricted reference class one is totally reliable. If the reference class is widened to the whole county, the frequency of correct judgments is reduced to 1 percent. So, relative to that less restricted reference class, one is quite unreliable. But since the customs of Barn-Facade County are quite parochial, the relative frequency of barn facades in the country as a whole is quite low. Relative to the nation as a whole, one is quite a reliable noninferential reporter of the presence of barns. Relative to the state, one's reliability will fall somewhere in between. One of the nice things about this example is that here the metaphor of *boundaries* is made concrete, and the difficulty of selecting the proper boundary is literally geographic.

Focusing on the relativity of reliability to decisions about where to draw these boundaries makes it evident that the question "Reliable or not?" is underdetermined in exactly the same way that the question "Regular or not?" is underdetermined. There are always *some* regularities that are being instantiated, and (in the case where the claim one is making is true) there are always *some* reference classes with respect to which one is reliable. Using these naturalistic notions to stand in for genuinely normative assessments works only relative to some way of privileging regularities or reference classes. The notions of regularity and reliability cannot do all the work by themselves. For the concept of regularity cannot discriminate between regularities, and that of reliability or probability of success relative to a reference class cannot discriminate between reference classes. In the sense in which, given a regularity, there is an objective matter of fact as to whether a further performance continues it, there is no objective matter of fact as to which of the various regularities exhibited by a given history of actual or dispositional performances is the right one to assess correctness with respect to. In the

sense in which, given a reference class of relevantly similar cases, there is an objective matter of fact as to what the probability that a certain skill exercised in those circumstances will yield a correct performance, there is no objective matter of fact as to which of the various possible reference classes to which the case in question might be assimilated is the right one to assess reliability with respect to. An objective or naturalistic theory of cognitive entitlement cannot be derived solely from considerations of reliability, any more than an objective or naturalistic theory of the correct application of concepts can be derived solely from considerations of regularity.

#### 4. *Taking or Treating as Reliable*

The general strategy of this work is to supply what is wanting in regularity theories of correct concept-application by appealing to the social (in an *I-thou* sense) and practical deontic attitudes of taking or treating a performance as correct or incorrect. The paradigm is taking or treating an assertion as correct in the sense of endorsing it, undertaking that commitment oneself, which is taking what it says to be true. It is these attitudes on the part of interpreters, of the deontic scorekeepers who attribute discursive commitments, that privilege some regularities over others and give a sense to the notion of *correct* use of expressions and so applications of concepts. The deontic status of *being* a correct application of a concept is to be understood in terms of the deontic *attitude* of *taking* or *treating* such an application *as correct*. That attitude (endorsing a claim, undertaking an assertional commitment) cannot be understood apart from its role in the essentially social practice of giving and asking for reasons, making and defending knowledge claims. The norms implicit in the application of concepts are social and perspectival, not (to begin with) objective and naturalistic.<sup>11</sup>

Regularity theories attempt to naturalize the normative status of correct claiming or concept-application. The countervailing idea pursued here is to explain that status by saying what it is for a performance to be taken or treated in practice as having such a significance. This is to focus on the deontic attitudes of acknowledging conceptual norms by attributing normative statuses and significances. Reliability theories attempt to naturalize the normative status of *entitlement* to the commitment undertaken by making a claim or applying a concept. The corresponding countervailing idea to be pursued here is accordingly to explain that status by saying what it is for a performance to be taken or treated in practice as having such a significance. This is to focus on the deontic attitudes that acknowledge that status and attribute that significance.

What in practice privileges some of the reference classes with respect to which reliability may be assessed over other such reference classes is the attitudes of those who attribute the commitment whose entitlement is in question. Each interpreter implicitly distinguishes between reference classes that are relevant and those that are irrelevant to the assessment of reliability,

and hence of entitlement to claims, by the circumstances under which that interpreter accords cognitive authority to those claims. The sort of authority in question here is not that acknowledged by the interpreter's own endorsement of the claim—that is, taking it to be correct in the sense of taking it to be *true* (which is the sense of correctness addressed by regularity theories rather than by reliability theories). The sort of authority in question is rather that of having an inheritable entitlement: the sort that supports successful deferrals by others (potentially including the interpreter). It is the scorekeeping social practices that actually govern the use of an expression (in particular the acknowledgment of entitlement to the commitments undertaken by its assertional use) that supply what is missing from pure reliability theories.

It is tempting, from the point of view of such theories, to think of the choice of reference class as a merely pragmatic matter—in a sense of ‘pragmatic’ that restricts it to what concerns the interests and goals of those performing speech acts. So it might be thought that for some purposes and in some contexts I should be counted as knowing that a sparrow is in front of me, even though I would believe that also if a sufficiently lifelike replica were there instead, while for other purposes and in other contexts (for instance where the stricter standards appropriate to discussions of principled skepticism are in force) I should not. No doubt there is such a variation in standards of entitlement depending on what is taken to turn on the issue, and it may be particularly acute in connection with the word ‘know’. But the contribution made by interpreters (those who attribute commitments and entitlements to commitment) to the determination of the boundaries with respect to which reliability is assessed are not “merely pragmatic” in this sense.

They make a fundamental contribution to the semantic content of empirical concepts. Indeed, this is one of the situations in which traditional ways of distinguishing semantic from pragmatic concerns can be seen to be inapposite. In particular, as will appear, what an interpreter takes to be the circumstances under which an expression can appropriately be used in noninferential reports—that is, when interlocutors are *entitled* to commitments because the acknowledgment of the commitment arises through the exercise of dispositions to respond differentially to various aspects of their environment—is an important feature of the empirical content the interpreter associates with that expression. The sort of authority accorded to noninferential reports, and the way the model of assertional practice can be extended to incorporate it, is discussed further below.

### III. OBSERVATION REPORTS AND NONINFERENTIAL AUTHORITY

#### 1. *Knowledge, Entitlement, and Understanding*

The topic of reliability theories of cognitive entitlement was introduced in connection with the thought that once the notion of entitlement

or positive justificatory status that matters for attributions of knowledge has been broadened by the recognition that a belief, claim, or commitment can in this sense *be justified* without having *been justified*—that justifying is not the only way that status can be acquired—the way seems open to dispensing entirely with inferential justifying in explaining the deontic status of entitlement. It was then pointed out that identifying the entitlement of a commitment with its being the output of a reliable process has the same sorts of difficulties with gerrymandering that plague its relatives that identify the correctness of a claim or application of a concept with its being in accord with a regularity exhibited by other claimings or applications of the concept. But these difficulties concern only one way of following out the original thought about the in-principle dispensability of inferential justification in explaining the status of knowledge claims. The deep mistake involved in completely decoupling justifying as giving reasons from cognitive entitlement has to do rather with the sort of *understanding* that is presupposed by claims to and attributions of knowledge.

An assertion, even if true, is not taken to express knowledge unless the one making it *understands* the claim being made. A practical grasp of the significance of making the claim is inseparable from an appreciation of its role as possible reason for other claims, and as something that reasons can in turn be offered for. It is being caught up in this way in the game of giving and asking for reasons that makes a performance the undertaking of a commitment (the making of a claim) in the first place. Unless one accords one's own performance such a significance (treats it as a move in that game), one is not making a claim, not undertaking a commitment that is eligible for the status of knowledge.

It is on this basis that Sellars objects to construals of cognitive entitlement exclusively in terms of reliability. Reliable differential responsive dispositions are only a necessary condition for observational knowledge. Parrots and thermometers can have such dispositions and so can be used by us in effect as measuring instruments to acquire knowledge. But what *they* have is not knowledge. For they do not understand the significance of their responses; they do not take those responses as reasons for further claims; and they do not understand claims as potentially in need of reasons. To decouple entitlement from reason-giving entirely is to jettison the inferential articulation in virtue of which the performances and commitments one is entitled to can be understood as *propositionally* contentful. It is to discard precisely what makes responses, however reliably produced, have the significance of undertaking *discursive* commitments. What is left is not a *cognitive* affair at all.

The most serious objection to a pure reliability theory accordingly is presented not by the general Wittgensteinian strand of thought concerning the significance of gerrymandering for attempts to construe norms as regularities, which Kripke expounds so forcefully. The most serious objection to reliabilism stems rather from the more particular Sellarsian insight concern-

ing the essential role played by the specifically inferential articulation even of noninferential reports. Sellars insists first that for a performance elicited by a reliable differential responsive disposition to be a candidate for expressing knowledge, it must count as an *endorsement* by the reporter of some claim, as the undertaking of a *commitment*. Furthermore, he recognizes that the identity and content of such commitments depend on their role in inference and justification, in giving and asking for reasons. He sees further that being capable of endorsing a claim requires grasping the role of that claim in inference and justification—that the official tripartite analysis of knowledge implicitly involves *understanding*, as part of what is required for *belief*.

Unfortunately, motivated by these insights, Sellars stakes out far too strong an antireliabilist position concerning the role of inferential justifying in entitlement to claims to observational knowledge. There is accordingly a danger that where the various strands of thought are not carefully sorted out, distaste for the epistemological internalism about cognitive authority that Sellars endorses will obscure the important lessons that ought to be drawn from his account. His basic point is that a noninferential reporter must be “in the space of giving and asking for reasons,” in addition to having the right differential responsive dispositions. That space is, for Sellars as here, articulated by relations of *authority*, and to be in that space one must be able to recognize or acknowledge the authority of claims. It is, in other words, in virtue of one’s capacity to adopt practical deontic attitudes, to take or treat something as having cognitive authority, that one counts as moving in the space of giving and asking for reasons.

## 2. *Sellars on the Authority of Noninferential Reports*

In the passages (from “Empiricism and the Philosophy of Mind”) that follow, Sellars is concerned with the nature of the authority (or as he sometimes puts it, “credibility”) of noninferential reports (Carnap’s *Konstatierungen*), which are claims to observationally acquired knowledge. The distinctive feature of such reports is that “the credibility of such tokens as ‘express observations’ [is] a credibility which flows from tokens to types.”<sup>12</sup> This contrasts with the credibility of nonobservational claims such as “Dogs are mammals,” which is attached to tokens in virtue of their being of types that are credible. The reliability approach then recommends itself as offering a simple and natural account of the source and nature of the credibility of sentence tokenings that report empirical observations: “An overt or covert tokening of ‘This is green’ in the presence of a green item is a *Konstatierung* and expresses observational knowledge if and only if it is a manifestation of a tendency to produce overt or covert tokens of ‘This is green’—given a certain set—if and only if a green object is being looked at in standard conditions.”<sup>13</sup>

What does such reliability have to do with *authority*? “The first hurdle to

be jumped concerns the *authority* which, as I have emphasized, a sentence token must have in order that it may be said to express knowledge. Clearly, on this account the only thing that can remotely be supposed to constitute such authority is the fact that one can infer the presence of a green object from the fact that someone makes this report.<sup>14</sup> This is an important move. The authority of reliability consists in its underwriting a propriety of *inference* (what might be called "the reliability inference"). The noninferential undertaking of a commitment by a reliable reporter can *inferentially authorize* another to undertake a commitment with that content. To take or treat someone as a reliable reporter (in certain circumstances) is for a scorekeeper to endorse the propriety of the move from *attributing* to the reporter a noninferentially acquired doxastic commitment to the scorekeeper's *undertaking* of a corresponding commitment (and taking others to be similarly entitled). This notion of how reliability fits into the giving of reasons is the key to understanding the special sort of authority characteristic of noninferential reports, which in turn is essential to the notion of *empirically* contentful claims.

As is by now familiar, Sellars has already taken issue with the sort of foundationalism that sees empirical knowledge as an inferential superstructure raised on an autonomous noninferential base. The target of his criticism

is the idea that there is, indeed *must be*, a structure of particular matter of fact such that

- (a) each fact can not only be noninferentially known to be the case but presupposes no other knowledge either of particular matter of fact, or of general truths; and
- (b) the noninferential knowledge of facts belonging to this structure constitutes the ultimate court of appeals for all factual claims—particular and general—about the world.<sup>15</sup>

Sellars, inferentialist and antifoundationalist though he is, does not deny either (b) or the first half of (a). His quarrel is with the second half of (a).

There are particular instances of believing or being committed that are noninferential in the sense that their acquisition was not the conclusion of an inferential process. There are no beliefs or discursive commitments that are noninferential in that what is expressed by a sentence can be *understood* without mastering inferential relations that content stands in to others. So a bit of knowledge (belief) can, and indeed all of it does, presuppose other knowledge (belief), even though it is not inferred from that other knowledge or belief. This possibility was not seriously examined by the classical epistemological tradition. It is a certain hierarchical picture of *understanding* (at this level a necessary condition of believing) that Sellars rejects. He does not object to a hierarchical picture of empirical *justification*, once that has been suitably disentangled from bad foundationalism concerning understanding. His claim that the authority that accrues to noninferential reports in virtue

of their being the results of reliable reporting or belief-acquiring mechanisms is a broadly inferential authority is in no way inconsistent with understanding observational knowledge to be authoritative in virtue of the reliable *noninferential* differential responsive dispositions that produce it. Inference need not be involved in the process that leads to a tokening of 'This is green'; but it is involved in grasping the type of authority that such noninferentially produced tokenings have, and so in understanding such tokenings, and so in their being potential expressions of knowledge.

The question is just what the relation is between mastery of this inference and possession by a tokening of the sort of authority characteristic of ground-level observational knowledge. Sellars's claim is that "to be the expression of knowledge, a report must not only *have* authority, this authority must *in some sense* be recognized by the person whose report it is."<sup>16</sup> The notion of claims *having* cognitive authority is indeed intelligible only in connection with practical attitudes of *taking* or *treating* claims as having such authority. Sellars has suggested that the authority distinctive of observational knowledge should be understood in terms of the correctness of an inference, from the making of a report such as 'This is green' by one whose differential responsive dispositions are taken to be reliable to the undertaking of a commitment to the effect that there is something green there.

He concludes: "In other words, for a *Konstatierung* 'This is green' to express observational knowledge, not only must it be a *symptom* or sign of the presence of a green object in standard conditions, but the perceiver must know that tokens of 'This is green' *are* symptoms of the presence of green objects in conditions which are standard for visual perception."<sup>17</sup> So Sellars's view is that the reliable reporter can count as being entitled to a noninferentially acquired commitment, and so the assertion by which that commitment is acknowledged can be cognitively authoritative in licensing or entitling others by the standard assertional mechanism of communicative entitlement inheritance, *only* if the reporter can inferentially *justify* the noninferential claim. Such a justification consists precisely in exhibiting the inference whose premises are the reliability of differential responsive dispositions to make such claims and responsive elicitation of the claiming in question and whose conclusion is another tokening of the claim itself.

### 3. *Attributing Reliability Is Endorsing an Inference: An Inferentialist Middle Way between Justificatory Internalism and Reliabilist Externalism*

There are two problems with this conclusion. First, Sellars takes it that for the claim of the reliable observer to *be justified*, the observer must be able to *justify* it inferentially—to offer reasons by displaying premises from which it follows. Second, he assumes that such justification must involve explicit invocation of reliability, that is, that a claim of reliability

must be one of the premises. Thus Sellars claims that “observational knowledge of any particular fact, e.g. that this is green, presupposes that one knows general facts of the form *X is a reliable symptom of Y*.<sup>18</sup>

This latter is an odd move for Sellars to make. He, after all, is the one who urged that material proprieties of inference not be everywhere traded in for suppressed premises. Why should it not be that reliability underwrites a propriety of inference, without the claim of reliability having to appear as a premise in the inferences so underwritten? Here Sellars overreaches himself. He is right to insist that reliability matters because it warrants inferences of a certain form. He is wrong to insist that this warranting must be understood in terms of endorsement of an explicit claim that can serve as a premise in inference.

The first claim also seems too strong. Securing entitlement to a claim need not always be assimilated to inferential justifying of the claim. The possibility of vindication of a commitment by deference rather than inference—by the invocation of communicational mechanisms depending on intracontent interpersonal testimony rather than intrapersonal intercontent justifying—shows that much. It would be a mistake to assimilate deferential entitlement inheritance to inferential entitlement inheritance by insisting that the one invoking the authority of another’s assertion be able to produce an explicit argument in which a claim as to the informant’s reliability would appear as a premise. Rather, deferring involves an *implicit* claim as to the reliability (more particularly the entitlement in this case) of the informant. One who accepts the deferring as successful (and so attributes the claimed entitlement to the one deferring) thereby implicitly endorses the propriety of a permissive inference from the informant’s claiming that *p* to *p*—which is just the inference that Sellars picks out as corresponding to reliability. But at the ground level, all of this can be made sense of as implicit in what is done in practice. Because it can, it is possible to explain the expressive role of the locutions that can be introduced at a later stage to make these attitudes explicit. Assimilating entitlement to the commitments acknowledged by noninferential reports, and therefore their authority, to that secured by explicit inferential justification, as Sellars does, is a mistake of the same sort.

A symmetrical mistake would be to assimilate the authority of noninferential reports to that of testimony, by understanding the invocation of such authority as a kind of deference to a “world-asserter.”<sup>19</sup> The structure of authority exhibited by noninferential reports is *sui generis*, to be reduced neither to that of inferential justification nor to that of testimony. These three are mutually irreducible—none can take over the function of any of the others. One of the primary explanatory aims of this work is to explain how commitments that are implicit in the fundamental practices that confer assertible conceptual content can eventually themselves be made explicit and assertible, expressed in a form in which reasons can be given and asked for them. The implicit attitudes that can in this way be explicitly expressed

once suitable vocabulary has been introduced include those involved in invocations and recognitions of the authority of both testimony and reports of observations. But the locutions that play these explicating roles can themselves be made intelligible only by understanding first the implicit structures they bring out into the assertional light of day.

In fact, Sellars's insight concerning the irreducible role played by inferential justification does not require insisting that noninferential reporters can be authoritatively entitled to their claims only if they can justify them. As just indicated, the inference from the noninferential undertaking of a commitment as the result of a reliable differential responsive disposition to endorsement of the claim thereby made may be implicit in practical attitudes, rather than explicit in claims offered as justifications for that endorsement. Nor is it necessary that the one who makes an observation report endorse the propriety of that inference, even in this implicit practical sense. Reliability may entitle the reporter to the knowledge claim, may qualify it as knowledge, even if the reporter does not even implicitly endorse the inference that is the practical acknowledgment of the authority of reliability. This is the primary insight that stands behind the justificatory *externalism* of reliability epistemologies, in contrast to the justificatory *internalism* Sellars exemplifies.

Suppose that Monique has been trained reliably to discriminate hornbeams by their leaves. As a result of the training, she is often disposed to respond to the visibility of leaves of the right sort by noninferentially reporting the presence of a hornbeam. She understands what it means to claim that something is a hornbeam and, in circumstances appropriate for such reports, actually comes to believe that there is a hornbeam present. She may still be uncertain of her discriminatory capacity long after she has in fact become reliable.<sup>20</sup> In such a situation she may have a true belief that there is hornbeam in front of her, yet be completely unable to justify that claim (for instance, by citing features distinctive of hornbeam leaves), and even deny that she is a reliable noninferential reporter of hornbeams.

Yet, the reliabilists point out, it can be entirely in order for one who does take her to be a reliable reporter of them, not only to come to believe that there is a hornbeam present on the basis of her report, but to cite her report (at least deferentially) as what warrants that belief. This is treating the claim as authoritative in just the way that is required for knowledge. Someone who thus takes her to be reliable can accordingly attribute to Monique the knowledge that there is a hornbeam in front of her, in spite of her protestations to the contrary. What makes her claim knowledge (according to the attributor) is the fact of her reliability (according to the attributor), regardless of her attitudes toward that reliability.<sup>21</sup> The status of her claim as knowledge is accordingly external to her attitudes—not only because of the truth condition on knowledge, but also because of the entitlement condition. Sellars is committed to withholding the attribution of knowledge in the absence of the

candidate knower's capacity to justify the claim, and so is committed to disagreeing with reliabilists about examples like this. Yet on this point the reliabilists are surely correct.

Sellars, however, is right that for a reliably elicited differential response to be a candidate for knowledge, the one making the knowledge claim must be in the space of reasons, must be capable of understanding the claim, and so must have some grip on its role in reasoning, hence on its use as a premise and conclusion of inferential justifications. Requiring this general capacity, of course, falls short of requiring that on each occasion the reporter must be able to justify the claim for it to count as the expression of observational knowledge. Furthermore, while reliabilism about cognitive entitlement and so cognitive authority is clearly correct that knowledge can be attributed even where the one to whom it is attributed cannot demonstrate entitlement to the claim inferentially, by providing a justification that appeals to other claims the putative knower endorses, it does not follow from this observation that reliability by itself is enough for entitlement and cognitive authority, apart from all consideration of attitudes of taking or treating the knower as reliable, as a thoroughgoing externalism about entitlement would have it. Sellars is also right to insist that attributions of knowledge require not just reliability but at least implicit endorsement of the inference that is the practical acknowledgment of reliability—the inference namely from the occurrence of a report, or the noninferential undertaking of the commitment such a report expresses, to the endorsement of the claim.

Where Sellars is wrong, as the sort of example just considered shows, is in thinking that the one who endorses this inference must be the one who *undertakes* the claim to observational knowledge. It has been pointed out that attributing knowledge is a hybrid deontic attitude involving not only the attribution of commitments but the undertaking of them. Not only does the attributor of knowledge take the candidate knower to endorse a claim; the attributor also must endorse that claim.

It is likewise the *attributor* of observational knowledge who must attribute reliability to the knower. Attributing such reliability is endorsing exactly the general form of permissive inference that Sellars points to—treating as appropriate the inference from the noninferential undertaking of a commitment (of the right sort, and in circumstances of the right sort) by the observer to the endorsement by others of the claim so elicited. Taking a report that is the outcome of a particular differential responsive disposition as entitling others to the claim (for instance accepting as entitling their deferrals to the reporter on such issues) just is treating the reporter in practice as reliable about such matters. Monique need not, *pace* Sellars, take herself to be a reliable reporter of hornbeams in order for her to count as knowing observationally that there is a hornbeam in front of her. But the one who attributes such knowledge must take her to be reliable. And adopting that practical attitude is endorsing the pattern of permissive inference that connects the

*attribution* by others to Monique of a noninferentially acquired belief about the visible presence of hornbeams with their *undertaking* of a commitment to the visible presence of hornbeams in Monique's vicinity.

Just as the truth condition on knowledge requires that the attributor of knowledge undertake, as well as attribute, commitment to the content of the knowledge claim, so satisfying the entitlement condition by mere reliability requires that the attributor of knowledge undertake (but not necessarily attribute) commitment to the propriety of the reliability inference. Where the language is rich enough to include the expressive resources necessary to make the reliability inference explicit (conditionals and 'claims that . . .' or 'believes that . . .'), attributors of knowledge can be challenged and called on to defend their endorsement of the conditional "If Monique claims (sincerely, responsively, and in appropriate conditions) that a hornbeam is visibly present, then (probably) a hornbeam is present." At this point, reliability could be invoked to *justify* the belief that there is a hornbeam present. But this is a sophisticated, late-coming possibility, built on the implicit acknowledgments already described. So full-blown reliabilist externalism about cognitive entitlement is mistaken in ignoring the necessity of such inferential attitudes on the part of attributors of knowledge, while full-blown Sellarsian internalism about cognitive entitlement is mistaken in insisting that the knower must have such attitudes. These are complementary ways of misunderstanding the essentially social structure of the cognitive deontic attitudes, in terms of which the status of a claim as knowledge must be understood.

#### 4. *Observational Knowledge and Empirical Conceptual Content*

The noninferential authority possessed by claims issuing from the exercise of reliable differential responsive dispositions—although reducible neither to the sort of interpersonal authority invoked by deferring nor to the sort of intercontent authority invoked by inferring—is not fundamental in the way that those structures of authority are. In the model of assertional practice that has been put on the table, communication and justification are two aspects of the game of giving and asking for reasons; neither is intelligible except in the context of the other. They are intelligible, however, in the absence of noninferential responsive authority. Practices that do not involve according any knowledge claims the significance of observation reports can nonetheless be understood as instituting specifically *assertional* significances, and so as conferring specifically *propositional* contents.

What is missing from such practices is claims with *empirical* content. Discourse recognizable as mathematical can be like this: reasons are given and demanded; claims communicated, challenged, and justified; and regresses of entitlement inheritance halted by appeal to axioms, free moves that anyone is treated as entitled to at any point in the conversation. Our

discourse is not in general like this, however, and the sorts of contents our claims have cannot be conferred by assertional practices that do not acknowledge some claims as having empirical authority stemming from their status as reports of observations. Indeed, it is essential to the contents of the ordinary concepts in terms of which we conduct our lives that they stand in inferential relations both to the acknowledgments of commitments resulting from what Sellars calls “language entry transitions,” in perception, and to the acknowledgments of commitments that result in what Sellars calls “language exit transitions,” in intentional action.<sup>22</sup> (The contribution of the latter *practical* empirical structure is discussed in the second half of this chapter. Attention is restricted here to the *cognitive* empirical structure.)

The practical significance characteristic of claims to observational knowledge is best understood in terms of the role they play in the default-and-challenge structure of entitlement. Noninferential reports can function as unjustified justifiers: claimings that are treated as having a defeasible default status as entitled. Properly made claims to observational or perceptual knowledge can accordingly provide entitlements that can then be inherited either inferentially or communicationally. So observation provides regress-stoppers, and in this sense a foundation for empirical knowledge. This is what stands behind Sellars’s endorsement of the claim (quoted above) that “noninferential knowledge of facts . . . constitutes the ultimate court of appeals for all factual claims—particular and general—about the world.”<sup>23</sup>

Default entitlements are of two sorts, depending on whether the entitlement attaches to a commitment in virtue of the type it instantiates or in virtue of the circumstances in which it is tokened.<sup>24</sup> There are sentence types that would require a great deal of work for one to get into a position to challenge, such as “Red is a color,” “There have been black dogs,” “Lightning frequently precedes thunder,” and similar commonplaces. These are treated as “free moves” by the members of our speech community—they are available to just about anyone any time to use as premises, to assert unchallenged. Noninferential reports, by contrast, have their default entitlement status as a result of the way in which the report tokening, or the particular acknowledging of the commitment that would be expressed by such a tokening, is elicited through the exercise of a reliable differential responsive reporting disposition. Treating such a claim as one the reporter is entitled to involves an implicit commitment on the part of the attributor to the actual circumstances being among those in which the reporter is responsively reliable concerning the sort of matters reported.

There will typically be some sorts of reports such that under appropriate reporting conditions (the same for all), essentially all the members of the linguistic community are reliable. Almost anyone can, under suitable circumstances, tell whether it is a warm day out or whether the marble one is holding is approximately round. Other sorts of reports involve not only more specialized circumstances but specialized training. Particle physicists are

trained reliably to respond noninferentially to the presence of mu-mesons in a bubble chamber by reporting the presence of mu-mesons. Not all of us can do this reliably. It is only someone who is taken not only to be looking at a bubble chamber but also to be properly trained to be reliable about these matters (and who has the right sort of collateral beliefs) whose reports will be accorded noninferential entitlement and the corresponding authority. As Quine says, what is observable varies from community to community.<sup>25</sup> He understands the status of being an observation report for a community (perhaps a proper subset of the whole linguistic community) in terms of what that community can agree on under concurrent stimulation, that is, in the same standard reporting circumstances.

The authority of noninferential reports requires the collaboration of both dimensions into which Sellars analyzes them: not only that they arise from the exercise of reliable differential responsive dispositions but that the response is to endorse a claim, to acknowledge a commitment, with a certain content. What makes it a mu-meson that the physicist is reporting rather than the hooked vapor trail that also forms part of the reliably covarying chain of events culminating in the report is to be understood not in terms of the differential responsive dispositions but in terms of the inferential role of the claim being made. The consequences that can be inferred from the presence of a mu-meson are quite different from those that can be inferred from the presence of a hooked vapor trail, covariant and concomitant though these phenomena may be. For instance, mu-mesons are much smaller, and move much faster, than the vapor trails they produce (see further at 7.1.6 below). As Quine argues further, it is important to understand that under the appropriate circumstances, which include the presence of a bubble chamber or similar device, and for the right community of observers, mu-mesons are literally observable—noninferentially reportable in much the same sense in which red things are for the rest of us. It is a mistake to think that what is really noninferentially observed is only the vapor trail and that the presence of mu-mesons is only *inferred*. Such an inference can be made, and learning to make it might be part of the training process that leads to becoming a reliable observer of mu-mesons (in bubble chambers). But coming to be disposed reliably to respond to the vapor trail, and *hence* to the presence of mu-mesons, by asserting or acknowledging a commitment to the presence of a mu-meson is learning to *observe* mu-mesons, to report them *noninferentially*. And this is so even if one is not totally reliable, in that there are circumstances in which one would *mistakenly* report the presence of a mu-meson *because* of the presence in the chamber of a vapor trail indistinguishable from those one has learned to respond to noninferentially by reporting mu-mesons (just as the fact that one can be fooled by a cunning replica does not preclude one from seeing a sparrow in the cases where one is not being fooled).

The claim is, then, that one is *directly* observing mu-mesons, in the sense

of *noninferentially* coming to be aware of them, to make claims about them, to know something about them—rather than indirectly, inferentially coming to a conclusion about mu-mesons on the basis of an *inference* (perhaps unconscious or implicit) from the presence of a vapor trail with a certain shape. This claim may seem implausible in light of the common practice of retreating, under certain suitable sorts of challenge, from the claim that a mu-meson is present to the claim that a hooked vapor trail is present. Is not such a retreat to be understood as relinquishing commitment to an inference, and therefore to its conclusion, while continuing to defend its (genuinely noninferential) premise? No.

Such cases ought to be understood as retreats (given a credible challenge to the effect that this might be one of the cases in which the exercise of a generally reliable capacity nonetheless leads one astray) to a claim that is *safer*. Being safer in this sense, however, is not a matter of withdrawing endorsement of an inference. One retreats to a different report with respect to which one is *more* reliable, as measured for instance by percentage of correct differential responses in the relevant circumstances, or by the same percentage of correct responses within a wider range of circumstances, or by the size of the community that does not share esoteric theoretical beliefs but does share the differential responsive disposition and corresponding capacity to make noninferential reports. Doing this can amount to offering an inferential justification of the original noninferential belief, by explaining how one was able to see it. An analogous case would be explaining that there was a mirror, not apparent to the audience assessing the authority of the claim, but apparent to the reporter, in order to explain how one was noninferentially able to report something that a challenger has pointed out is around a corner and so ought to be invisible.

The possibility of such an inferential justification of a claim on the basis of a safer claim does not show that the original claim should be understood as itself the product of a process of inference, any more than the capacity of sophisticated reporters to offer justifications of their claims to observational knowledge by citing their reliability and appealing to the reliability inference shows that their original claim was arrived at as a result of an inference from that premise. Nor does the fact that the capacity to make certain sorts of noninferential reports depends on collateral beliefs show that those reports are really inferences from something more basic, together with those collateral beliefs. One must have many beliefs about mu-mesons in order to be able to understand and so to make *any* claims about them, noninferential or otherwise. That does not preclude one from coming to be able to observe them.

The basis of observational knowledge, then, is that it should be possible to train individuals reliably to respond differentially to features of their environment by acknowledging doxastic commitments. Those commit-

ments are inferentially related to others that not only play inferential roles but also are themselves appropriately elicited noninferentially by features of the environment. These cross-connections put constraints on endorsements of inferences relating expressions whose circumstances of appropriate application include noninferential ones. Both oranges and orange things can be noninferentially reported, so someone who reports the presence of an orange after tasting but not seeing it and then infers from its being an orange to its being orange in color is liable to be challenged by another who is in a position to report it noninferentially as purple. For the commitment entitlement to which was acquired noninferentially is incompatible with that entitlement to which was acquired inferentially. Either the identification of the orange by taste, which formed the premise of the inference, or the identification of its color as purple might itself bear further challenge and investigation; but if these stand up, the reliability of the inference from being an orange to being orange in color will be impugned. In this way the possession of noninferential circumstances of appropriate application of some concepts imbues them with empirical content—recognizable as conceptual content in virtue of its inferential articulation and as empirical in virtue of its dependence on the noninferential acquisition of commitments to those contents (and of entitlements to those commitments).

Similarly, the inferences from circumstances to consequences of application (which are implicit in conceptual contents) are subject to empirical criticism in virtue of inferential connections among the contents of commitments that can be acquired noninferentially. So it may happen that one uses the term 'acid' in such a way that a substance's tasting sour is a sufficient condition for applying it, and that it will turn litmus paper red is a necessary consequence of applying it. Finding a substance that both tastes sour and turns litmus paper blue shows that such a concept is inadequate. Conceptual contents can accordingly be criticized, groomed, and developed empirically in a way parallel to the sort of Socratic process discussed in Chapter 2. In virtue of their inferential connections to concepts that can be used to make reports, even purely theoretical concepts (those whose *only* circumstances of appropriate application are inferential) inherit empirical content and have the inferences they are involved with constrained by the commitments and entitlements actually thrown up by what is responded to noninferentially. That the reliable differential responsive dispositions underlying this structure of noninferential authority are dispositions to acquire *commitments* and *entitlements* to those commitments, that is, to alter deontic status, means that the practices they appear in must include corresponding practical deontic *attitudes*. Something practitioners can do must be the taking or treating of performances as having the significance of noninferential reports, the recognition of the status of some claims as deriving their entitlements from their being expressions of reliable differential responsive dispositions to ac-

knowledge commitments. For one cannot make sense of normative significance, even the normative significance of reliability, apart from consideration of its uptake or attribution.

### 5. *Attributing Observational Entitlement*

It is straightforward to extend the model of assertional practice as outlined so far to encompass the structure of authority in virtue of which claims can have and be treated as having the significance of noninferential reports. What is primarily required is to say what it is for one interlocutor to attribute noninferential or observational authority to the claim of another, thereby recognizing or acknowledging it as having a special sort of entitlement. The authority involved is entitlement heritable by the usual intrapersonal intercontent inferential and interpersonal intracontent communicational pathways. What is distinctive of observational authority is that such authority is accorded to particular *tokenings* of acknowledged commitments (rather than to their types) and the way in which that authority depends on a special combination of content-based and person-based features. For the imputed reliability of an observer varies from content to content within each observer, and from observer to observer—someone who is taken to be able reliably and noninferentially to discriminate mu-mesons in bubble chambers may not be taken to be able to do so for '52 Pontiacs in traffic.

Furthermore, if the topic is fixed (the concepts used in the reports being assessed are specified), imputed reliability still varies depending on circumstances. For each particular observable, there will be an associated set of appropriate circumstances of reporting, according to the one attributing reliability and so observational authority. The authority of a reliable reporter is conditional on the obtaining of those appropriate circumstances. Even a lookout who is in general a reliable reporter of whales must be facing in the direction of what is being reported, cannot see well in the direction of a horizon-hugging sun, is less reliable if there are large walruses about, and so on. These appropriate circumstances of reporting, associated with the observable content (and perhaps the individual reporter), figure as commitments undertaken by the one attributing or assessing the responsive authority of a claiming.

So associated with each sort of noninferential authority a given interlocutor grants to another (the product of a person and a kind of content, for example, reports of the presence of whales, or mu-mesons), there is a set of enabling conditions (looking in the right direction, looking in a bubble chamber) and a set of defeating conditions (presence of many walruses in the vicinity, physicist drunk and woozy), and it is the interaction of these, according to the commitments undertaken by the one assessing noninferential authority, that determine whether responsive entitlement is attributed or

not. If the assessor undertakes commitment to a suitable range of the enabling reporting conditions and does not undertake commitment to any of the defeating reporting conditions, the reporter's claim is treated as having a default status as entitled. In this way the empirical authority of some attributed commitments, on the basis of implicit inferential acknowledgment of reliability under suitable circumstances, is distinguished from the type-based default status of "Red is a color" and "There have been black dogs," which do not exhibit a similar relativity to person, content, and the environing conditions as they are taken to be by the assessor. Observational authority is accordingly another hybrid deontic status: attributing it involves not only *attributing* commitments and entitlements but also *undertaking* or *acknowledging* them by endorsing reliability inferences.

Once the attitude of taking or treating someone's performance as having the significance of a noninferential report whose authority is grounded in the local and conditional reliability of the observer is in place, it is possible to introduce a type of performance that is the claiming of or petitioning for such authority by an observer. A certain sort of noise or gesture (perhaps a shrug) can come to have the significance of *invoking* observational authority. Then if a report is challenged, it can be vindicated (its entitlement demonstrated) by invoking observational authority rather than by deferral or inferential justification. But there is no strict need for practices encompassing empirical conceptual contents to include a speech-act kind with this significance. It is enough if interlocutors sometimes accord such authority, and thus take the commitments acknowledged by noninferential reports in some circumstances to be vindicated (implicitly, according to the assessor's attitudes) by the fact of the reporter's reliability. Where there is such a speech act, it would implicitly mean something like "I see it (for example, that it is red)." An explicit assertion to this effect can be introduced as well, but just how will not be clear until Chapter 8, where pragmatically explicating locutions such as 'believe that' and 'claim that' are officially introduced into the model of assertional practices. The significance of an invocation of observational authority does not depend on any assertionally explicit content that the invocation might have. (Compare Wittgenstein's suggestion that if challenged to say how one knows that the thing in front of one is red, one might say simply, "I speak English.")

#### 6. Expressions of Belief That Are Not Claims to Knowledge

The account here of assertions as claims to knowledge turns on the implicit obligation to vindicate the commitment undertaken by demonstrating one's *entitlement* to it. The foregoing discussion of observation and reliability focused on the importance for the *attributor* of observational knowledge claims of implicitly attributing reliability. Adopting that attitude requires endorsing the inference from the attribution of a noninferentially

responsively elicited acknowledgment of a *commitment* (under suitable circumstances and for a qualified observer) to the attribution of *entitlement* to that commitment. It is in the context of concern with entitlement to assertional commitments that the complaint was levied against reliabilists that they ignore the inferentially articulated attitude in which recognition or attribution of entitlement consists. The corresponding objection to Sellars was that, while appreciating the significance of that hybrid practical deontic attitude, he inappropriately insists that the reliability inference it involves be endorsed by the one *making* the observation rather than the one attributing or assessing it.

But this concern with entitlement can seem out of place in a discussion of a sort of discursive commitment that is intended to do the sort of explanatory work characteristically performed by a notion of *belief*. If belief is to be understood in the first instance as the state or status expressed by assertional speech acts, it seems wrong to treat assertions as also involving a claim to *knowledge*. For expressing a belief and claiming to know are different.

When an idiom is developed to the point that it has the expressive resources provided by the English locutions 'believes that' and 'knows that'—which make the pragmatic status being attributed or undertaken explicit as part of the *content* of what is claimed—it becomes possible to say of someone else, "He believes that Arnauld did not write *The Art of Thinking*, but he does not know it." The case has already been considered where what is expressed is the attitude of a scorekeeper who attributes commitment to a claim but does not endorse the attributed claim—that is, does not take it to be true. It is also possible, however, to distinguish expressions of mere belief from claims to knowledge in the first-person case, in which the claim *is* being endorsed or taken-true. In such cases, the social-perspectival distinction between attributions of knowledge and attributions of belief cannot get a grip.

For although undertaking an assertional commitment *is* taking-true the claim, a difference can arise precisely over the issue of entitlement or justification. The attribution of knowledge may be withheld by a scorekeeper who attributes a commitment without attributing a corresponding entitlement. Indeed, sometimes we make claims while fully aware that they may legitimately be challenged and that we are not in a position to vindicate them by demonstrating our entitlement to them. This is the implicit attitude that becomes assertionally explicit in claims such as "I believe that Arnauld did not write *The Art of Thinking*, but I don't claim to know it." For this sort of reservation can concern not the truth of the belief but my capacity to justify it. I may continue to take the claim to be true, to endorse it, to acknowledge the commitment it expresses, and yet not be prepared to shoulder the justificatory burden associated with a knowledge claim. This might be because I have forgotten the source of my conviction, or it might be an expression of a claim's having a ground-level status for me as an unjustified justifier that I do not take to be widely shared—I just believe that people with beards cannot be trusted, or that house cats are dangerous.

The speech acts that express such attitudes are what might be called *bare* assertions,<sup>26</sup> corresponding to *mere* beliefs, without the implicit claim to entitlement that is demonstrable should someone become entitled to challenge it (paradigmatically by expressing an entitled commitment to a claim incompatible with it). Does not the possibility of such bare expressions of commitment without claim of entitlement, of conviction without warrant, show that it is a mistake to understand claims on the model of claims to *knowledge*? No. Such claims are intelligible only as exceptions against a background of practices in which claims typically have the significance of claims whose authority is redeemable by demonstration of warrant. The possibility of bare assertion is parasitic on the possibility of assertions that implicitly involve undertaking a conditional task-responsibility to demonstrate the asserter's entitlement to the commitments undertaken by the performance of speech acts of that kind.

For bare assertions and the commitments they express would be completely idle if they could not figure as premises in inference and could not be passed along in communication. This is to say that bare assertions involve something of the *authority* of full-blooded assertions, while disavowing the corresponding *responsibility*. Yet that authority (licensing inferences by the asserter to commitments with other contents and the undertaking of commitments with the same contents by other interlocutors) makes sense only in a context in which inferential and deferential invocation of such authority can be demanded. What assertions are *for* is justifying other assertions. To accept someone's bare assertion is to take it to be a claim from which conclusions can be drawn. But giving reasons presupposes the possibility of asking for them, or at least the possibility that claims often stand in need of reasons. A game of giving and asking for reasons cannot consist exclusively in the exchange of speech acts that are accorded the significance of bare assertions. Within the broader context of full-blooded assertions (which do involve the demonstration of entitlement by inference and deference), however, it is possible to make sense of treating some claims as having the significance of bare assertions. Assertionial commitments essentially involve the dimension of entitlement. Assertions are paradigmatically knowledge claims, and the sort of belief they express is unintelligible except in relation to the possibility of assessing beliefs for their status as knowledge, as warranted and true.

#### IV. RATIONAL AGENCY

##### 1. *Methodological Constraints on the Conception of Practical Rationality*

Beliefs make a difference both to what we say and to what we do. They manifest themselves both linguistically, in assertions, and practically,

in actions. A basic criterion of adequacy for any theoretical account of this fundamental sort of intentional state is that it explain both of these ways in which beliefs can be expressed in behavior, and the relation between them. The methodologically parsimonious idea that one or the other of them ought to be accorded explanatory priority is the motivation common to both of what Stalnaker (in the rough-and-ready botanization alluded to in the previous chapter) distinguishes as the "linguistic" and the "pragmatic" approaches to intentionality. Theories of the sort he calls "linguistic" construe believing by analogy to claiming: as a kind of inner asserting of sentences. They are accordingly obliged, first, to explain assertions without appeal to their role as expressions of belief and, second, to explain the norms that determine the role of belief in rational agency in terms of the proprieties that govern the public use of sentences. Theories of the sort he calls "pragmatic" (such as the one Stalnaker himself endorses), in contrast, take the role of belief in intentional action to be primary. They then owe both a nonlinguistic explanation of rational agency and an account of speech acts, paradigmatically assertion, in terms of intentional states so understood.

The approach pursued here takes belief to be intelligible only in the context of social-linguistic practice. But it is a relational, rather than a reductive, linguistic theory. Although doxastic commitment (the sort of deontic status corresponding to the intentional state of belief) cannot be made sense of apart from the possibility of expressing such commitments by performing speech acts that have the significance of assertions, neither can assertional significance be made sense of without reference to the commitments such speech acts undertake and acknowledge. As regards asserting and believing, the theory is even-handed; it accords explanatory priority to neither one. It nonetheless deserves to be called a *linguistic* account of intentionality (in a sense broader than Stalnaker's<sup>27</sup>) inasmuch as it does accord explanatory priority to the linguistic manifestation of belief in assertion over its practical manifestation in action.

As Dennett and Davidson have emphasized, attributing propositionally contentful intentional states such as beliefs to a creature is taking it to be *rational*. Thus Kant uses the rubrics of *theoretical* and *practical* rationality to distinguish the sort of normative competence manifested in giving and asking for reasons for *claims* or *judgments* from the sort of normative competence manifested in giving and asking for reasons for *actions*—judgments and actions being picked out precisely as the sorts of things reasons can be given for and for which reasons can be asked. Being rational is understood here generically as being able to play the game of giving and asking for reasons, which is to engage in a specifically *linguistic* social practice. For one cannot give reasons unless one can make claims. Doing so requires mastery of the normative dimension of inference: a practical grasp of the notion of *right* reasoning, of the distinction between correct and incorrect inference. Assessing performances as correct or incorrect is adopting normative atti-

tudes that are intelligible only in a context of interpersonal scorekeeping—even though in such a context it is possible to make sense both of self-assessment and of assessments of objective correctness, for which no one's scorekeeping attitudes are counted as decisively authoritative.<sup>28</sup> First-person deliberation is the internalization of such third-person assessment.

To take this line—identifying the rationality that qualifies us as sapient with being a player in the normative game of offering and assessing, producing and consuming reasons—is to deny two widely held reductive conceptions of rationality: one that identifies rationality with *logical* competence, and another that identifies it with prudence or *instrumental* competence. All parties can agree that to be rational is to distinguish good inferences from bad inferences. The disagreement concerns whether 'good inference' in this formula can be restricted to *logically* good inferences, or again to *instrumentally* good inferences—ones whose correctness is determined by their utility in satisfying desires or maximizing preferences. Logical competence is mastery of the use of locutions by means of which inferential proprieties are made *explicit* as the contents of claims. This theoretical ability to codify practices as principles accordingly presupposes prior practical mastery both of *implicit* inferential proprieties and of the use of the ordinary, nonlogical claims they articulate and govern. When this expressive role of logical vocabulary is appreciated, the identification of rationality in general with its manifestation as logical manipulation is unmasked as another form of the intellectualism that insists on discerning a propositionally explicit principle underlying every implicit property of practice—a form of platonism whose remedy is a complementary pragmatism.

Identifying rationality in general with the sort of instrumental rationality manifested in rational agency also inverts the proper order of explanation.<sup>29</sup> For the propositional contents of the intentional states appealed to in practical reasoning presuppose assertional-inferential proprieties, and hence linguistic social practices. (Though to say this is not to deny that proprieties of practical inference also contribute to the propositional contents of the states and expressions caught up in them.) To make out this claim it is necessary to say something about the practical reasoning that is implicitly attributed in interpretations of individuals as rational agents. In particular, just as it was shown how the capacity for logical reasoning is to be made intelligible in terms of (as the explication of) a conceptually prior capacity for nonlogical reasoning, it must also be shown how the capacity for practical reasoning incorporates and depends upon a conceptually prior capacity to give reasons for claims, rather than for actions.

Only an account of assertion of the sort introduced in Chapter 3 leaves room for the pursuit of such an order of explanation. Everyone ought to agree that asserting is putting forward a sentence as true. Following Davidson's lead, it has been suggested that distinguishing practical attitudes as taking or treating something *as true* requires a specifically linguistic social context of

mutual interpretation—that is, attribution of doxastic commitments, of the sort that has been elaborated as assertional-inferential scorekeeping. The next chapter develops the idea that this principle is best exploited by starting with an antecedent notion of assertional significance and then moving via that principle to an understanding of what is involved in talk of truth.

Commitment to understanding rational agency in terms of linguistic practice, rather than the other way around, strongly constrains the construal of the putting-forward portion of the principle that asserting is putting forward a sentence as true. For the claim that our first grip on the paradigmatic intentional state of belief (taking-true) is as what is expressed by assertions—rather than as what makes certain nonlinguistic performances intelligible in a way that is made explicit by exhibiting a piece of practical reasoning—is evidently incompatible with understanding asserting instrumentally, as a means intentionally adopted by a rational agent in order to achieve certain desired ends. If the linguistic practice of making and assessing claims (the game of giving claims as reasons and demanding reasons for claims) is an essential element of the context required to make sense of the notion of *propositional* intentional content (assertible, believable contents, which in English can be made explicit by the use of 'that' clauses), then what has been called "agent semantics" is not entitled to the conceptual raw materials it employs. In particular, one may not appeal to the *intentions* of the asserter—for instance intentions to say something true, or to make the audience believe that what is said is true, or to make the audience believe that it is uttered with the intention of saying something true or of engendering the corresponding beliefs. For what one is attributing can be identified as intentions to bring about various states of affairs only in virtue of the role such states play in a larger practical whole—one that includes the possibility of attributions of beliefs that the corresponding states of affairs obtain.

Less obviously, this order of explanation also precludes appeal to *conventions*, at least as commonly understood. The influential account offered by Lewis, for instance, takes a convention to be a social regularity that is sustained in a special way by the beliefs, intentions, and desires of the parties to the convention.<sup>30</sup> They are required not only to conform to the regularity but, among other conditions, to believe that others do so, to conform themselves because of that belief, to prefer that everyone conform, and to believe that everyone else has such beliefs and preferences. The present view is that on such a construal of convention, as Davidson concludes, "philosophers who make convention a necessary element in language have the matter backwards. The truth is rather that language is a condition for having conventions."<sup>31</sup> Construing the putting-forward bit of the principle that for a sentence to have assertional significance is for it to be put forward as true in terms of social conventions rather than individual intentions is also not an acceptable move according to this explanatory strategy. Certainly conventions of the sort that Dummett tries out—conventions to the effect that one

is to be understood as trying to utter sentences only with the intention of uttering true ones—are of no avail in the context of these explanatory commitments.<sup>32</sup> That is why it was necessary to move beyond explaining asserting in terms of *intentions* or *conventions* to explaining it instead in terms of *practices*, which themselves can be explained without appeal to intentions or conventions.

The next task is to show how that account of practices can be extended so as to encompass deontic statuses corresponding to the other sorts of intentional states that figure in the giving of reasons for *nonlinguistic* performances: the intentions and desires that play an essential role in the *practical* reasoning implicitly attributed by interpretations of individuals as rational *agents*. The aim is to provide a broadly Kantian account of the will as a *rational* faculty. By exploiting the analogy between discursive *entry* transitions in *perception* and discursive *exit* transitions in *action*, the rational will can be understood as no more philosophically mysterious than our capacity to notice barns or red things. A scorekeeping account can pick out performances (largely nonlinguistic ones) as intentional (under some specification) and hence as actions (under any specification) insofar as they are expressions of deontic attitudes—acknowledgments of a certain kind of commitment. Practical reasoning can then be understood as leading to performances with this sort of deontic significance. And on that basis, the expressive role of distinctively *normative* vocabulary can be specified in terms of its role in making explicit the endorsement of patterns of practical reasoning.

## 2. Acting and Perceiving

The general claim is that there are two species of discursive commitment: the cognitive and the practical. Acknowledging commitments of these two sorts is adopting deontic attitudes that correspond to the intentional states of *believing* and *intending*, respectively. A practical commitment is a commitment to *act*. The content of such a practical commitment is to *making-true* a claim. These commitments and their contents are intelligible only in a context that includes also the *taking-true* of claims. For it is in terms of such assertional taking-true that the success of actions, the fulfillment of practical commitments, must be understood. The category of cognitive discursive commitments accordingly enjoys a certain explanatory priority over that of practical discursive commitment. Each is essentially something that reasons can be given for and for which reasons can be asked, and one cannot give reasons unless one can acknowledge doxastic commitments by making claims.

The practical dimension of discursive practice can be understood by exploiting two ideas. The first is that practical commitments are like doxastic commitments in being essentially inferentially articulated. They stand in

inferential relations both among themselves and to doxastic commitments. The second idea is that the noninferential relations between acknowledgments of practical commitments and states of affairs brought about by intentional *action* can be understood by analogy to the noninferential relations between acknowledgments of doxastic commitments and the states of affairs that bring them about through conceptually contentful *perception*. The causal dimension of acting for reasons—acknowledging practical commitments by acting on them—involves the exercise of reliable differential responsive skills on the *output* side of the game of giving and asking for reasons, just as perception does on the *input* side. Elaborating the first idea involves examining the sense in which practical reasons are *reasons*; elaborating the second idea involves examining the sense in which practical reasons are *causes*.

Adding practical commitments to the model of discursive practice enriches the propositional contents that such practice can be understood to confer on states and their expressions in a way analogous to the enrichment provided by including the empirical authority of observationally acquired doxastic commitments. In each case the general category of assertional commitments and their contents can be understood in advance of the enrichment. The three structures of authority that the model of assertion, as presented thus far, comprises are mutually irreducible, but not all are equally fundamental. The inferential authority invoked by justification and the testimonial authority invoked by deference are intelligible apart from the default authority of noninferential reports; but inferential and deferential practice are two sides of one coin, apart from which the authority of noninferential reports is not intelligible. Thus *empirical* content represents an enrichment of the generic sort of propositional content specifiable in abstraction from the contribution of observation. Similarly, *practical* content represents an enrichment of the generic sort of propositional content specifiable in abstraction from the contribution of action. The empirical and practical involvements of claims—even those that are purely theoretical in the sense that they are only inferentially connected to claims that have direct empirical and practical significance—make a fundamental contribution to their contents. Only a model that incorporates both of these not purely inferential dimensions of discursive articulation has any prospect of generating propositional contents that resemble those expressed by the declarative sentences of natural languages.

The best way to understand the place of action in the deontic model of discursive practice is to exploit the analogy between action and perception. Sellars divides the “moves” that can be made in a language game into three kinds: intralinguistic moves, language entry moves, and language exit moves.<sup>33</sup> The first kind consists of inferential moves. These are moves in which a position within the language game (paradigmatically the endorsement of a claim) is responded to by the adoption of another such position.

The second kind consists of noninferential reports of observations. These are moves in which a nonlinguistic situation is responded to by the adoption of a position within the language game (paradigmatically the endorsement of a claim). The third kind consists of deliberate actions. These are moves in which a position within the language game (for instance, endorsement of a plan) is responded to by bringing about a nonlinguistic situation.

Following Sellars's lead, language entry moves have been analyzed in the first three sections of this chapter in terms of two components in their content: their inferential articulation and their noninferential elicitation. In virtue of the former they are *conceptually* contentful, and in virtue of the latter they are *empirically* contentful. These components and their interaction have been elaborated here in the idiom of deontic scorekeeping, into which Sellars's framework has been transposed. In that idiom, noninferential reports count as *entries* in the sense that they are responses that consist in changes of deontic scorekeeping attitude, elicited by stimuli that do not themselves consist in changes of deontic score. As such they contrast with inferential moves, in which an alteration of deontic attitude—for instance the undertaking or attributing of a commitment—has as a scorekeeping consequence another alteration of deontic attitude. The language or discursive scorekeeping *exits* (intentional *actions*) are to be understood by analogy to these *entries* (perceptual *observations*). In action, alterations of deontic attitude, specifically acknowledgments of practical commitments, serve as stimuli eliciting nonlinguistic performances.

Observation depends on reliable dispositions to respond differentially to states of affairs of various kinds by acknowledging certain sorts of commitments—that is, by adopting deontic attitudes and so changing the score. A competent observer under suitable circumstances responds to the visible presence of a red ball by coming to acknowledge a commitment to the claim that there is a red ball present. The content of the commitment responsively undertaken is jointly determined by the chain of reliably covarying events that culminates in its acquisition and by its inferential connection to other contents (including those empirical conceptual contents that themselves incorporate a responsive observational component). Action depends on reliable dispositions to respond differentially to the acknowledging of certain sorts of commitments (the adoption of deontic attitudes and consequent change of score) by bringing about various kinds of states of affairs.<sup>34</sup> A competent agent under suitable circumstances responds to the acquisition of a commitment to flip the light switch by flipping the light switch. The content of the commitment so expressed is jointly determined by the chain of reliably covarying events that its acknowledgment initiates and by its inferential connection to other contents (including both other contents that themselves incorporate a practical component and those empirical conceptual contents that incorporate a responsive observational component).

In any given situation, interlocutors can be taught to be reliable noninfer-

ential reporters of only certain sorts of circumstances. Unaided by special instruments, we cannot reliably discriminate the presence of X rays, and we cannot tell automobiles that will at some point in the future be painted green from those that will not; we can reliably discriminate the presence of loud noises, and we can tell automobiles that are now painted green from those that are not. Similarly, interlocutors can be taught to be reliable performers of only certain kinds of acts. Unaided by special tools, we cannot reliably produce X rays, and we cannot make an automobile have been painted green some time in the past; we can reliably produce loud noises and paint automobiles green. What can be noninferentially reported varies from reporter to reporter and from situation to situation. Only a properly trained physicist can noninferentially observe the presence of a mu-meson, and then only with a bubble chamber; only a properly trained pianist can noninferentially produce a performance of the Moonlight Sonata, and then only with a piano.

Observation requires reliable responsive dispositions to acquire acknowledged commitments, while action requires reliable responsive dispositions to fulfill acknowledged commitments. Reliability in the first case concerns the relation between the state of affairs responded to and the content of the commitment acknowledged. Reliability in the second case concerns the relation between the content of the commitment acknowledged and the state of affairs brought about. In each case, assessments of reliability require some independent access to the eliciting or the elicited state of affairs—assessments of the *truth* of the claim the perceiver has noninferentially come to make and of the *success* of the performance the agent has noninferentially come to produce. Attributions of reliability consist in endorsements of score-keeping inferences *from commitments attributed to reporters or agents to commitments undertaken by the attributor of reliability* (commitments concerning the state of affairs reported or produced). Thus my noninferentially acquired doxastic commitment to the effect that there is a red thing in front of me is, under appropriate conditions, a good reason for others inferentially to acquire a doxastic commitment to the effect that there is a red thing in front of me. My acknowledging a practical commitment to the effect that I will raise my arm in the next minute is, under appropriate conditions, a good reason for others to undertake a doxastic commitment to the effect that I will raise my arm in the next minute.

In observation, the elicited commitment-acknowledgment is an attitude toward a *doxastic* discursive deontic status. In action, the eliciting commitment-acknowledgment is an attitude toward a *practical* discursive deontic status. The first sort of attitude corresponds to believing or taking-true—in one sense of believing, namely the causally relevant sense that depends on what one would *acknowledge* commitment to, not the ideal sense in which if  $p$  entails  $q$ , then believing that  $p$  is believing that  $q$ , whether one knows it or not. The second sort of attitude corresponds to intending or making-true—in one sense of intending, namely the causally relevant sense that depends on what one would *acknowledge* commitment to, not the ideal sense in

which if doing *A* entails doing *B*, then intending to do *A* is intending to do *B*, whether one knows it or not.

The wider ideal senses of ‘believe’ and ‘intend’ correspond to the deontic *statuses* of doxastic and practical commitment, rather than to the deontic *attitudes* of acknowledging them. These senses are to be understood in terms of the fundamental scorekeeping principle that *undertaking* a commitment (to begin with, by acknowledging it) licenses others to *attribute* it, and the attributions that are thereby authorized can outrun what one is disposed to acknowledge. The only function of the concept *deontic status* in the idiom in which the model of discursive practice is formulated is its use in keeping score. It is a creature of the activity of scorekeeping on deontic attitudes; deontic statuses figure only as the objects of attitudes, as what is undertaken and attributed.

Understanding practical discursive commitments (commitments to act) is accordingly a matter of understanding their pragmatic significance: the way they depend on and influence the deontic score interlocutors keep by acquiring and relinquishing attitudes toward their own and others’ deontic statuses. Practical commitments, like doxastic or assertional commitments (including noninferentially acquired empirical ones), are discursive or conceptually contentful commitments in virtue of the *inferential* articulation of their pragmatic significance. The scorekeeping significance of practical commitments is analogous to that of doxastic commitments—indeed the inferential and incompatibility relations that the contents of practical commitments stand in are largely inherited from those of corresponding doxastic commitments, except for their role in the sort of practical reasoning that connects them inferentially with doxastic commitments proper. Thus one practical commitment can have others as consequences; a commitment to drive to the airport today entails a commitment to go to the airport today, because the inference from ‘*X* drives to *s*’ to ‘*X* goes to *s*’ preserves *doxastic* commitments. In the same way, one practical commitment can be incompatible with another, as are a commitment to drive to the airport today and a commitment to spend the day snoozing in a hammock under a shade tree—again because of the incompatibility of the corresponding doxastic commitment contents.

The instrumental inferences corresponding to the principle “Who wills the end wills the means,” like inferences generally, come in two flavors: committive and permissive. Some instrumental inferences (those whose premises specify goals one is committed to and whose conclusions specify the necessary means to those ends) are also commitment-preserving. If cutting down a tree is the only way to get across the ravine, then undertaking or attributing a commitment to getting across the ravine has as a scorekeeping consequence undertaking or attributing a commitment to cutting down a tree. But some means-end reasoning is permissive in nature; there may be more than one way to skin a cat.

Inferences whose premises express commitments to secure certain ends and whose conclusions express sufficient (but not necessary) means to those

ends are entitlement-preserving rather than commitment-preserving. One who is entitled to a practical commitment to secure an end is entitled thereby to a practical commitment to performances that would (according to the one whose scorekeeping is being elaborated) bring about that end—in the absence (as is always the proviso with permissive inferences) of collateral commitments incompatible with such a commitment. Entitlement to a practical commitment to achieve some end may simultaneously entitle one to each of a set of mutually incompatible alternative means; entitlement to a commitment to cross the ravine may instrumentally entitle one to cut down the tree at the edge of the ravine, and it may entitle one to anchor a rope bridge to the top of that tree, even though doing one of these things precludes doing the other. In the same way, permissive inferential relations (paradigmatically inductive ones) among the contents of doxastic commitments can result in entitlement to each of a set of incompatible conclusions. In each case, choosing one, committing oneself to a conclusion or a means, relinquishes entitlements to those incompatible with it. For incompatibility is a relation involving both deontic statuses: two contents, whether doxastic or practical, are incompatible in case commitment to one precludes entitlement to the other. Entitlement to both without commitment to either is not ruled out. Neither, of course, is commitment to both; making this possibility straightforwardly intelligible is one of the cardinal advantages deontic normative construals of belief have over causal-functional ones.

### *3. Asymmetries between Practical and Doxastic Discursive Commitments*

It is in their relations to their corresponding entitlements that practical discursive commitments differ most markedly from doxastic discursive commitments. The significance of undertaking a doxastic commitment, paradigmatically through its overt acknowledgment by assertion, was explained in terms of the interactions between the coordinate dimensions of authority and responsibility. The responsibility involved is to vindicate the commitment, by demonstrating or displaying one's entitlement to it, if it is brought into question by a suitable challenge (an incompatible assertion with an equal, *prima facie* claim to entitlement). Default entitlements aside, this responsibility can be discharged by appeal to the authority of other doxastic commitments; credentials for the commitment are secured by displaying its entitlement as inherited from that attached to other commitments.

The authority of the commitments undertaken by assertion exhibits a dual structure, corresponding to two different sorts of routes by which entitlement can be passed on for use in discharging the responsibility associated with other commitments. On the one hand, a doxastic commitment to which one interlocutor is entitled licenses further commitments (with different

contents) by that same interlocutor. These are its *inferential* consequences (committive and permissive). This sort of authority is invoked to vindicate those consequential commitments by presenting a *justification*, which appeals to the authorizing claims as premises. On the other hand, a doxastic commitment to which an interlocutor is entitled licenses further commitments with the same content, by other interlocutors. This is its authority as *testimony*. It is invoked to vindicate the commitments it authorizes, by *deferral* to the one whose testimony is relied upon.

The first way in which the structure governing the attribution of entitlements to practical discursive commitments differs from that governing the attribution of entitlements to doxastic ones is that there is nothing corresponding to the authority of testimony in the practical case. The issue of entitlement can arise for practical commitments, as for all discursive commitments. But the (conditional) responsibility to vindicate such commitments is, in the practical case, exclusively a *justificatory* responsibility. Default entitlements aside, it is only by exhibiting a piece of reasoning having as its conclusion the practical commitment in question that entitlement to such commitments can in general be demonstrated or secured.

This feature of the deontic scorekeeping model of discursive commitments reflects a fundamental asymmetry between expressing a belief by making a claim and expressing it by performing an action. What I take-true I thereby, *ceteris paribus*, authorize you to take-true. Though there can be various complications about the transfer of title (because of differences in collateral beliefs), in general what serve me as good reasons for belief can serve you also as good reasons for that same belief. What I (seek to) make-true, however, I do not thereby in general authorize you also to (seek to) make-true. What serve me as good reasons for action may or may not be available to you as good reasons for action, even bracketing differences in collateral beliefs. For you and I may have quite different ends, subscribe to different values, occupy different social roles, be subject to different norms. That I have good reasons to drive to the airport today does not imply that you do. If you form a similar intention, you cannot in general show that you are entitled to it by deferring to me ("Well, *he's* going"). Only some kinds of reasons that entitle me to an intention and action are automatically available to you. You might be in a position to make the same argument I can, but if so, that in general is independent of my being in a position to use that line of thought; there is no general (even defeasible) presumption of heritability.<sup>35</sup>

Committing oneself to a claim is putting it forward as *true*, and this means as something that everyone in some sense *ought* to believe (even though some unfortunates will for various reasons not be in a position to do so and need not be blameworthy for that failure). Committing oneself to a course of action need not be like this. It need not (though in special cases it can) involve putting it forward as something that everyone else ought to do (even subject to the recognition that some unfortunates will for various reasons not

be in a position to appreciate this, and need not be blameworthy for that failure). Some kinds of reasons for actions, paradigmatically moral ones, have a permissive or committive force that is independent of interpersonal differences. But reasons for action in general do not have this kind of force. What an agent has reason to do can depend on what that agent wants (or on what institutional role that agent occupies). Differences among agents as to desires and preferences (or institutional roles) need not have the significance of indications of normative failures. Whenever two believers disagree, a diagnosis of error or ignorance is appropriate for at least one of them. Though agents with differing practical commitments can also be criticized on the grounds of error and ignorance, mere difference of desire or preference is not sufficient in general to make them liable to such criticism.

We come with different bodies, and that by itself ensures that we will have different desires; what is good for my digestion may not be good for yours; my reason to avoid peppers need be no reason for you to avoid peppers. Our different bodies give us different perceptual perspectives on the world as well, but belief as taking-true incorporates an implicit norm of commonality—that we should pool our resources, attempt to overcome the error and ignorance that distinguish our different sets of doxastic commitments, and aim at a common set of beliefs that are equally good for all. Talk about belief as involving an implicit commitment to the Truth as One, the same for all believers, is a colorful way of talking about the role of testimony and challenge in the authority structure of doxastic commitment—about the way in which entitlements can be inherited by others and undercut by the incompatible commitments they become entitled to. The Good is not in the same way One, at least not if the focus is widened from specifically moral reasons for action to reasons for action generally, so as to include prudential and institutional goods. Desires and preferences can supply reasons for actions (can entitle agents to practical commitments) in the sense of ‘entitle’ that corresponds to that at stake in the discussion of doxastic commitments, and desires and preferences can vary from individual to individual. That there is no implicit normative commitment that plays the same role with respect to desire (and therefore intention and action in general) that truth plays with respect to belief consists simply in the absence (in the structure according to which entitlements to practical commitments are inherited) of anything corresponding to the interpersonal dimension of testimony and vindication by deferral.

It is of course possible to add an interpersonal dimension of practical authority as a superstructure to the basic game of giving and asking for reasons for actions. Where within a certain sphere of practical activity the performance of one individual licenses or compels performances by others, there exists an authority relation of superior to subordinate. In a practice in which reasons can be given (and so asked for) at all—that is, a linguistic practice, one in which some performances are accorded the significance of

assertions—the authorizing performances can be speech acts with the significance of imperatives and permissives. The superior issues an order, which specifies what the subordinate is obliged to do (what the subordinate thereby acquires a commitment to do), by displaying the assertion that must be made-true, the assertible content of the doxastic commitment that anyone must be entitled to undertake (perhaps observationally) upon completion of the task. Or the superior offers permission in the form of a licence, which specifies what the subordinate is entitled to do by displaying the assertion that can be made-true.

This sort of practical authority structure is like that of testimony in some ways. When the issue of the agent's entitlement to a practical commitment (perhaps claimed implicitly by deliberate action) is raised, rather than the entitlement being inherited from reasons that could be cited for it—either by the agent in terms of doxastic and practical commitments undertaken or by other scorekeeping assessors in terms of doxastic and practical commitments attributed to the agent—that entitlement can be inherited from the superior who ordered or permitted it. Such authority can be invoked by deferring to the issuer of the command or license (a mode of vindication codified in the legal doctrine of *respondeat superior*). So besides intrapersonal entitlement inheritance invoked by inference, there can be a mode of interpersonal entitlement inheritance invoked by deference, in the practical as well as the doxastic case.

There are many disanalogies between these two cases as well, however. First, the licensing is restricted as to subject matter and the interlocutors involved, to those situations in which a prior superior/subordinate authority relation has been established. The employer can authorize or compel only certain sorts of performances, and only on the part of certain individuals. Perhaps this difference does not go very deep. The limitation is characteristic of a society in which such authority relations are established and limited by explicit contracts. In a society based on status rather than contract, the superior/subordinate relations are fixed once and for all in advance, appearing as part of the nature of things, and need not be restricted as to subject matter at all. Furthermore, such restrictions can arise, *de facto* or even *de jure*, in the case of assertional authority as well: the teacher of secret doctrines may not authorize their repetition or answer for them except to favored students. And there are what amount to hierarchies of assertional authority regarding technical topics such as mu-mesons and quarter horses.

One difference that does go deep, however, is an asymmetry between the authorizing performance and the authorized performances, in the case of commands and the issuing of licenses. The asserter licenses members of the audience to perform speech acts with just the same content and significance as the original assertion. They are authorized to authorize others in the same sense in which they are authorized. Assertion, at least as it is construed in the ideal Sprachspiel presented here, is an egalitarian practice in a sense in

which commanding and giving permission is not. Only in very special cases does the practical license one is given authorize the further issuance of such licenses; only in very special cases does the command one is given compel or permit one to offer such commands to others.

The structure of default entitlement and calling to account by challenging entitlements to action are similarly asymmetric in the case of interpersonal practical authority structured by superiors and subordinates. There is in general nothing corresponding to assessments of reliability underlying the default authority of superiors (though analogs exist for special cases). Entitlement to challenges must similarly be relativized to superior/subordinate relations, if commands and licenses are to have any significance. The point of rehearsing these asymmetries is just that the fundamental differences between doxastic and practical structures of authority and entitlement inheritance remain even in the case where the normatively significant social status of individuals as superior or subordinate is widely or universally instituted by the attitudes of those keeping score on commitments and entitlements. If the asymmetries characteristic of superior/subordinate relations are removed, making interpersonal practical authority look more like assertional authority as here conceived, nothing remotely resembling the issuing of orders or the giving of permission results.

In the interests of simplicity, the deontic scorekeeping model of assertional significance has been talked about as though assertional authority is always made universally available throughout the community and is always universally recognized. Where testimony has this sort of catholic significance, the community can be thought of as engaged in the search for a single common body of truths, for anyone's entitlement to any claim is open to challenge from any quarter. Doxastic practice need not be so monolithic, of course. There may be many subcommunities, distinguished precisely by what sorts of authority they acknowledge, and so what sorts of challenges to entitlements they take to be in order. Specialists may recognize the authority only of other specialists. Members of one speech community may be divided into competing schools of thought on various topics and may not recognize the entitlements or therefore the challenges of those from other groups, as regards claims concerning those topics. Yet within those subcommunities it is essential that the authority granted by an assertion include a reassertion license—a license to do just what the asserter did. This feature makes it possible for the claims of one interlocutor to have the significance of challenges to the claims of another.

The importance of this structure is particularly evident in the case of empirical practice, for it is by testimony that observations by one interlocutor can be assessed and adjusted by confrontation with the observations of another. The notion of entitlement to a doxastic commitment depends on the in-principle heritability of interpersonal authority. Because an assertion that would be defended by appeal to testimony can have just the same

entitlement status as one that would be defended by providing a justification or by the invocation of noninferential responsive reliability (observational prowess), the credentials of each claim do not need to be traced back to their source before it can be treated as having the significance of a *prima facie* challenge to incompatible claims. This is a basic feature of the assertional default-and-challenge structure. That a status or performance whose entitlement is inherited from another should have just the same authority as the status or performance that authorized it (according to the subcommunity that recognizes such authority) accordingly distinguishes doxastic discursive commitments from practical discursive commitments. For as has been pointed out, if subordinates have the same authority as their superiors in virtue of being commanded or licensed by them, the entitlement of a superior to issue a command would be subject to challenge by commands issued by subordinates, not just by other superiors, and the hypothesized asymmetry between superior and subordinate would disappear.

## V. PRACTICAL REASONING: INFERENCES FROM DOXASTIC TO PRACTICAL COMMITMENTS

### *1. Acting for Reasons and Acting Intentionally*

Giving and asking for reasons for *actions* is possible only in the context of practices of giving and asking for reasons generally—that is, of practices of making and defending *claims*. The structure of entitlement characteristic of practical discursive commitments is not autonomous but presupposes that of doxastic ones. This dependence appears in two ways. On the side of the circumstances of acquisition of practical deontic statuses, it appears in the role of practical reasoning; practical reasoning requires the availability of doxastic commitments as premises. On the side of the consequences of acquisition of practical deontic statuses, it appears in the essential role that propositional (= assertible) contents play in specifying conditions of success—that is, what counts as fulfilling a commitment to act. With regard to this latter role, it has already been pointed out that practical commitments inherit some of their inferential relations from the propositional contents that specify their conditions of success. If doxastic commitment to *p* has as a scorekeeping consequence doxastic commitment to *q*, then a practical commitment to make-true *p* has as a scorekeeping consequence a practical commitment to make-true *q*. Understanding what one has committed oneself to by undertaking a practical commitment to bring it about that *p* accordingly requires mastery of the inferential role *p* plays in doxastic discursive practice.

The relation between doxastic and practical commitments that is most important for extending the deontic scorekeeping account to include both

species of discursive status, however, is that exhibited in practical reasoning. Because interpersonal inheritance of entitlements is not an essential part of the scorekeeping structure that institutes practical deontic discursive statuses, the conditional responsibility to demonstrate entitlement that is part of undertaking a commitment to act is a specifically *justificatory* responsibility. Justifying a practical commitment is exhibiting a suitable piece of practical reasoning in which it figures as the conclusion. It is in terms of practical inferences that we give reasons for action, make our own and each other's conduct intelligible, exhibit it as rational. Practical reasoning accordingly forms the core of intentional explanations of nonlinguistic deportment.

In what follows, an abstract account of practical reasoning is sketched in the deontic scorekeeping terms familiar from the treatment of theoretical reasoning concerning doxastic commitments.<sup>36</sup> The explanatory framework in which the notion of practical reasoning is to function is the Kantian one, according to which to treat a performance as an *action* is to treat it as something for which it is in principle appropriate to demand a *reason*. Not everything an agent does is an action. If I am walking along the top of a cliff and stumble and fall off, stumbling and accelerating at 32 feet per second per second are both things I do (in the sense that they are bits of my behavior), but they are not actions of mine; walking and grabbing a bush as I topple over the edge are. Actions are the things agents do *intentionally*. In the terms to be employed here, acting intentionally is noninferentially producing a performance that either is the acknowledgment of a practical commitment (in the case of intentions in action) or results from exercising a reliable differential disposition to respond to such an acknowledgment (in the case of prior intentions). The acknowledgment of the practical commitment can be thought of as the intention with which the performance is produced.

One can act *with* a reason, but unintentionally (for instance in a case in which one is unaware of the commitments that supply the reason that an attributor might cite). But only what is done intentionally can be done *for* a reason—though one can act intentionally but without a reason.<sup>37</sup> Only rational beings can be agents, but there are such things as irrational actions: for instance where one acts intentionally, but on impulse rather than according to what one has reason to do. In the deontic framework, such irrational actions are *intentional* in that they are acknowledgments of practical *commitments* (or arise from the exercise of reliable noninferential dispositions to respond differentially to them), and they are irrational in that the practical commitment in question is not one the agent is *entitled* to by a good practical inference from premises that agent is committed and entitled to—either because one has *no* reason or because one has an overriding reason to do something incompatible with what one in fact does. Since to be so entitled requires having a *reason* for performing the action, practical commitments, and therefore actions (intentional performances), are attributed only to those who are in the space of giving and asking for reasons—that is, to those who are (treated as) rational.

Undertaking any discursive commitment involves a conditional responsibility to demonstrate entitlement to it. In the case of practical commitments this takes the form of a specifically justificatory responsibility. Only against the background of a general capacity to comprehend and fulfill such a justificatory responsibility—to assess and produce reasons for practical commitments—can what one does have the significance of an acknowledgment of a practical commitment, that is, the significance of acquiring or expressing an intention. Given such a general capacity or status as rational, however, one can in particular cases undertake practical commitments to which one is not entitled by reasons, and so act irrationally. Intentional but irrational actions are perfectly intelligible within the deontic framework, in the same way and for the same reasons that, on the side of doxastic rather than practical discursive commitments, incompatible beliefs are—namely as commitments lacking the corresponding entitlements. These phenomena cause explanatory difficulties for other sorts of accounts (for instance those that construe intentional states exclusively in terms of causal-functional role), difficulties that simply do not arise when those states are construed in terms of deontic statuses instituted by scorekeeping attitudes.

To be entitled to a practical commitment is to have suitable reasons for it. Practical inferences—as distinct from the doxastic inferences that have been considered in previous chapters—are those that have practical commitments as their conclusions.<sup>38</sup> Reasons for such commitments, and hence for the actions elicited by the acknowledgment of such commitments, are the premises of good practical inferences. It has already been pointed out that intentions can serve as reasons for other intentions—the intention to bring it about that  $p$  serving as a reason for intending to bring it about that  $q$  if that  $q$  is true is necessary or sufficient for bringing it about that  $p$ . What about reasons for commitments to act that are not themselves commitments to act? Facts, as acknowledged in doxastic commitments, can provide reasons for practical commitments.

## 2. Three Patterns of Practical Reasoning

Consider the following three bits of practical reasoning:

- (α) Only opening my umbrella will keep me dry,  
so I shall open my umbrella.
- (β) I am a bank employee going to work,  
so I shall wear a necktie.
- (γ) Repeating the gossip would harm someone, to no purpose,  
so I shall not repeat the gossip.

'Shall' is used here to express the significance of the conclusion as the acknowledging of a practical commitment.<sup>39</sup> The corresponding doxastic commitment would be acknowledged by a standard assertion using 'will'.

The role of the speech acts performed by uttering sentences of this form can be understood in scorekeeping terms from their fundamental pragmatic significance as acknowledging a practical commitment, together with the inferentially articulated content that results from combining its involvement in the inferences deriving from the corresponding 'will' statements with its involvement in the sort of basic practical inferences of which  $(\alpha)$ ,  $(\beta)$ , and  $(\gamma)$  are examples.<sup>40</sup>

There are two ways to think about inferences like these, which move from doxastic premises to practical conclusions, from beliefs to intentions. What is perhaps the standard approach is that taken by Davidson.<sup>41</sup> He defines a *primary reason* as the pair of a belief and what he calls a pro-attitude. He allows that sometimes one or the other is cited by itself as a reason, but insists: "In order to understand how a reason of any kind rationalizes an action it is necessary and sufficient that we see, at least in essential outline, how to construct a primary reason."

In other words, inferences such as those exhibited by  $(\alpha)$ ,  $(\beta)$ , and  $(\gamma)$  are *enthymemes*, in which a premise necessary for the correctness of the inference has been suppressed or omitted. In the first inference, what is missing is some such premise as:

(a) Let me stay dry,

an expression of a desire, preference, or pro-attitude that would be explicitly self-ascribed by something like "I desire (prefer) to stay dry."<sup>42</sup> The second inference might be underwritten by something like:

(b) Bank employees are obliged (required) to wear neckties.

In the third case, the suppressed premise is something such as:

(c) It is wrong to (one ought not) harm anyone to no purpose.

Each of these supplies the missing pro-attitude required to make the premises into primary reasons.

As appears in these examples, the notion of pro-attitude encompasses not only wants, desires, and preferences but also more general evaluative attitudes. This assimilation represents an important insight, which will be exploited below. In fact, Davidson thinks that all pro-attitudes are expressed by sentences that are in a broad sense normative or evaluative.

There is no short proof that evaluative sentences express desires and other pro-attitudes in the way that the sentence "Snow is white" expresses the belief that snow is white. But the following considerations will perhaps help show what is involved. If someone who knows English says honestly "Snow is white," then he believes snow is white. If my thesis is correct, someone who says honestly "It is desirable that I stop smoking," has some pro-attitude towards his stopping smoking.

He feels some inclination to do it; in fact he will do it if nothing stands in his way, he knows how, and he has no contrary values or desires. Given this assumption, it is reasonable to generalize: if explicit value judgments represent pro-attitudes, all pro-attitudes may be expressed by value judgments that are at least implicit.<sup>43</sup>

Davidson thinks of evaluative expressions as expressing something like desires, but the connection can equally well be exploited in the other direction. For conversely, one who desires or prefers  $p$  to  $q$  (say desires that one eat pears rather than that one eat peaches) thereby attaches some value to  $p$  over  $q$ , takes  $p$  to be preferable to, or more desirable than,  $q$ . What is important is to see that normative claims and expressions of desire and preference are species of a genus defined by the role they play in completing primary reasons.

Pro-attitudes must be included in primary reasons, on this account, to bridge the gap between what one believes and what one decides to do. My preference to stay dry makes my belief that I can stay dry only by opening the umbrella relevant to the practical issue of whether to open the umbrella. The fact that bank employees are obliged to wear neckties makes my working at the bank relevant to the practical issue of whether to wear a necktie. And the negative value of causing pointless harm (the fact that it is wrong or that one ought not to do it) makes the consequence of gossiping relevant to the practical issue of whether to gossip. In the context of different pro-attitudes, those same beliefs would provide reasons for quite different intentions and actions.

### 3. Normative Vocabulary Makes Explicit Material Properties of Practical Reasoning

There is another way of construing the relation between  $(\alpha)$ ,  $(\beta)$ , and  $(\gamma)$ , on the one hand, and  $(a)$ ,  $(b)$ , and  $(c)$ , on the other hand. That relation could be modeled on the relation between *materially* good inferences and the *conditionals* whose addition as premises would turn them into *formally* (logically) good inferences. In that case (considered in Section IV of Chapter 2), it turned out to be a fruitful strategy to consider the apparently enthymematic inferences as in order just as they stood, and to treat the conditionals not as suppressed *premises* but as making *explicit* (expressing in the form of a claim) what is *implicit* in the endorsement of the inferences. Part of the payoff from considering things this way around is an understanding of the expressive role played by conditionals; they can be understood as making *inferential* commitments propositionally explicit (= assertible). What makes that approach possible is an account of proprieties of inference as deontic social statuses instituted by scorekeeping attitudes, so that commitment to a material propriety of inference can be understood in terms of what it is to

take or treat an inference as correct in keeping score by attributing and acknowledging assertional commitments.

There is no bar to understanding  $(\alpha)$ ,  $(\beta)$ , and  $(\gamma)$  as materially good inferences in this sense. The fact that endorsement of claims incompatible with  $(a)$ ,  $(b)$ , and  $(c)$  would void these inferences does not show that they function as suppressed premises—any more than the fact that endorsement of  $p \rightarrow \neg q$  would void the inference from  $p$  to  $q$  shows that the conditional  $p \rightarrow q$  is a suppressed premise in the material inference from  $p$  to  $q$ . The claims  $(a)$ ,  $(b)$ , and  $(c)$  might, like  $p \rightarrow q$ , be understood rather as codifying material-practical inferential commitments. The payoff from doing so would be making it possible to understand the expressive role played by the broadly evaluative words (such as ‘prefer’, ‘obliged’, and ‘ought’) used to express these pro-attitudes, in a way analogous to the understanding suggested for conditionals. According to such an account, although the sort of practical inference instanced by  $(\alpha)$ ,  $(\beta)$ , and  $(\gamma)$  does not need supplementation to be correct, in a language with sufficient expressive resources it is possible to make the inferential commitments that are implicit in endorsing such inferences explicit in the form of claims.

The benefits of doing so are the familiar benefits of propositional explicitness: once expressed in the form of claims, these commitments are themselves subject to challenge and justification, rather than simply being accepted or rejected. Two interlocutors who disagree about the correctness of an inference such as  $(\gamma)$  can now argue about whether  $(c)$  is true, challenge entitlement to such a claim, and offer counterclaims to it. A new venue is opened up for resolving disagreements about what follows from what, about which claims rationalize which actions. Davidson’s view (transposed into the deontic idiom) that in order to see how a doxastic commitment can rationalize a practical commitment we must be able to see (“at least in essential outline”) how to construct a primary reason is correct in the following sense.

Once the expressive resources provided by terms such as ‘prefer’, ‘obliged’, and ‘ought’ are available, it must be possible to use them to make explicit the implicit practical inferential commitment underlying bits of practical reasoning such as  $(\alpha)$ ,  $(\beta)$ , and  $(\gamma)$ . But there is nothing incoherent about an idiom that lacks those expressive resources. Practical reasoning can still take place in it, and there is still a perfectly serviceable distinction between good and bad inferences available within such an idiom. It is by comparing the more primitive practices of giving and asking for reasons for action to the sophisticated ones made possible by the introduction of inference-explicitating locutions such as ‘ought’ that we can understand (in terms of deontic scorekeeping) the expressive role those locutions play.

The broadly normative or evaluative vocabulary that Davidson understands as expressing the pro-attitudes needed to turn the incomplete reasons offered as premises in  $(\alpha)$ ,  $(\beta)$ , and  $(\gamma)$  into complete reasons is actually used to make explicit in assertible, propositional form the endorsement of a *pat-*

tern of inferences. Different patterns of inference correspond to different sorts of norms or pro-attitudes. Consider someone assessing the propriety of the practical inference in  $\{\alpha\}$ , in a primitive idiom that does not yet permit the formulation of  $\{a\}$ . Suppose that the scorekeeper who assesses the practical reasoning attributes to someone commitment to the premise of  $\{\alpha\}$ , and also entitlement to that commitment. The question is whether entitlement to the doxastic commitment serving as the premise is inferentially heritable by the practical commitment serving as the conclusion.

To take it that it is, for a particular interlocutor, just is implicitly to attribute a desire or preference for staying dry. If the inferential commitment that underwrites this piece of practical reasoning is as expressed by  $\{a\}$ , then  $\{\alpha\}$  is just one of a whole family of inferences that stand or fall together. For instance, an attributor who takes  $\{\alpha\}$  to be entitlement-preserving will also take the following two inferences and a host of similar ones to have that status.

- ( $\alpha'$ ) Only standing under the awning will keep me dry,  
so I shall stand under the awning.
- ( $\alpha''$ ) Only remaining in the car will keep me dry,  
so I shall remain in the car.

To attribute a preference for staying dry to an individual is just to take inferences of this form to be entitlement-preserving, for that individual.

#### *4. Varieties of Prima Facie Reasons for Action*

Of course there can be competing entitlement-preserving inferences, corresponding to other desires. For recall that permissive inferences generally, whether doxastic or practical, can entitle one to incompatible conclusions<sup>44</sup>—though once an interlocutor endorses one of them, the undertaking of that commitment removes any entitlements that may hitherto have been available for competing claims. So endorsing this pattern of inferences as entitlement-preserving for an individual—which is implicitly attributing the preference that one could explicitly attribute either by attributing commitment to  $\{a\}$  or by undertaking commitment to the ascriptional claim “*A* desires to stay dry”—does not require attributing to that individual the practical commitment expressed by the conclusion in case commitment to a premise of the proper form is attributed. This is another way of saying that even in the presence of the desire, the belief need not lead to the formation of an intention, for there may be competing desires or other considerations in play. The notion of entitlement-preserving inferences accordingly provides a pragmatic analysis, in deontic scorekeeping terms, of the notion of *prima facie reasons* (whether doxastic or practical).

That a scorekeeper treats inferences of the form common to  $\{\alpha\}$ ,  $\{\alpha'\}$ ,

( $\alpha''$ ) . . . to be entitlement-preserving for interlocutor  $A$  does not involve any commitment to treating them as entitlement-preserving for interlocutor  $B$ , even apart from any consideration of the attribution of incompatible commitments (doxastic or practical, or either species of inferential). Treating these inferences as permissively good for  $A$  but not for  $B$  is just what attributing the relevant preference to  $A$  but not to  $B$  consists in. This is not how endorsement of doxastic inferences (even permissive ones) works. Endorsing a doxastic inference (one whose premises and conclusions are claims, that is, expressions of possible beliefs), treating that inference as entitlement- or commitment-preserving for *one* interlocutor, involves treating it as good for *all* interlocutors—subject, as always, to disqualification by commitment to incompatible claims, and with the proviso that differences in collateral doxastic commitments can make a difference in what premises are available as auxiliary hypotheses in such inferences.

This difference in generality is a fundamental difference between doxastic inference and this sort of practical reasoning. Desire is multifarious and different from individual to individual, but truth is one; so, according to each scorekeeper, the inferences that can be described unofficially as good in the sense of truth-preserving are one, while those practical inferences that are underwritten by desires are many. Of course, beliefs may differ from individual to individual as much as desires, and with it the endorsement of inferences whose propriety is underwritten by particular doxastic commitments, although the social institution of the status of objective information by the interpersonal dimension of assertional authority and the justificatory responsibility to respond to challenges incorporates an implicit norm of common belief that has no analog for desire. The difference being pointed to here is rather that attributions of *conative* commitments are construed here as fundamentally a kind of inferential commitment, linking doxastic and practical commitments, while *cognitive* or doxastic commitments and practical commitments are inferentially articulated and inferentially significant, but not themselves inferential commitments. That (in informal terms) desires vary from individual to individual, as beliefs do, is accordingly reflected in a structure of inferential commitments in the conative case that differs from that of the cognitive case.

Permissive proprieties of practical reasoning, endorsement of which is implicitly attributing—or in the reasoner's own case, acknowledging (which is self-attributing)—desires or preferences (pro-attitudes in a strict sense, as represented by the example of ( $\alpha$ )), are, however, only one species. Those represented by the example of ( $\beta$ ) need not be understood as having anything in particular to do with desires or preferences. The norm, rule, or requirement that bank employees wear neckties is what makes going to work into a reason for wearing a necktie, for bank employees. Taking it that there is such a norm or requirement just is endorsing a pattern of practical reasoning—namely, taking ( $\beta$ ) to be an entitlement-preserving inference for anyone

who is a bank employee. This inferential pattern is different from that exhibited by  $(\alpha)$  in two ways. First, there need not be for each interlocutor for whom  $(\beta)$  is taken to be a good inference a set of other inferences corresponding to  $(\alpha)$ ,  $(\alpha')$ ,  $(\alpha'')$ . Second, the scorekeeper will take  $(\beta)$  to be a good inference for any interlocutor  $A$  such that the scorekeeper *undertakes* doxastic commitment to the claim that  $A$  is a bank employee. Thus the way in which the scorekeeper's endorsement of the inference  $(\beta)$  as permissively good is distributed across various interlocutors who might reason that this way is different from that of  $(\alpha)$ .

Here the norm implicitly underwriting the inference is associated with having a certain status, as employee of a bank, rather than with exhibiting a certain desire or preference. Whether one has a good reason to wear a necktie just depends on whether or not one occupies the status in question. This pattern—where what matters is the scorekeeper's undertaking of a commitment to  $A$ 's occupying the status, rather than  $A$ 's acknowledgment of that commitment—corresponds to an objective sense of 'good reason for action' (according to the scorekeeper). In this sense, that  $A$  is preparing to go to work can be a good reason for  $A$  to wear a necktie, even though  $A$  is not in a position to appreciate it as such. The scorekeeper might take it that  $A$  is entitled to a practical commitment to wear a necktie, even though  $A$  could not justify it by producing the reasoning in  $(\beta)$ . In wearing a necktie,  $A$  would be acting *with* a reason, even if not *for* a reason. This corresponds to taking a reliable noninferential reporter to be entitled by that reliability to various observations, even in the case where the reporter is not in a position to appeal to that reliability in justifying those claims.

In another sense, of course, for the norm that the scorekeeper takes to be in force to supply a reason *for A*, the claim that  $A$  is a bank employee must also be acknowledged by  $A$ . For  $A$  to be able to justify a commitment to wearing a necktie by rehearsing the reasoning of  $(\beta)$ ,  $A$  must also endorse the pattern of inference codified explicitly in  $(b)$ . For a scorekeeper to take  $A$  to have a good practical reason in this stronger sense—that not just the scorekeeper, but  $A$  could produce it—requires that the scorekeeper attribute to  $A$  endorsement of an *inference*. In the model as presented thus far, this can be done only by attributing commitment to a claim codifying that inference. Depending on the expressive resources available, this might either be  $(b)$  or a set of corresponding conditionals.

One final stronger sense of *reason for A* is sometimes invoked by philosophers who insist that even  $(b)$  together with  $A$ 's acknowledgment of being a bank employee fall short of providing one unless supplemented by  $A$ 's desire to do what is required as a bank employee. It is indeed always possible, by supplying "suppressed" premises as needed, to assimilate all practical reasonings to the form of  $(\alpha) + (\alpha)$  (assimilating them to belief-desire reasons), so that norms and evaluations appear only in the role of objects of preference, as *staying dry* does in  $(\alpha) + (\alpha)$ . And it remains true that the role of  $(\beta)$  would

be quite different if (b) were conjoined with some claim *incompatible* with attributing to *A* such a desire. But this is just another instance of the phenomenon mentioned above as motivating but not warranting an enthymematic view. The inference from *p* to *q* is also undercut by the *denial* of the conditional *p* → *q*, but this does not show that the conditional is a suppressed premise in the original argument. The present approach requires no such instrumental reductionism, however, for it is possible to say what it is for scorekeepers to *treat* various other (from this point of view truncated) forms of practical reasoning *as* correct, as entitling agents to their practical commitments. Doing so makes it possible to explain how various normative vocabulary works (what it expresses)—including the vocabulary of commitment and entitlement that is employed in laying out the deontic scorekeeping model of discursive practice.

Endorsement of practical reasoning of the sort of which  $\{\gamma\}$  is representative, codified in the form of a normative principle by (c), corresponds to an inferential commitment exhibiting a pattern different from those involved in either (a) or (b). For a scorekeeper who takes  $\{\gamma\}$  to be entitlement-preserving for *A* takes it to be entitlement-preserving for *anyone*—regardless of desires or preferences and regardless of social status. Inferential commitments displaying this pattern are made explicit by *unconditional* ‘ought’s, whereas those displaying the other two patterns are made explicit by *prudential* ‘ought’s (in the case of (a)) and *institutional* ‘ought’s (in the case of (b)). Unconditional ‘ought’s, which correspond to this agent- and status-blind pattern of endorsement of practical inferences as entitlement-preserving, are one candidate that has been proposed as a good thing to mean by “moral ‘ought’.” Some thinkers insist rather that to treat reasons as moral reasons requires treating them as overriding; this amounts to saying that the ‘ought’ in (c) is a moral ‘ought’ only if  $\{\gamma\}$  is not only entitlement-preserving but also *commitment*-preserving—that anyone committed to the doxastic premises is thereby committed to the practical conclusion. It is not the point of this discussion to take a stand on how to distinguish specifically moral norms. Nor is the point to try to provide an exhaustive catalog of the sorts of norms (or “pro-attitudes,” in the broad sense) there can be. The point is just to show that various important sorts of norms (or pro-attitudes) can sensibly be thought of in deontic scorekeeping terms as corresponding to different patterns of endorsement of practical *inferences*.

To endorse a practical inference as entitlement-preserving is to take the doxastic premises as providing reasons for the practical conclusion. To exhibit a piece of good practical reasoning whose conclusion is a certain intention is to exhibit that intention, and the action (if any) that it elicits, as *rational*—that is, as reasonable in the light of the facts cited and the commitments exhibited in the premises. So *all* of the ‘ought’s that make explicit species of practical reasoning taken as examples here (the prudential or preferential ‘ought’, the social or institutional ‘ought’, and the unconditional

'ought') are different kinds of *rational* 'ought'. Being rational is just being in the space of giving and asking for reasons, and being a rational agent is being in the space of giving and asking for reasons for what one does. When the proprieties of practical inference that articulate that space are made explicit in the form of claims, they take the form of norms—of rational 'ought's. Rationality is the genus to which all these species of 'ought's belong.

There is no a priori reason to identify the rational with some one of the species of practical reasoning (for instance the prudential) and cut and paste the rest into suitable shape to be assimilated to the favored one. Being a reason is to be understood in the first instance in terms of what it is for a community to *treat* something in practice as such a reason, on the practical side of reasons for action just as on the doxastic side of reasons for claims. In neither case is this approach to normative status (what one is *really* entitled or committed to) through normative attitude (what one is *taken* to be entitled or committed to) incompatible with making eventual sense of *objective* norms, which underwrite the possibility that everyone's attitudes toward them are wrong. But understanding what is meant by such objective proprieties—what is *really* a good reason, as opposed to just what is *treated* as one—comes at the end of the story.<sup>45</sup> It is not something that can be understood a priori and imposed as a constraint at the outset.

## VI. INTENTIONS

### 1. *Reasons and Entitlement to Practical Commitments*

Exhibiting a piece of practical reasoning *rationalizes* the practical commitment or intention that is its conclusion. It displays reasons for that intention, offers a rational justification for it, shows how one might become rationally entitled to it. Accepting a practical inference as entitling someone to a practical commitment in this sense requires endorsing the *inference* as permissively good (and so only as providing a *prima facie* case for commitment to the conclusion, defeasible by incompatible commitments) for the agent whose conduct is being assessed. It does not require that the inference be accepted as one that would be (permissively) good in the scorekeeper's own case; the scorekeeper need not share the desire, preference, or institutional status that is implicitly attributed by treating some practical inferences as good for some agents. Nor does it require that in all cases the scorekeeper assessing that entitlement endorse the *premises*; a requirement of that sort picks out the special sense of objective entitlement. As Davidson says about prudential or preferential practical reasoning (the only kind he acknowledges): "When we talk of reasons in this way, we do not require that the reasons be good ones. We learn something about a man's reasons for starting a war when we learn that he did it with the intention of ending all wars [for Davidson this is equivalent to 'because he desired to end all wars'],

even if we know that his belief that starting a war would end all wars was false. Similarly, a desire to humiliate an acquaintance may be someone's reason for cutting him at a party though an observer might, in a more normative vein, think that that was no reason. The falsity of a belief, or the patent wrongness of a value or desire, does not disqualify the belief or desire from providing an explanatory reason.<sup>46</sup>

An agent can fulfill the justificatory responsibility involved in undertaking a practical commitment by exhibiting a piece of practical reasoning in which a commitment with that content serves as the conclusion. In the doxastic case, what is an entitling justification for one is an entitling justification for all, except for disqualifications due to commitment to claims incompatible with the premises or the conclusion. In the practical case, entitling justifications need not be portable across agents in this way. Displaying an intention as the conclusion of a piece of practical reasoning that is good in this sense makes it intelligible by showing reasons that could entitle the agent to it. To secure an attribution of entitlement to a practical commitment in this way, the practical reasoning in question may be offered as a justification by the agent, perhaps upon being challenged to do so. But it also may be attributed by the scorekeeper, who constructs the practical argument from premises already attributed to that agent, according to patterns of practical inference the scorekeeper endorses for that agent. These might be patterns of inference the scorekeeper endorses for everyone (unconditional 'ought's), or endorses for the agent on the basis of status (institutional 'ought's), or endorses only for the agent, thereby implicitly attributing idiosyncratic desires or preferences (prudential 'ought's), or of some other kind.

A scorekeeper who in this way takes an agent to be entitled to a practical commitment on the basis of its being the conclusion of a practical inference taken to be good for the agent, and who also attributes to that agent commitment to the premises of that inference, need not take it that acknowledgment of the practical commitment actually arose as the result of a process of inference by the agent from acknowledgment of those premises. As Davidson says "We cannot suppose that whenever an agent acts intentionally, he goes through a process of deliberation or reasoning, marshals evidence, and draws conclusions. Nevertheless, if someone acts with an intention, he must have attitudes and beliefs from which, had he been aware of them at the time, he could have reasoned that his action was desirable (or had some other positive attribute)."<sup>47</sup> For Davidson, acting intentionally and acting for reasons are the same thing. From the present point of view, this position involves conflating the two deontic statuses of practical commitment and entitlement to such a commitment. An act is intentional if it is (or is, as the exercise of a reliable differential responsive disposition, noninferentially elicited by) the acknowledgment of a practical commitment. To act for reasons is to be *entitled* to that practical commitment. One can in particular cases act intentionally but without reasons, even though there is no making sense of inten-

tions apart from their liability to the demand for reasons. But Davidson's point survives this confusion. An intention can be rendered intelligible as rational (a practical commitment can be displayed as one the agent is entitled to) by displaying reasons for it (premises from which it could legitimately have resulted as the conclusion of a good practical inference), even in cases where it was not in fact arrived at by such a process.

Contemporary thought about action begins with Anscombe's insight, developed with great force and clarity by Davidson, that the difference between actions and other performances—the answer to Wittgenstein's challenge to explain the difference between my raising my arm and my arm going up—is that actions are performances that are intentional under some description.<sup>48</sup> Any performance can be specified in many ways: Davidson moves his finger, flips the switch, turns on the light, alerts the burglar, causes a short-circuit in the wiring, starts a fire. These are all things he *does*; the different descriptions are different ways of specifying one action he performs. Not all of these are specifications under which what he does is intentional. But they are all specifications of an action, so long as what he does is intentional under *some* specification, for instance turning on the lights. In Davidson's slogan, being an action is an *extensional* property of an event (whether a given event is an action or not is not sensitive to how the event is specified), while being intentional is an *intensional* property of an event (whether a given event is intentional or not is sensitive to how the event is described). The very same event is intentional as turning on the lights but unintentional as alerting the burglar, causing a short-circuit, and starting a fire. The extensional property of performances, *being an action*, is defined in terms of the intensional property of performances of *being intentional* by existential quantification over descriptions or specifications of the performance: if it is intentional under any one of them, it is an action under all of them.

By this strategy the problem of explaining what privileges some (but not all) of an agent's performances as actions is reduced to the problem of explaining what privileges some (but not all) descriptions or specifications of an action as ones under which a performance is intentional. Davidson's solution to this problem in turn is that a performance is intentional under a description if that description figures as the conclusion of a piece of practical reasoning that exhibits the agent's reasons for producing it. These two moves together—the account of actions as performances that are intentional under some description, and the account of performances as intentional under just the descriptions that appear as the conclusions of practical inferences that rationalize those performances by giving reasons for them—reduce the problem of explaining what is special about action to that of explaining the giving of practical reasons. Davidson's account of primary reasons as pairs of beliefs and pro-attitudes is then offered to explain what it is for reasons to rationalize a performance according to a practical inference.

The Davidsonian explanatory structure provides a recipe, then, for turning

an account of practical reasoning into an account of action. In such a context, the present account of practical reasoning in terms of deontic scorekeeping on inferentially articulated practical commitments and entitlements to such commitments has some advantages over the one Davidson himself endorses. As originally presented, Davidson's theory eschews intentions entirely, in favor of beliefs and desires: "The expression 'the intention with which James went to church' has the outward form of a description, but in fact it is syncategorematic and cannot be taken to refer to an entity, state, disposition, or event. Its function in this context is to generate new descriptions of actions in terms of reasons; thus 'James went to church with the intention of pleasing his mother' yields a new, and fuller, description of the action described in 'James went to church'."<sup>49</sup> The account offered here, by contrast, explains deontic statuses corresponding to beliefs and intentions and defines those corresponding to desires (and other <sup>s</sup>expressions of pro-attitudes<sup>s</sup>) in terms of them.

Acknowledging intendings as full-fledged intentional states (or attitudes toward deontic statuses) avoids at the outset a difficulty that forced Davidson to modify his earlier account. For there are cases where someone has an intention (not just a reason for action), but no action arises from it. The possibility of intention without action is a symptom of the limited scope of Davidson's original discussion. As he puts the point in his introduction to the collection of his essays on this topic: "When I wrote ['Actions, Reasons, and Causes'] I believed that of the three main uses of the concept of intention distinguished by Anscombe (acting with an intention, acting intentionally, and intending to act), the first was most basic. Acting intentionally, I argued . . . was just acting with some intention. That left intending, which I somehow thought would be simple to understand in terms of the others. I was wrong . . . Contrary to my original view, it came to seem the basic notion on which the others depend; and what progress I made on it partially undermined an important theme of ['Actions, Reasons, and Causes']—that 'the intention with which the action was done' does not refer to an entity or state of any kind."<sup>50</sup>

## 2. Two Sorts of Intention

Explaining intentional action requires only what Searle calls *intentions in action*. Explaining pure intending requires also what he calls *prior* intentions. The distinction is motivated by the fact that "I can do something intentionally without having formed a prior intention to do it, and I can have a prior intention to do something and yet not act on that intention."<sup>51</sup> Pure intendings are special cases of prior intentions. In the deontic idiom, both sorts of intentions are (acknowledgments of) practical commitments—that is, commitments to act. Cases of intentions in action without prior intentions are those in which the performance that is accorded by a scorekeeper

the significance of an acknowledgment of a practical commitment is just the action itself. Consider Searle's example: "Suppose I am sitting in a chair reflecting on a philosophical problem, and I suddenly get up and start pacing about the room. My getting up and pacing about are clearly intentional actions, but in order to do them I do not need to form an intention to do them prior to doing them. I don't in any sense have to have a plan to get up and pace about. Like many of the things one does, I just do these actions; I just act."

In scorekeeping terms, undertaking a commitment is doing anything that makes it appropriate for it to be attributed. This may involve a distinct explicit acknowledgment, as in the case of asserting; it may be consequential, as in undertaking commitment to the consequences of a claim that is asserted; or it may be a default matter. Scorekeepers will take one not only to be entitled but to be committed to the claims that orange is a color and that there have been black dogs, unless these have been overtly disavowed (and even this may not be sufficient, if collateral commitments that are not disavowed entail these claims). The sort of case Searle considers concerns actions that have the default significance of intentional actions, as ones accompanied by a commitment for which the question of entitlement by reasons is in principle in order. It might turn out that what is, when described in a suitably impoverished vocabulary of motions of limbs, exactly similar behavior is not intentional at all but automatic, involuntary, compulsive behavior, triggered ineluctably by pheromones. In that case the default attribution of a practical commitment would be defeated. But the undertaking or acknowledging of a commitment to act need not be a performance distinguishable from the act one is committed to perform.

Nonetheless, in many cases, it is a separately datable event. In such cases of prior intention, acknowledging the commitment antedates fulfilling it (or not, as the case may be). I can now acknowledge a commitment to get on the bus *when* it arrives. My mastery of the practical conceptual (because inferentially articulated) content of that commitment includes my mastery of the *noninferential* differential disposition to respond to it and the joint fulfillment of its condition (the bus arriving) by getting on the bus. When the bus arrives, the condition is fulfilled. My getting on the bus is an acknowledgment of a practical commitment to get on the bus *now*—an intention in action. If I have the reliable noninferential differential disposition to respond to the acknowledgment of a practical commitment to do A when (if) C by doing A when (if) C, then my prior intention to get on the bus when it arrives will *mature* into a corresponding intention in action (marked above by the 'now' in the linguistic expression of the intention).

Prior intentions must specify the actions one is committing oneself to perform in *general* terms; they would be expressed by statements of the form

I shall get on the bus when it arrives,

in which the action is specified by a description. Intentions in action can be thought of as specifying the action one is committing oneself to perform in *demonstrative* terms; they could be expressed by statements of the form

I shall now do *this*.

Only an intention in action can be directed at a particular unrepeatable action. The process by which a prior intention ripens into an intention in action—the exercise of the practical skill of reliably responding to the undertaking of a commitment to bring about a state of affairs of such and such a description when such and such conditions obtain by doing so—involves recognizing when doing *this now* would fulfill such a commitment. (See further discussion below in 8.5.2.) What are here called “prior intentions,” Sellars calls simply “intentions,” and what are here called “intentions in action,” Sellars calls “volitions”:

A simple case of the relation of intending to volition can be illustrated by considering Jones, who has formed the intention of raising his hand in ten minutes. Suppose that no alternative course of action recommends itself to him. Then we may picture the situation as follows:

I shall raise my hand in *ten* minutes.

...

I shall raise my hand in *nine* minutes.

...

I shall raise my hand *now*.

(which culminates in action, if Jones happens not to be paralyzed).<sup>52</sup>

In Sellars’s idiom, a volition is an intention whose time has come.

Sellars takes the capacity respond reliably to prior intentions whose time has come by the formation of intentions in action (here, acknowledgments in the form of suitable performances) to be part of grasping the meaning of what is expressed by ‘shall’, and of the practical content of the particular concepts that articulate the content of those intentions. As in the account endorsed here, he takes these capacities to be part of the “important similarity between learning to make the language-entry transition of responding to presented red objects by saying ‘This is red,’ and learning the language-departure transition” involved in exercising those capacities.<sup>53</sup> Just as in the case of language entries or noninferentially elicited but inferentially articulated doxastic commitments, the existence of reliable differential responsive dispositions is compatible with making mistakes, so in the case of language exits or performances noninferentially elicited by inferentially articulated practical commitments, the existence of reliable differential responsive dispositions is compatible with failure. Mistakes of observation are diagnosed by scorekeepers by comparison of the contents of the doxastic commitments attributed to the observer and those undertaken (whether noninferentially,

inferentially, or by testimony) by the scorekeeper—for instance when the scorekeeper attributes a commitment to the claim that the ball is red and undertakes a commitment incompatible with the claim that the ball is red. Failures of action are diagnosed by scorekeepers by comparison of the contents of the practical commitments *attributed* to the agent and the contents of the doxastic commitments *undertaken* by the scorekeeper.

For instance the scorekeeper who attributes a commitment that could be expressed (Sellars-wise) as

Shall [The basketball goes through the hoop]

may also be obliged (perhaps observationally) to acknowledge a separate commitment incompatible with the claim that the basketball goes through the hoop (for example that expressed by “The ball missed the hoop”). A practical commitment may also remain a “pure” intending, eventuating in no action, successful or unsuccessful, for one of two reasons. It might be that the condition of maturation of a prior intention into an intention in action never is satisfied—the bus never arrives, the commitment expires. Or it may be that the commitment is relinquished, perhaps in favor of one with an incompatible content—the agent undergoes a change of mind.

### 3. Acknowledgments of Commitments Can Cause and Be Caused

When prior intentions are made explicit, they specify in general descriptive terms the performance the agent is committed to produce. When intentions in action, which are implicit in suitable performances, are made explicit, they specify those performances demonstratively. The way in which prior intentions elicit suitable performances, and so intentions in action, is a causal process. The analogy between action and observation (between language-exit transitions and language-entry transitions) is intended to illuminate the nature of the process involved.

Davidson's original essay famously endorses the claim that reasons are causes. The difference between a commitment's being *a reason* for an agent's action and its being *the reason* for that action must be explained in terms of differences in the causal roles played by various states. Primary reasons, conceived of as pairs of a set of beliefs and a set of pro-attitudes, rationalize actions (which accordingly count as intentional) first by providing reasons for them and second by serving to bring them about (“in the right way”). The account offered there has been criticized here for running together the notion of being *committed* to act in a certain way and being *entitled* to do so by reasons; even though the first deontic status cannot be made sense of apart from the second (any more than doxastic commitments can be made sense of apart from practices of giving and asking for reasons entitling interlocutors

to them), nonetheless these are distinct statuses, and one can be committed without being entitled, can act intentionally without having reasons for doing so. This distinction between practical commitments and entitlements to such commitments also, it has been claimed, opens up a space for the notion of *prior intention* (besides that of intention in action), of which pure or unconsummated intendings are a species.

What becomes of the doctrine that reasons are causes, when intentional states are construed in terms of social scorekeeping on deontic statuses and the deontic statuses of commitment and entitlement are appropriately distinguished? The account of the social practices that institute deontic statuses appeals to such statuses only as scorekeeping devices. The significance of being committed to a certain claim or assertible content is normative. It has to do with what else one is committed or entitled to. It is articulated by *proprieties* of scorekeeping and consists of the *proper* antecedents and consequences of that status. In the same way, that there are two strikes on a batter is a status properly acquired by various performances (just which depending on the antecedent score), a status that alters the significance various further performances have for the subsequent score.

Any effect that such elements of the score have on what performances are *actually* produced is indirect, mediated by the *attitudes* of those who keep score. The score determines only what *ought* to be done, what would be *proper*. What ought to be done and what is proper affect what players do only insofar as they are trained to respond in various ways to *taking* a certain course of action to be proper. The only access that deontic statuses have to the causal order is through the deontic attitudes of the scorekeeping practitioners.

Inferential relations among propositional contents are a matter of normative relations among deontic statuses: commitment to the claim that lions are mammals entails commitment to the claim that lions are vertebrates. Inferring, by contrast, is a causal process that relates deontic *attitudes*: *acknowledging* (and equally, attributing to someone else) commitment to the claim that lions are mammals will, under various circumstances and in those well versed practically in the inferential relations among deontic statuses, have as a causal consequence acknowledging (or, correspondingly, attributing) commitment to the claim that lions are vertebrates. Unless the members of a linguistic community are pretty good at keeping score by altering their attitudes as they ought to according to the contents associated with the deontic statuses in terms of which they keep score, there is no point in interpreting them as engaging in the practices specified by those proprieties of scorekeeping. Nonetheless, normative status is one thing, the attitudes of attributing and undertaking those statuses, the alteration of which is what scorekeeping consists in, is another.

As it is with the inferential articulation of the conceptual contents conferred on states, attitudes, performances, and expressions by deontic score-

keeping practices, so it is with their noninferential involvements, which confer empirical and practical conceptual contents on them. What observable states of affairs causally elicit in perception, according to reliable differential responsive dispositions, is in the first instance deontic *attitudes* rather than statuses: *acknowledgments* of doxastic commitments. What in action causally elicits the production of performable states of affairs (by the exercise of reliable differential responsive dispositions) is in the first instance deontic *attitudes* rather than statuses: *acknowledgments* of practical commitments. ("In the first instance" because acknowledging a commitment is one way of undertaking one, so those deontic attitudes have scorekeeping consequences for the deontic statuses of those whose attitudes they are.) That a particular doxastic commitment was elicited by the exercise of such dispositions is another way of putting the condition on perception that the belief not only be caused by the state of affairs reported but be caused by it "in the right way." That a particular practical commitment elicits a performance by the exercise of such dispositions is another way of putting the condition on action that the performance not only be caused by the intention but be caused by it "in the right way." Mastering the two sorts of reliable differential responsive dispositions connecting noninferentially acquired acknowledgments of doxastic commitments to their appropriate causal antecedents and noninferentially efficacious acknowledgments of practical commitments to their appropriate causal consequents is part of grasping, in one's scorekeeping practice, the empirical and practical components of the contents of concepts employed in observation and action, and of those theoretical concepts inferentially related to these.

Thus just as 'belief' is ambiguous in scorekeeping terms, referring sometimes to a deontic status and sometimes to a deontic attitude (sometimes to doxastic commitment and sometimes to acknowledgment of such a commitment), so 'intention' is ambiguous in scorekeeping terms, referring sometimes to a deontic status and sometimes to a deontic attitude (sometimes to practical commitment and sometimes to acknowledgment of such a commitment). Believing in the sense that entails one's readiness to avow what one believes and to act on it corresponds to acknowledging a doxastic commitment. Intending in the sense that entails one's readiness to act on it (and, should the expressive resources for doing so exist in the linguistic practices in question, to avow it with a 'shall' claim) corresponds to acknowledging a practical commitment. In this sense of 'intention', then, intentions are causes, for in the properly trained agent, acknowledgments of practical commitments reliably causally elicit performances. In this sense of 'belief', when beliefs provide reasons that entitle one to a practical commitment, they may function also as causes. They do just in case the acknowledgment of the practical commitment in fact arose by inferring it from an acknowledgment of the belief playing the role of premise in practical reasoning.

One, however, may have intentions without reasons, practical commit-

ments to which the agent is not entitled by doxastic commitments suitably related as premises of practical inferences. In that case there can still be action, but it will not be caused by reasons. Again, the agent may be entitled to the practical commitment, according to a scorekeeper, by doxastic commitments that are attributed by the scorekeeper but not acknowledged by the agent. This might happen where the scorekeeper takes the agent to have undertaken corresponding doxastic commitments as inferential consequences of others that are acknowledged, but where the agent has never been through the process of inference that would lead to acknowledging those consequences. Commitments of this sort could still entitle the agent to the practical commitment, even though only the scorekeeper, and not the agent, would be in a position to exhibit the practical reasoning that secures that entitlement. In such cases, too, the agent's reasons for the action would not be functioning as causes. So once the deontic statuses of commitment and entitlement to commitments are properly sorted out, it turns out that a nonlinguistic performance can have at least three different sorts of scorekeeping significance:

1. an agent's acting intentionally—that is, acknowledging a practical commitment by producing a performance or exercising a reliable non-inferential differential disposition to respond to acknowledgments of practical commitments by producing a performance,
2. an agent's having reasons for action or acting *with* reasons—that is being entitled to a practical commitment, and
3. an agent's acting *for* reasons, the action being caused by (attitudes toward) reasons for action—that is the acknowledgment of the practical commitment having arisen by a process of inference from acknowledgment of the commitments that provide the entitling reasons.

The first does not entail the second (nor vice versa), nor does the second entail the third, though they are all compatible; one can act intentionally either with or without reasons, and one may or may not act for the reasons one has.

#### *4. Acknowledging Commitments Need Not Be Modeled on Promising*

Davidson's own view about intentions (once he comes to countenance them at all) identifies them as all-things-considered judgments, in the light of all the agent's primary-reason-providing beliefs and desires, that an action of a certain kind is desirable, good, or ought to be performed.<sup>54</sup> From the present point of view this is an unsatisfactory conclusion, both because of its appeal to unanalyzed notions of desirability, good, and what ought to be done, and because it does not say what it is for the attitudes these locutions express to become explicit in the form of an evaluative judgment.

He has not laid out the practices that could confer such a content on a judgment, so he has not explained how to understand fully the implicit commitment that is being made propositionally explicit by the use of this vocabulary. Such a complaint is of exactly as much interest as the concrete alternative account that is recommended; it is better to light a candle than to curse the darkness. The deontic scorekeeping account of practical commitments, built on that of doxastic commitments, is meant to supply this want. In order to recommend that idiom over the one Davidson endorses, however, it is necessary to confront the argument against construing intentions as a sort of commitment that he offers along the way to his identification of intentions as a special kind of judgment.

The leading idea of the present account is that acting intentionally is doing something that has the deontic scorekeeping significance of acknowledging a practical commitment (in the case of intentions in action), or noninferentially producing a performance by exercising a reliable differential disposition to respond to the acknowledgment of a practical commitment (where prior intentions are involved). Intentions are identified with such acknowledgments of commitments, and the reasons for or with which an agent acts with the attitudes or facts that entitle that agent to those practical commitments, according to the role they play as premises in practical inferences. This normative (more specifically deontic) approach to intention and action is rooted in Sellars's discussion of the giving and asking for reasons for action, which has been elaborated along different lines by Castañeda.<sup>55</sup> Although the details of their accounts are different, the overall approach is very similar. Sellars never actually talks about intentions in terms of commitments, but this way of putting it is implicit in his account.<sup>56</sup> If there is something wrong with thinking about intentions in terms of commitments, then this whole approach is broken-backed. So it is of the first importance to consider Davidson's arguments against it.

Davidson begins by considering theories that focus on the speech act of expressing an intention (the speech act that Sellars regiments using 'shall'). He observes that "saying, under appropriate circumstances, that one intends to do something, or that one will do it, can commit one to doing it; if the deed does not follow, it is appropriate to ask for an explanation."<sup>57</sup> The suggestion that forming an intention is performing a speech act of this sort (perhaps addressed to oneself)—a performative theory of intention—is rejected because "the performative character of commands and promises which makes certain speech acts surprisingly momentous depends on highly specific conventions, and there are no such conventions governing the formation of intentions."<sup>58</sup> Indeed it seems enough to observe that, although for Davidson as for the deontic scorekeeping account, one must be able to talk in order to have intentions (because it is only in the context of linguistic practices of giving and asking for reasons that anything could be accorded the significance of an intention), there is no necessity that there actually be a term 'shall' that

overtly marks a special speech act that attaches to propositional contents the significance of explicit undertakings of practical commitments. The present explanatory strategy demands rather that that implicit force first be explained in scorekeeping terms so that a clear sense can be made of the introduction of locutions whose expressive function is to make that force explicit in the form of an assertion. The present account does not understand the undertaking of a practical commitment as requiring a special speech act. What is important is the attitude of acknowledging a commitment to act; any connection with special performative speech acts comes later.

But Davidson objects as well to the invocation of commitments in this context: "Promising involves assuming an obligation, but even if there are obligations to oneself, intending does not normally create one. If an agent does not do what he intended to do, he does not normally owe himself an explanation or apology, especially if he just changed his mind; yet this is just the case that calls for explanation or apology when a promise has been broken. A command may be disobeyed, but only while it is in force. But if he does not do what he intended because he has changed his mind, the original intention is no longer in force."<sup>59</sup> There are a number of points being made here; they turn on disanalogies between forming an intention and making a promise, which serves for Davidson as the paradigm of the undertaking of a commitment. To begin with, he offers the implicit suggestion that there may be problems with the notion of making a promise to oneself. Promises (like commands) are made to someone, while no one else is typically addressed by the formation of an intention. Then the central objection is presented, that there seems to be no *sanction* associated with failure to perform as one is committed to perform. If there is a commitment, then fulfilling it or failing to fulfill it ought to make some sort of difference, as it does in the case of failing to fulfill a promise or to carry out an order from a suitable authority. Yet once an intention has been formed, it can be withdrawn without penalty—the agent can have a change of mind. Promises would not be promises, would not involve the undertaking of *commitments*, if they could be canceled at the whim of the promiser. How could sense be made of a commitment that was in force only as long as the one committed decided to keep it in force but that could be relinquished without penalty at any time? Davidson concludes that the disanalogies are too great and that forming an intention cannot sensibly be conceived as undertaking a commitment.

These disanalogies between intending and promising, even promising oneself, should be acknowledged. But the conclusion follows only if there is no other model of acknowledging or undertaking commitments available besides that of promising. The deontic scorekeeping account of acknowledging *assertional* or doxastic commitments shows that this is far from being the case. Assertional commitments, after all, can be withdrawn without penalty by the asserter who undergoes a change of mind. Commitments of this sort

are put in force by the performance of a speech act and, except for special cases (where one is also consequentially committed to the claim in virtue of other, unwithdrawn assertions), can be relinquished without penalty by another speech act. It is true that nothing resembling promising could work this way, but the model of asserting shows that there are other ways to conceive the undertaking and acknowledging of commitments. While a doxastic commitment is in force, that fact has consequences; the undertaking of such a commitment has a significance for the deontic score. Commitment to one content entails commitment to others and precludes entitlement to yet others.

It is the same with practical commitments as here presented. Undertaking one is not without significance simply because it can be voided, withdrawn, or overridden essentially at the whim of the agent. For when such a commitment is in force (according to a scorekeeper who attributes it), it is significant. It entails various further commitments and precludes various entitlements. It can license the attribution of doxastic commitments (standing in for beliefs) that would warrant it, according to an attributed piece of practical reasoning. Like doxastic commitments, practical commitments involve a (conditional) justificatory responsibility to vindicate the commitment by demonstrating entitlement to it (upon suitable challenge). This forms part of the significance of these commitments, on the side of antecedents rather than of consequents, for it determines the circumstances under which it is appropriate to acquire these statuses. The disanalogies between promising and undertaking doxastic commitments do not make the latter sort of status unintelligible as a species of commitment, and the same disanalogies between promising and the undertaking of practical commitments, which are modeled closely on doxastic ones, do not make that sort of status unintelligible as a species of commitment.

It might be objected that the disanalogies between doxastic and practical commitment reinstate the difficulty. For on the one hand, doxastic commitments are like those undertaken by promising, and unlike intendings, in that they are intelligible only in terms of a speech act that has the significance of an overt public acknowledgment of them. And on the other hand, a scorekeeping sanction for failing to fulfill the justificatory responsibility associated with undertaking a doxastic commitment is the loss (in the eyes of the scorekeeper who attributes the failure of entitlement) of its authority, its capacity to license commitment by others to that same content. But the lack of this sanction is precisely one of the important points of disanalogy between doxastic and practical commitments.

Each of these points might have force if practical commitments were conceived as autonomous—that is, as statuses that could be instituted by practices that did not also institute doxastic deontic statuses. This sort of autonomy is claimed only for doxastic commitments and entitlements, however, not for their practical counterparts. In this context, the disanalogies

between doxastic and practical deontic statuses do not reinstate Davidson's objections. The first point just shows that because of the essential role played by the overt public acknowledgment of doxastic commitments by the performance of speech acts accorded the significance of assertions, doxastic commitments are more like those undertaken by promising than are practical commitments. For practical commitments need stand in no such intimate relation to speech acts accorded the significance of acknowledgments of them. Explicating locutions permitting the production of speech acts of this sort can be introduced, but practical commitments are intelligible even in their absence. Practical commitments as here conceived are unintelligible apart from all reference to the overt undertaking of commitments by speech acts; that is why they are an essentially linguistic phenomenon. But as here described, the only sort of speech act they presuppose is assertion, the acknowledgment not of practical but of doxastic commitments.

The second point was that intentions do not implicitly claim the sort of interpersonal authority that assertions do, so that the consequences of attributing commitment without entitlement cannot be in the practical case what they are in the doxastic case, namely the undercutting of that authority. But this does not show that entitlement to practical commitments is unintelligible, only that it is different in detail from entitlement to doxastic commitments. Entitlement to practical commitments still has an intrapersonal significance, for instance in connection with the incompatibility of practical commitments (which is linked to that of doxastic commitments), and so with their permissive entailments. Again, this feature does not threaten to make practical commitments unintelligibly private, both because of the irreducibly social character of the deontic scorekeeping, in terms of which such statuses are explained by theory and sustained by practice, and because of the connection with doxastic commitments via practical reasoning.

Entitlements aside, treating a performance as (or as elicited by) an acknowledgment of a practical commitment—that is, treating it as intentional—has scorekeeping consequences not only for the attribution of further practical commitments but also for the attribution of doxastic commitments. It is often possible to infer what an agent believes from what that agent does. Committing oneself to act in a certain way may be committing oneself to taking the world to be a certain way, in the eyes of a scorekeeper who attributes a suitable background of other commitments (of both discursive species). The doxastic commitments an agent is taken consequentially to have undertaken in this way may be incompatible with other doxastic commitments the scorekeeper attributes, in which case entitlement to all of them is undercut. So in part in virtue of the intimate connections between them, the asymmetries between practical and doxastic commitments do not threaten the intelligibility of the deontic scorekeeping significance of the former.

### 5. 'Should' and 'Shall'

Once these points are appreciated, it becomes clear that Davidson's considerations provide no reason not to understand forming an intention as acknowledging a commitment, provided commitment is properly understood according to the practical deontic scorekeeping model of doxastic commitment rather than on the model of promising. But it also becomes clear that there is surprisingly little difference between his construal of intentions as all-out evaluative judgments and the deontic scorekeeping construal of them as acknowledgments of practical commitments. For, first, though Davidson does not think of them that way, in the context of the model presented here, taking intentions to be a kind of judgment *is* taking them to be a kind of commitment. Davidson uses 'judgment' as the genus and allows cognitive and conative, or descriptive and evaluative, species corresponding to beliefs and intentions. The idiom presented here uses 'commitment' as the genus and allows doxastic and practical species corresponding to beliefs and intentions.

For Davidson, there are two sorts of "evaluative" judgment: those that express pro-attitudes suited to be elements of primary reasons for action (that is, those that express merely *prima facie* or *ceteris paribus* evaluations), and those that express intentions and are directly responded to by the production of suitable performances. In the case of practical reasoning whose conclusion is the formation of an intention, the agent is noninferentially disposed to respond reliably by producing suitable performances—which are qualified as actions by having such a provenance; these two sorts of evaluations appear in Davidson's account in the role of *premises* and of *conclusions*, respectively. The account offered here denies that what is expressed by the *prima facie* evaluative judgments Davidson understands as codifying pro-attitudes need appear as explicit premises in such reasoning; they correspond to the endorsement of a pattern of practical inference as entitlement-preserving. Such practical inferential commitments may be made explicit in the form of doxastic commitments with assertible contents (and so be available for duty as explicit premises) if suitable explicating vocabulary is available—just as theoretical inferential commitments may, but in general need not, be codified explicitly by the use of conditionals. In this use, then, normative expressions exemplified by 'should' as it appears in rules of conduct play an inference-explicating role on the practical side that is analogous to that played by 'if . . . then . . .' on the doxastic side;<sup>60</sup> in neither case does the omission of a premise that codifies a material propriety of inference, whether practical or doxastic, result in an enthymeme. Besides this *permissive* use of normative locutions such as 'should', which corresponds to Davidson's *prima facie* evaluative judgments, there is also a *committive* use, which corresponds to the "all-out" evaluative judgments that serve for him as intentions.

Recall the discussion of intentional explanation in Chapter 1. One of the ideas advanced there in connection with the suggestion that intentional states be understood in terms of deontic statuses and [propositional] attitudes toward them is that the conclusions of intentional explanations in the strict sense are *normative*, rather than *descriptive* claims. One attributes beliefs and desires (or other evaluations or pro-attitudes) and concludes from those attributions, not that the agent *will* perform an action of a certain kind, but that the agent is *committed* by those beliefs and desires to do so, that in the light of those other attitudes the agent *ought* (rationally) to do so. Intentional explanation illuminates what was done by showing why the intentional agent was *committed* to acting in that way. Under various circumstances it is possible to continue the inquiry and to ask why the agent acted in accord with that commitment. The response to such a question is not an intentional explanation, however, but a different sort of account—one showing why it is useful to offer intentional explanations of this individual, why treating the individual as a rational agent is a useful predictive and explanatory strategy. Explanations of this supplementary sort may appeal to how the organism is wired up and how it was trained so as to be able to respond reliably to the acknowledgment of a practical commitment by producing a performance of the sort specified in the content of that commitment. Such considerations are offstage from the point of view of intentional explanations proper, for these go only as far as showing what an agent should (rationally) do, what the agent is committed to do by the doxastic and inferential commitments that agent acknowledges.

Intentional explanations display sample pieces of practical reasoning, attributing theoretical and practical deontic statuses as premises and attributing a practical commitment as a conclusion. To serve as an intentional explanation of something the agent did, or to draw a conclusion about what the agent should do on which a prediction might be based about what the agent will do, these must be treated as commitment-preserving inferences. For the conclusion is that the intentional agent was or is *committed* to act in a certain manner. When the deontic scorekeeper attributes various commitments and concludes that therefore the agent *should* perform an action satisfying a particular description, the evaluative judgment expressed is of Davidson's second, all-things-considered kind.

So on the deontic scorekeeping approach there are two sorts of 'should', corresponding to the two sorts of evaluative judgment that Davidson considers: one involving *prima facie* evaluations suitable to serve as premises in practical reasoning, and one involving all-in evaluations suitable to serve as conclusions in practical reasoning. The first sort of 'should' is used to make explicit the endorsement, undertaken or attributed, of a pattern of practical reasoning, as in "Bank employees should wear neckties." The second sort of 'should' is used to make explicit commitments to act, which are attributed as the conclusions of committive practical inferences attributed in the course

of intentional explanation. To say this is to say that in the latter sort of use, 'should' expresses in the third person what is expressed in the first person by 'shall'. (Indeed, 'shall' and 'should' are etymologically linked in just the way suggested by this doctrine.)

Translated into the official scorekeeping terminology of deontic attitudes, this is the claim that while 'shall' is used to make explicit the *acknowledgment* (and therefore the *undertaking*) of a practical commitment to make some claim true, 'should' is used to make explicit the *attribution* of such practical commitments. The same piece of practical reasoning can be presented from either social perspective. In first-person, deliberative terms, the agent may acquire a practical commitment that would be made explicit (if the idiom encompasses sufficient expressive resources) by an overt utterance of "I shall wear a necktie," as the result of an inference from acknowledged commitments that would be made explicit (perhaps in response to a challenge to demonstrate entitlement to the practical conclusion) by an overt utterance of "I am a bank employee." In third-person, scorekeeping terms, the scorekeeper may attribute a practical commitment, adopting an attitude that would be made explicit (if the idiom encompasses sufficient expressive resources) by an overt utterance of "He should wear a necktie," as a result of an inference from attributed commitments that would be made explicit by an overt utterance of "He is a bank employee." The same piece of practical reasoning can be exhibited either by the one undertaking a practical commitment or by the one attributing it—significant either in deliberation regarding action or in assessment of such action. Indeed, deliberation—my considering various practical inferences in order to decide what I *shall* do—is just the internalization of assessment, the consideration of what anyone, given the relevant collateral commitments and circumstances (as they are taken by the assessor to be), *should* do.

## 6. Weakness of the Will

Of course 'should' has first-person uses as well. Some of these express only endorsement of patterns of permissive inference and so correspond to Davidson's merely *prima facie* evaluative judgments: "I should wear a necktie" (since I am a bank employee, but only if there is no better reason not to do so). But some also are self-attributions of commitments, in which one takes up a third-person perspective toward oneself, drawing conclusions about what one's reasons commit one to do: "I should (all things considered) drive to the airport." With practical commitments, as with doxastic ones, although acknowledging a commitment entails attributing it to oneself, the converse is not the case; though attributing a commitment to oneself is one way to *undertake* that commitment, this can be a consequential undertaking, rather than an acknowledgment. In particular, a self-attribution of the 'I should . . .' variety need not trigger the reliable differential dispositions to

respond to an acknowledgment of a practical commitment by producing a suitable performance. The noninferential significance of the deontic attitude that is made explicit by 'I should . . .' can be different from that of the deontic attitude that is made explicit by 'I shall . . .' Thus one can self-attribute a practical commitment without acknowledging it in the sense that matters for eliciting action.<sup>61</sup>

This possibility is one of the phenomena philosophers have discussed under the heading of weakness of the will, or *akrasia*—knowing the better and doing the worse. The attitude expressed by 'I should . . .' in its all-in sense does indeed commitment-entail that expressed by the corresponding 'I shall . . .' statement. But the difference between acknowledging a commitment and consequentially undertaking it depends on the fact that one does not always acknowledge the consequences of commitments that one acknowledges. In scorekeeping terms one can nonetheless be said to undertake those consequential commitments because the initial acknowledgment licenses others to attribute them. This distinction remains even when one takes up a third-person point of view toward oneself, as in deliberation about various possible courses of action when the agent traces out what commitments would be undertaken consequentially were certain others acknowledged. For in deliberating, an agent considers what commitments would be attributed by scorekeepers, under various circumstances. It is for this reason that one must be able to assess the conduct of others in order to deliberate about one's own.

The akratic's deliberations and intentions are out of step; the commitments acknowledged in the third-person theoretical way are incompatible with those acknowledged in the first-person practical way. The akratic agent is accordingly the analog on the practical side of the believer who undertakes incompatible doxastic commitments. It is one of the cardinal strengths of the deontic scorekeeping approach to intentional states in terms of normative statuses that there is nothing conceptually mysterious about the possibility of such incompatible commitments. Difficulties in coherently understanding akratic action and endorsement of incompatible beliefs arise from exclusive emphasis on a causal-functional model of intentional states.

The account of action presented here is a thoroughly Kantian one. For Kant, *will* is just the capacity for practical reasoning—that is, the capacity to derive performances from a conception of laws.<sup>62</sup> In the terminology introduced here, this is just the capacity to respond reliably to acknowledgments of commitments (the pragmatic version of "deriving from conceptions of laws") by producing suitable performances—suitable in terms of the way their descriptions line up with the contents of the practical commitments they either acknowledge or by the acknowledgment of which they are responsively elicited. For Kant the expressive role of 'ought' or 'should' (*Sollen*) is to make norms explicit in the form of imperatives. Specifically, such normative vocabulary "indicates the relation of an objective law of reason to a will

which is not in its subjective constitution necessarily determined by the law."<sup>63</sup> As construed here, normative vocabulary (of which 'ought' and 'should' are paradigmatic) has the logical expressive function of making explicit in the form of something that can be *said* (put in the form of a claim) an attitude that otherwise could be implicit only in what is *done*—namely, the endorsement of a pattern of practical reasoning. The propriety of a form of reasoning is the practical correlate of Kant's "objective (= valid, binding) law of reason," and its acknowledgment as constraining actual practical reasoning is its "subjective" relation to the will.

The rational will as described here is not a particularly puzzling phenomenon. Its normative dimension is explained by extending the account of discursive commitments to encompass not only doxastic but practical deontic statuses. Its causal dimension is explained by exploiting the analogy between discursive entries and exits, between action and perception. The relation between the normative and the causal aspects of rational willing or practical reasoning is explained by appealing to the causal efficacy of the deontic attitude of *acknowledging* commitments: acknowledgments of doxastic commitments can be reliably differentially elicited as responses to environmental situations in perception, and acknowledgments of practical commitments can reliably differentially elicit performances as responses in action. Reasons can be causes because deontic scorekeeping attitudes can play both normative and causal roles. There is much still to be learned about the empirical details of the differential responsive dispositions that make possible these discursive entry-and-exit practices, but it is not hard to understand in principle how there can be such things. We are rational creatures exactly insofar as our acknowledgment of discursive commitments makes a difference to what we go on to do—on the side of action, insofar as we incorporate a connection between what is expressed by 'should' and what is expressed by 'shall'.



---

## **PART TWO**

---



# The Expressive Role of Traditional Semantic Vocabulary: 'True' and 'Refers'

---

It was when I said,  
"There is no such thing as the truth,"  
That the grapes seemed fatter.  
The fox ran out of his hole.

You . . . You said,  
"There are many truths,  
But they are not part of a truth."  
Then the tree, at night, began to change.

WALLACE STEVENS, "On the Road Home"

---

## I. FROM INFERENCE TO TRUTH, REFERENCE, AND REPRESENTATION

### 1. *Discursiveness, Sapience, and Sentience*

We are the creatures who say 'we'—who can explicitly take or treat someone as one of us. Adopting this practical attitude is adopting a discursive normative stance: attributing propositionally contentful commitments and entitlements, and attributing to those to whom we attribute them in turn the practical recognition of such deontic statuses by their corresponding acknowledgment and attribution of them. Sapience of the sort distinctive of us is a status achieved within a structure of mutual recognition: of holding and being held responsible, of acknowledging and exercising authority. The specifically *discursive* character of that normative social structure—what makes it appropriate to interpret the statuses we institute by our deontic scorekeeping activities as having their significance determined by *propositional* contents—consists in the *inferential* articulation of those cognitive practices. We are the ones who give and ask for *reasons* for what we say and do.

Propositionally contentful commitments are picked out in the first instance as those that can both serve as and stand in need of reasons; it is playing an appropriate role in the practices of giving and asking for reasons that confers propositional contents on commitments and the performances

that express them. Offering a reason is making a claim—performing a speech act with the force or significance of an assertion, the undertaking (by overt, explicit acknowledgment) of a doxastic commitment. Practices in which performances are accorded assertional significance deserve for that reason to be called *linguistic* practices. That what one interprets a community as producing and consuming are *reasons* is not something that can be achieved simply by a stipulation to that effect on the part of the interpreter. Rather, the interpretation must attribute to the community practices that incorporate a suitable structure of *inheritance* of *entitlement* to the commitments the practitioners are understood as undertaking and attributing. Such a structure arises out of the interaction of two different dimensions: the intercontent, intrapersonal *consequential* inheritance of entitlement, and the intra-content, interpersonal *testimonial* inheritance of entitlement. Each deontic attitude on the part of a scorekeeper is the attribution to (or acknowledgment by) an interlocutor of a deontic status having a certain content. What it means for commitments with *different* contents to be undertaken by or attributed to the *same* interlocutor, and for commitments undertaken by or attributed to *different* interlocutors to have the *same* content, is to be understood in terms of this broadly inferential structure of justification and communication.

We are sentient creatures as well as sapient ones, but our sentience is different from that of those who cannot give and ask for reasons. Described in the language of physiology, our sensing may be virtually indistinguishable from that of nondiscursive creatures. But we not only sense, we also perceive. That is, our differential response to sensory stimulation includes noninferential acknowledgment of propositionally contentful doxastic commitments. Through perception, when properly trained and situated, we find ourselves passively occupying particular positions in the space of reasons.

We are practical creatures, as well as linguistic ones, but our purposive activity is different from that of those who cannot give and ask for reasons. Described in the language of physiology, our motor activity may be virtually indistinguishable from that of nondiscursive creatures. But we not only produce performances, we perform actions. The performances we produce include noninferential responses to acknowledgments of propositionally contentful<sup>1</sup> practical commitments. Through action, when properly trained and situated, we can respond to the particular positions we occupy in the space of reasons by actively altering the nondiscursive environment.

Our mammalian cousins, primate ancestors, and neonatal offspring—who are sentient and purposive but not discursive creatures—are interpretable as perceiving and acting only in a derivative sense. An interpreter can make sense of what they do by attributing propositionally contentful intentional states to them, but the interpreter's grasp of those contents and of the significance of those states derives from mastery of the richer practices of giving and asking for reasons for the sort of doxastic and practical discursive com-

mitments that are not attributed to these simpler folk. The activities they are interpreted as engaging in do not suffice to confer anything recognizable as propositional contents on their states, attitudes, and performances. Our discursive practices make us semantically autonomous in a sense in which their nondiscursive practices do not.

## *2. Making and Taking Reasons, Seeking and Speaking Truths*

An account has been offered of practices that are, it is claimed, sufficient to confer genuinely propositional contents on our states, attitudes, and performances. For a sentence to express such a content is for its proper use to be determined by the norms implicit in assertional practices such as those described in Chapters 3 and 4. Even if that account is accepted, however, it is not obvious from what has been said so far why various features of the discursive scorekeeping model ought to be treated as *necessary* for propositional contentfulness. It might be granted that treating a community as engaging in the sort of linguistic social practices described in Chapters 3 and 4 ought to count as interpreting them as instituting doxastic and practical deontic statuses with assertible and therefore propositional contents—as in this sense doing what we discursive interpreters do—while insisting that the linguistic and social dimensions of these attributed practices are not needed for the states creatures are taken to have to deserve to count as propositionally contentful. Responding to this sort of objection requires looking more closely at the notion of propositional content.

The account elaborated in Part 1 of this work focuses exclusively on one way of understanding what it is to exhibit such contents: in terms of inferential articulation, above all playing a premissory role in reasoning. As was acknowledged at the outset of this investigation, however, there is another, perhaps more familiar route to understanding the sort of sapience being called on to do demarcational duty. It begins with the concept of *truth*, rather than that of *inference*. We are believers, and believing is taking-true. We are agents, and acting is making-true. To be sapient is to have states such as belief, desire, and intention, which are propositionally contentful in the sense that the question can appropriately be raised under what circumstances what is believed, desired, or intended would be *true*. Understanding such a content is grasping the conditions that are necessary and sufficient for its truth.

Something has been said about why understanding ourselves as makers and takers of reasons entails understanding ourselves as seekers and speakers of truths; acknowledging a doxastic commitment is taking-true, and producing an assertional performance is putting a claim forward as true. This discussion at best offers only a starting point, however, from which an account of the relation between the theoretical concepts of inference and truth might depart. It needs to be supplemented by an account of the sort of socially articulated inferential role something must play in order to have a

propositional content in the sense of having truth conditions. Theorists who begin by associating sentences or states with truth conditions—whether by bare stipulation or, more ambitiously, by exhibiting practical proprieties of use or function sufficient to confer such contents (so the association is established by the creatures being interpreted rather than the interpreter)—can then derive from that association various norms governing inference. For the inference from  $p$  to  $q$  can be understood as correct (in the sense of commitment-preserving) just in case the truth conditions of  $p$  are a subset of the truth conditions of  $q$ .<sup>2</sup> An account such as the present one, which pursues the converse direction of explanation by understanding propositional contents in the first instance in terms of proprieties of inference, accordingly owes a corresponding story about assessments of truth and the association of truth conditions with sentences and the beliefs they express.

In order to turn a construal of the practical attitude of taking-true (as undertaking an inferentially articulated doxastic commitment) into an account of the sort of grasp of the propositional content of a state or utterance that consists in associating truth conditions with it, it is necessary to understand the expressive role played by ‘true’ in its various uses. What must an interpreter be doing for it to be correct to take what is being associated with the utterances and intentional states of those interpreted to be *truth* conditions? An answer to this question has to explain how an expression (possibly in an alien language) needs to be used—what practical proprieties its employment must be subject to—in order for it properly to be understood as meaning ‘true’. In the context of the present project, such an account is subject to two further criteria of adequacy.

First, the vocabulary used to specify the expressive role played by ‘true’ and its cognates must be that provided by the inferentially and socially articulated deontic scorekeeping model of discursive practice. That is, it is necessary to show how that model can be extended to include truth talk. For the only locutions officially acknowledged to be intelligible within the confines of this work are those that can be understood as having just the contents that are conferred on them by the role they play in discursive practice. So the issue is what one would need to do to add the expressive power provided by ‘true’ to the game of giving and asking for reasons as so far specified.

Second, the inferentialist order of explanation is committed to showing why expressions, performances, and intentional states having semantic contents that are propositional in the sense of being assertible—here construed roughly as both being capable of serving as a reason and potentially standing in need of reasons—therefore are also appropriately talked about as having contents that are propositional in the sense of having truth conditions. What is the connection between these two ways of picking out the propositional? Both of these explanatory demands can be addressed by showing what it is about the inferentially and socially articulated deontic scorekeeping prac-

tices (which according to the story being told here confer propositional contents) that 'true' has the job of expressing explicitly (that is, in assertible, so propositional, form).

### 3. *The Representational Dimension of Discursive Practice*

The need for an account of truth talk in broadly inferential terms is only part of a larger methodological requirement: that the *representational* dimension of intentionality and discourse be addressed. The concept of representation lies at the center of the theoretical idioms typically used today to discuss intentional and semantic contentfulness. It has been at least since Descartes forged his worldview in the bifurcated mold, one side of which consists of representings and the other of representeds—the former conceived as discursive, on the model of algebraic equations, the latter as geometric, on the model of the figures those equations determine. Four aspects of that representational idiom deserve special attention.

First, in addition to the use of 'true' to formulate a semantic approach to sentences in terms of truth values and truth conditions, the tradition to which the later Frege gave birth uses 'refers' or 'denotes' to express the representational relations terms and predicates stand in. Tarski's discussion of truth definitions for formalized first-order languages<sup>3</sup> provides a paradigm of how the use of 'true' and of 'refers' or 'denotes' ought to be understood to be related to each other.

Second, in tandem with this semantic focus on truth and reference is a distinction between extensional and intensional contexts. Central among the latter (which are the source of special explanatory difficulties in the representationalist context) are propositional-attitude-ascribing locutions, paradigmatically ". . . believes that . . ." An account must be offered of the different semantic behavior of these two classes of locutions—a difference that can be roughly rendered in the traditional terminology as the distinction between contexts for which only what is represented matters for the inferential significance of sentences exhibiting those contexts, on the one hand, and contexts in which representings themselves are somehow involved in determining such inferential significances, on the other.

Besides these technical semantic notions crafted by self-conscious theorists in the representationalist tradition, there are locutions in ordinary language that are used to express claims concerning what people are talking or thinking *about*. Prime among these ordinary representational locutions is the use of 'of' in *de re* ascriptions of propositional attitudes, as in "Royce believed of the author of *Dreams of a Spirit-Seer* that he was a great philosopher" or "What James said about tough- and tender-minded thinkers is true of Kant, but not true of Hegel." The nontechnical notion of intentional *aboutness* that figures in the fundamental pretheoretical distinction between what people are talking or thinking about and what they say or think about it is to be

understood in terms of what is expressed by the use of these *de re* specifications of the contents of intentional states and the utterances that express them. It is from these roots in ordinary conversational practice that more technical representational concepts are elaborated.

Finally there is the crucial idea of *objective* representational proprieties of judging and inferring. It is a feature of our assessments of the application of at least some concepts that we take it that besides correctness of discursive attitude in the sense of entitlement to the commitment undertaken—the concept-user being *justified* in the application, either noninferentially by observable circumstance, inferentially by concomitant commitments and entitlements, or deferentially by available testimony—there is also a kind of correctness that is determined by how things are with what is represented. Whether the application of a concept is correct or incorrect in this sense is independent of the attitudes of the one applying the concept. ("Full many a gem of purest ray serene / The dark unfathom'd caves of ocean bear: / Full many a flower is born to bloom unseen, / And waste its sweetness in the desert air.")<sup>4</sup> Making sense of this fundamental characteristic of our linguistic practice means funding a notion of discursive *success* (both doxastic and practical) that transcends our attitudes (our takings and tryings) and answers instead to the *objects* those attitudes address. Such success is determined by the properties of those objects and by the relations they stand in. A certain sort of social and inferential articulation of attitudes must be shown to institute proprieties and confer contents such that what it is correct to conclude or to claim and what one has actually done depends on how the objects referred to, talked about, or represented in one's discursive attitudes actually are. Talk of objective representational conceptual contents involves both representation of objects and objective proprieties of representation. It remains to be seen what these are and how they are related.

On the face of it, acknowledgment and analysis of conceptual proprieties that are objective in this sense would seem simply to be incompatible with a social-phenomenalist approach to norms. For what is distinctive of such an approach is precisely its treating the norms that govern our discursive conduct as instituted ultimately by our *attitudes*. How, then, can such an approach hope to make intelligible a notion of attitude-*transcendent* discursive norms? The answer takes the form of specifying a certain sort of *stance*—a constellation of attitudes that can be understood as taking or treating certain sorts of assessments *as* answering to facts concerning the objects represented rather than to anyone's attitudes toward or claims about them. Understanding what it is to adopt such a representational stance toward (or interpretation of) what we do is understanding the *implicit* practical attitudes that are expressed *explicitly* by our use of representational locutions—both the technical semantic tropes involving 'true' and 'refers' and the *de re* ascriptions of propositional attitude used in ordinary language to make clear what has been achieved in the way of communication or action. Indeed, it is just

the aspects of conceptual content made explicit by these representational locutions that turn out to require that the practices conferring those contents be specifically *linguistic social* practices of the sort outlined in the previous chapter.

#### 4. Inference, Substitution, and Anaphora

The various representational tropes just considered all involve associating semantic contents with subsentential expressions—above all, singular terms and predicates. The understanding of reference to objects, of the distinction between extensional and intensional occurrences of expressions, of *de re* locutions, and of the notion of the truth of at least some claims as a normative status or propriety that answers to the properties objects have (and the relations they stand in) rather than those anyone *takes* them to have (or stand in) all involve grasping—either implicitly as a practitioner or explicitly as a theorist—conceptual contents that are not propositional. No account of this sort of conceptual content has yet been offered here. The semantic contentfulness of sentential expressions has been derived from the role they play in inferences, as premises and conclusions; the semantic contentfulness of expressions in the grammatical categories of singular terms and predicates cannot consist in their playing inferential roles in this sense. The notion of conceptual content must accordingly be broadened beyond the propositional in order to give it application beyond the category of sentences.

The inferentialist order of semantic explanation was given its modern form by the young Frege, whose definition of *begriffliche Inhalt* motivates one strand of the model developed here. (The social, practical, and specifically linguistic dimensions of that model—and to some extent the understanding of the normative one as well—are rooted elsewhere.) One of Frege's most important ideas is his strategy for extending broadly inferential notions of content to subsentential expressions, which cannot serve as premises and conclusions of inferences. The key theoretical concept he introduces for that purpose is *substitution*. His idea is that the way in which sentences are related to one another when one results from the other by substituting one subsentential expression for another confers an *indirectly* inferential role on the occurrence of subsentential expressions. Roughly, subsentential expressions can be sorted into equivalence classes that can be thought of as having the same conceptual content in an extended sense. For they can be assimilated insofar as substitution of one for another does not alter some feature of the inferential role of the sentences they are substituted into—paradigmatically, insofar as such substitution does not turn any materially good inferences those sentences are involved in into materially bad inferences. It is this methodology for carving up sentences into semantically significant subsentential parts by noting inferential invariants under substitution that he elaborates.

rates into the theory of *functions*, which throughout his later work he takes to be one of his greatest intellectual contributions.

In Chapter 6 Frege's notion of substitution is used to investigate the fine structure of the inferential roles conferred on sentences (and the states whose contents they express) by discursive scorekeeping practices of the sort already discussed. The key notion is that of substitution *inferences*: those whose conclusions are substitutional variants of their premises. The scorekeeping significance of the occurrence of singular terms and predicates is explained by their association with substitution-inferential *commitments*. These are commitments to the propriety of a particular range of substitution inferences: those that exhibit a certain pattern preserving the deontic status of doxastic or practical commitment in the passage from premise to conclusion. In these terms it is possible to explain what it means for propositionally contentful sentences to represent objects as exhibiting properties and standing in relations—and as a consequence to offer an account of the distinction between extensional and intensional sentential contexts in which such subsentential expressions may occur. The official theoretical semantic apparatus is accordingly deepened and made more powerful by adding to the notion of *inference* that of *substitution*.

Both the sentential inferential story about conceptual contents and the subsentential substitutional one rely essentially on the *repeatability* of intentional states and their linguistic expressions. Sharpening the analytic focus, Chapter 7 explores the structures by which conceptual repeatables are constructed out of unrepeatables. Descending to this level is necessary, to begin with, in order to understand the conceptual content (and so pragmatic scorekeeping significance) of deictic or demonstrative responses to the non-discursive environment within which discursive practice is conducted. Such unrepeatable but conceptually articulated responses make a crucial contribution to empirical cognition, in particular in its relation to practical action. Indexicals (the paradigm of expression-types, tokenings of which uttered by different individuals on different occasions are grammatically guaranteed to have different significances) are only a particularly obtrusive special case of a much more general phenomenon, however—one that is fundamental to the very possibility of interpersonal communication and comprehension. Even if attention is restricted to nonindexical expression types, the ubiquity of differences in collateral beliefs or commitments across interlocutors ensures that lexically cotypical utterances issuing from different mouths will often have very different deontic significances.

It will emerge that talk of the *representational* dimension of discourse is in fact addressed to the *social-inferential* mechanisms by means of which communication can nonetheless be secured across such doxastic gaps. These mechanisms are examined in deontic scorekeeping terms in Chapters 7 and 8. Chapter 8, which discusses various forms of *ascription* of conceptually contentful deontic statuses and attitudes, looks at the expressive power of the locutions used to make various aspects of these mechanisms proposition-

ally *explicit*. As always, such explicitating vocabulary is intelligible only against the background of an understanding of the *implicit* practices that such vocabulary makes it possible to express in explicit (assertible) form.

The most important underlying structure is that of *anaphora*, the paradigm of which is the relation a pronoun stands in to its antecedent. In Chapter 7, treating one expression tokening as anaphorically dependent on another (whether they are uttered by the same or by different interlocutors) is explained as taking the substitution-inferential significance of producing the dependent tokening to depend in a certain way on the substitution-inferential significance of producing its antecedent tokening. In other words, anaphora is construed as a special mechanism for the *inheritance* of substitution-inferential commitments. The third component of the theoretical semantic idiom employed here is accordingly *anaphora*. The result is a three-tiered semantic structure: *inference*, *substitution*, and *anaphora* (ISA for short).

### 5. 'True' and 'Refers'

This chapter looks at the expressive function of the technical semantic vocabulary that has been employed by theorists outside the minority inferentialist tradition pursued here. On the side of sentences, the central locution is 'true', while on the side of subsentential expressions, it is 'refers' or 'denotes'. The discussion of the way these expressions work plays a dual role in the exposition of the account of discursive practice presented here: retrospectively as offering some of the justification for the decision to focus on *inferential*, rather than *representational*, semantic primitives (announced in Chapter 2), and prospectively as motivating the descent to *substitutional* inferential substructures in Chapter 6 (culminating in the treatment of *anaphora* in Chapter 7). For on the first, or retrospective, point, it is argued here that once the *expressive* function of 'true' and 'refers' is properly understood, it is seen to be incompatible with the *explanatory* function those locutions have been accorded in the dominant semantic theoretical tradition.<sup>5</sup> Insofar as this case can be made out, it makes urgent the demand for alternatives to these semantic primitives—precisely the demand that the inferentialist semantic program pursued here seeks to satisfy.

On the second, or prospective, point, the account of the expressive role of 'true' and 'refers' that is arrived at in this chapter construes them as functioning *anaphorically*. They are anaphoric proform-forming operators; the paradigmatic use of 'true' is to construct a special kind of prosentence, while the paradigmatic use of 'refers' is to construct a special kind of pronoun. The expressive role of traditional representationalist semantic vocabulary can be rendered only in the broadly inferentialist idiom used to specify the model of discursive practice offered here, then, inasmuch and insofar as that model can be extended so as to encompass the attribution by deontic scorekeepers of specifically *anaphoric* relations between the significance of various expres-

sion tokenings. This is a promissory note that is redeemed two chapters later, after the substitutional raw materials required have been brought onboard. Once that has been done, an account will have been offered in systematic terms of why it is useful to talk about intentional contentfulness in terms of truth and reference, even though the practices that confer the contents these locutions express should be understood rather in terms of reasons and properties of inference.

The previous two chapters presented a model of social linguistic practices. A fundamental criterion of adequacy for such a social-practical model of discursiveness or intentionality is that it be possible to show in the idiom it provides that the normative significances those practices institute are sufficient to confer genuinely propositional contents on states, attitudes, performances, and expressions that are suitably caught up in them. To do that, it is necessary to explain how the social-inferential articulation of those practices can amount to objective representational content. Such an explanation in turn requires an account, in terms of the social and inferential articulation of the pragmatic significances constituted by deontic scorekeeping, of what we are saying when we make claims about what our claims and beliefs represent or are about. The philosophical elaboration of these representational notions uses the technical vocabulary of truth and reference. But what are we saying when we say that a claim is true, or that a term refers to an object?

It is natural, and not obviously inappropriate, to understand these notions themselves in representational terms. To do so is to take '*... is true*' to denote a property of claims and beliefs, and so of the sentences that express them, and to understand '*... refers to ...*' as denoting a relation between words and the world, between linguistic or intentional items and nonlinguistic, nonintentional ones. Now a great many of our beliefs *are true*, and a great many of our terms *do* refer to objects. But it is necessary to be careful about what assumptions are made concerning what is being said when these claims are made. For exactly what we are saying when we make these true claims—indeed, what we are *doing* when we do so—is of critical importance for determining what *explanatory* role such claims are suited for.

The discussion of representational locutions in the rest of this work moves on two tracks: one having to do with technical representational terms employed by semantic theorists to make explicit various important features of the conceptual contents they address, the other with the locutions of ordinary language practitioners use to make explicit the representational dimension of the doxastic and practical commitments they attribute and undertake. Together these discussions seek to make clear what the representational purport implicit in our discursive practices consists in. This chapter looks to specify the actual expressive role played by the technical terms '*true*', '*refers*', '*denotes*', and their cognates. In order to do this precisely and clearly, it seeks to show how vocabulary playing that expressive role can be introduced into and understood in terms of the fundamental model of discursive practice

offered in Chapter 3 and Chapter 4. It also argues that once this expressive role is properly understood, it becomes clear that representational locutions are not suited to play the role of primitives in a semantic theory. Nailing down this point requires looking at their actual expressive role, which is not capturable in the sort of representational semantic theory that uses these notions as primitives; it is therefore necessary to work in another one. The conclusion one arrives at by such an analysis is that some other sort of semantic theory is needed. So the broadly inferential approach introduced in Chapter 3 serves both as the use-language in which the argument is conducted and as the answer to the question that is raised by seeing the inadequacy of truth and reference locutions to serve as fundamental semantic explanatory vocabulary—namely, if semantics ought not to be done in terms of truth conditions for sentences and reference and denotation for singular terms and predicates, what semantic primitives ought to be employed?<sup>6</sup>

The rhetorical situation at this point is complex, and there is a danger that the conclusion that the traditional semantic primitives are inappropriate for explanatory purposes (though quite in order for expressive ones) might be taken to be a result that depends on the inferentialist semantic metalanguage being employed. In fact this thought about the use of 'true' and 'refers' can be motivated independently, and that is the tack taken here. So the analysis of traditional representational semantic vocabulary in this chapter could have been presented before inferential approaches to semantics were introduced in Chapter 2, in order to motivate the search for some new semantic primitives—since most theorists use truth and reference as much because they cannot think of an alternative as for any other reason. But this order of explanation would give too much prominence to the notions of truth and reference; the model being pursued here can be motivated apart from the troubles of this semantic tradition. The expressive role of traditional semantic vocabulary is explained here in terms of *anaphora*, and the argument that this is the right way to understand 'true' and 'refers' is conducted independently of the considerations of the previous chapters. The notion of anaphora that is appealed to is in principle equally available to the representationalist, who will think of it just as grammatically guaranteed coreference.<sup>7</sup> This neutral analysis is then supplemented by showing, in Chapter 7, how anaphora is to be understood according to the inferentialist model of Chapter 3.

## II. TRUTH IN CLASSICAL PRAGMATISM

### 1. *Stereotypical Pragmatism*

The popular conception of the theory of truth of classical pragmatism is summed up in the slogan "The truth is what works." According to this view, the pragmatists were trying to give a theory of truth in the sense of offering necessary and sufficient conditions for possession of that property.

Their innovation is then seen to consist in taking the possession of this property by a belief to consist in a relation not simply to what is believed but also to what is desired. Working, or being satisfactory, involves a further argument place beyond the standard representational or correspondence notion, for it is relative to preferences, purposes, interests, needs, or some such satisfiable practical states.<sup>8</sup>

A theory of truth, on this line, is generically a pragmatic one if it treats truth as the property of conducting to the satisfaction of some state associated with the believer—paradigmatically desire. Specific versions of this genus of explanation will be distinguished by how they understand the state, its subject, and the sort of satisfactoriness involved. Thus within the pragmatic genus truth might be identified with properties as various as evolutionary adaptiveness for a species and optimality for felt-preference maximization by a time-slice of an individual agent.

This sort of understanding of truth as a property of utility for some end—a matter of how useful, in some sense, it is to hold the belief that is a candidate for truth—may be called ‘stereotypical pragmatism’. It is important to notice what sort of a theory it is. Pointing out the apparent appropriateness on some occasions of questions such as, I believe that the theory works (for instance, makes correct predictions) but how do I know it is true? already shows that this sort of pragmatism is very implausible if it is conceived as elucidating our concept of truth.<sup>9</sup> As Dewey was well aware, views of this stripe can best be maintained as revisionary proposals—not as accounts of what we mean by ‘true’ but as suggestions that we stop using that concept and get along instead with the pragmatist’s notion of utility.

Any assessment of the merits of such a proposal depends on an account of what the role of the concept of truth is, what explanatory uses the property of truth is wanted for. For only in that context can it be argued that some utility notion would better serve those ends or play that role. The significance of the classical pragmatists in the present story derives from their contribution to that antecedent question—the question of what expressive and explanatory work is and ought to be done by the truth concept. Although their account of the role of truth talk cannot, in the end, be counted as correct, it nevertheless provides the central idea around which an adequate account can be constructed. The answer that eventually emerges as to the role of ‘true’ makes it hard to see how stereotypical pragmatism, even as a revisionary proposal, can amount to anything other than changing the subject, sharing only a homonym with ordinary truth talk.

## 2. Five Theses of Classical Pragmatism

There is no question that the classical American pragmatists at times commit themselves to what has just been called ‘stereotypical pragmatism about truth’. But there is a deeper and more interesting explanatory

strategy that the pragmatists pursue as well. According to this way of setting out their account, concern for what ‘works’ or is satisfactory is only the final move in an innovative rethinking of the nature of truth and belief. The early moves are worthy of attention, even though the final one does not in the end prove satisfactory. The essential point of a theory such as James’s is to treat calling something true as doing something more like praising it than like describing it.<sup>10</sup>

Five separable theses can be distinguished in the elaboration of this approach. First is the performative, antidescriptive strategy, emphasizing the *act* of calling something true rather than the descriptive content one thereby associates with what is called true. Next is an account of that act as the personal taking up of a certain sort of normative stance or attitude. Taking some claim to be true is endorsing it or committing oneself to it. Third is a particular understanding of that stance or attitude. Endorsing a claim is understood as adopting it as a guide to action, where this in turn is understood in terms of the role the endorsed claim plays in practical inference, both in first-person deliberation and in third-person appraisal. Fourth, and least important, is the view that an advantage of understanding the appropriateness or correctness of adopting an attitude of endorsement in terms of its role in guiding action consists in the possibility for some sort of not merely subjective measure of that appropriateness, namely, the success of the actions it leads to. This is the only strand of the argument acknowledged or embraced by stereotypical pragmatism.

Finally and, it will be argued, most significantly, the theory claims that once one has understood acts of *taking-true* according to this four-part model, one has understood all there is to understand about truth. Truth is treated, not as a property independent of our attitudes, to which they must eventually answer, but rather as a creature of *taking-true* and *treating-as-true*. The central theoretical focus is on what one is *doing* when one takes something to be true—that is, our use of ‘true’, the acts and practices of taking things to be true that collectively constitute the use we make of this expression. It is then denied that there is more to the phenomenon of truth than the proprieties of such takings. Theories of this general sort may be called ‘phenomenalist’, in recognition of the analogy with the paradigmatic subjective phenomenalist concerning physical objects, whose slogan was “*esse est percipi*.” The significance of such a move in the context of the present approach should be obvious, for *taking-true* is just asserting or judging. The classical pragmatist line of thought accordingly holds out the possibility of understanding the use of ‘true’ in terms of what we are doing when we make a claim, putting forward a sentence *as true*.

According to this decomposition of their view into five theory-features, the pragmatists start with the idea that in calling something true, one is *doing* something, rather than, or in addition to, *saying* something. Instead of asking what property it is that we are describing a belief or claim as having

when we say that it is true, they ask about the *practical significance* of the *act* we are performing in attributing that property. We accomplish many things by talking, and not all of them are happily assimilated to describing how things are. One ought not to conclude that because truth ascriptions are expressed in the same subject-predicate grammar that descriptions are, they must for that reason be understood to function as descriptions do. The pragmatic approach, centering on the act of calling something true rather than the content one thereby characterizes it as displaying, has much to recommend it. It has been seized upon by a number of authors who would not go on to accept the account of the act in question that the pragmatists offer. For, stripped of those further commitments, the recommendation is for a *performative* analysis of truth talk. In Fregean terms it is the suggestion that 'true' is a force-indicating, rather than a sense-expressing, locution.

Wittgenstein notoriously warned against thoughtless assimilation of sentence use to fact-stating, and of term use to referring. In the wake of Austin's discussions, theorists such as Strawson offered accounts of 'true' as a performative.<sup>11</sup> Its use was to be assimilated to other sorts of commitment-taking, in a way parallel to that expressed by the explicit performative 'I promise . . .' In the same spirit, other contemporary accounts were offered of 'good' as expressing a kind of commendation, as taking up an attitude or expressing one's own relation to something, rather than as describing it by attributing some objective property. This is the sort of assimilation James had been urging in saying that truth is "what is good in the way of belief." Such remarks are often misinterpreted as claiming descriptive equivalence, or coextensiveness of the predicates 'true' and 'what it is good for us to believe'. On such a reading, the allegedly uncontroversial claim "It is good for us to believe the truth,"—that is, the truth is among the things it is good for us to believe—is turned on its head. Necessary conditions are treated as sufficient, and truth is defined as whatever it is good for us to believe. James' intent was rather to mark off 'true', like 'good', as a term whose use involves the taking up of a nondescriptive stance, the undertaking of a commitment that has eventual significance for action. In the vocabulary of this work, it is adopting an implicitly normative practical attitude.

What motivates such a performative analysis, for the pragmatists no less than for later theorists, is the special relation that obtains between the force or practical significance of an act of taking-true (which one might, before the performative possibility has been broached, uncritically have called an act of 'describing as' true) and the force or significance of a straightforward assertion. In asserting "It is true that *p*," one asserts that *p*, and vice versa. The force or significance of the two claims is the same. On the face of it, this redundancy or transparency of force, the fact that adding the operator 'It is true that . . .' to what one is going to assert does not change the force or significance of that assertion, might be explained in either of two different ways. One might take it that the content that is expressed in a truth ascrip-

tion is special, and that the redundancy of force of truth claims arises out of features of the property a claim or belief is said to exhibit when it is described as true. One must then go on to offer an account of why attributing that property has the consequences that it does for the force of one's attribution. This can be compared to treating claims using 'good' or 'ought' as describing properties of actions, and then needing a theory explaining the special motivational role that attributions of these properties must be taken to have for the attributor. The pragmatic theories being considered adopt the more direct path of taking the transmitted force of truth claims as the central phenomenon, one that is merely obscured by the misleading grammar of property ascription.

Dewey's assertibilist theory of truth develops these ideas along explicitly performative lines using the model of utterances of 'I claim (or assert) that *p*'.<sup>12</sup> The claim that the force of freestanding utterances of this type and of 'It is true that *p*' are equivalent is especially liable to misinterpretation as the claim that the contents expressed by these utterances are the same. As will be seen below, it is easy to show that that is not so. In any case, as a revisionist, Dewey did not even claim equivalence of force, though that was the dimension along which he assessed the relationship between his views and the tradition. He has accordingly often been 'refuted' on the basis of misunderstandings of theories that he did not subscribe to in the first place.

To this performative, antidescriptive explanatory commitment, the pragmatists add a particular sort of account of the act of taking-true as adopting a *normative stance* toward the claim or belief. In treating something as true, one is praising it in a special way—endorsing it or committing oneself to it. The stance is normative in involving what the claim to which one has taken up a truth-attitude is good for or *appropriately* used for. For treating something as the truth is plighting one's troth to it, not just acknowledging that it has some property. Truth undertakings are taken to be personal in that the proprieties of conduct one thereby commits oneself to depend on one's other commitments, commitments to choose (representing preference, desire, interest, need, and so on) as well as commitments to say (assert and believe). One is expressing or establishing one's own relation to a claim, in taking it to be true, rather than recognizing some independent property that claim already had. Again the model of promising is important. This important emphasis on the normative character of cognitive undertakings was their acknowledgment of the central Kantian legacy that has been rejuvenated for us by Wittgenstein. Its expression is often obscured (Peirce is, as so often, an exception) by the pragmatists' further commitment to the sort of naturalism about the norms involved that gives rise to the attribution to them of stereotypical pragmatism.

Their understanding of the sort of commitment undertaken in taking-true is as a commitment to rely on the belief or claim in question in guiding practical activity. This in turn is understood as a commitment to using the

claim as a premise in practical inferences, whose conclusions are not further claims but actions—that is, performances under a description that is privileged by its relation to deliberation and appraisal. Relative to the truth-taking commitment, one ought to reason practically in one way rather than another. The proprieties of practical inference concerning whether to bring an umbrella are different for one who takes-true the claim that it is raining than for one who does not. The force of such proprieties is normative in that, although they may be ignored, the significance or force of the agent's commitment is to the effect that they *ought* not to be. It is these prudential 'ought's that appraisal of actions assesses.

Thus the stance or attitude that one adopts in treating something as true is to be understood by its role in orienting action when activated by a contextualized attempt to satisfy the desires, preferences, and so forth that one finds oneself with. The discussion of action in the previous chapter indicates that such a naturalistic, instrumental picture of the 'ought's involved in action and belief is not the only candidate. But the fact that the classical pragmatists go on to offer a naturalistic account of norms along these lines should not distract attention from their appreciation that the phenomenon of taking-true they address concerns *normative attitudes*, and that attributing truth is for them attributing a *normative status*.

Pragmatism in the stereotypical sense arises when one conjoins the ideas of a performative analysis of taking-true, of the relevant kind of performance as undertaking a personal commitment, and of the commitment as specifying the appropriate role of a claim in action-orienting deliberation, with the further idea that the measure of the correctness of the stance undertaken by a truth-attributor is the success of the actions it guides. Although the point will not be pursued here, it may be noted in passing that it is not easy to say what relation to the success of individual actions some substantive property of beliefs ought to stand in for it to play the sort of explanatory role that classical pragmatism envisions. It is by no means the case that for each belief in each deliberation, an agent is more likely to succeed in accomplishing the desired result if that belief is true than if it is false. The success of a plan can be thwarted by collateral false beliefs, and in the context of such beliefs, it may actually be helpful to have a further, compensating false belief. Someone who desires to cross a ravine and has formed the plan of felling a tree if one tall enough can be found near the edge will be more likely to succeed if equipped with the false belief, of the only likely candidate tree, that it is more than forty feet tall, than with the true belief that it is only thirty feet tall, in the case where that individual also believes falsely of the twenty-five-foot-wide ravine that it is forty feet across. Although the possibility of this sort of case is no reason to prefer in general to act on false beliefs, it does show that success is associated with a property of sets of beliefs, rather than individual beliefs. Where this point is acknowledged, the desired pragmatist principle is often formulated as the claim that "if all our beliefs were true, then the actions we undertook on those beliefs would satisfy our desires."<sup>13</sup>

But even this does not seem right. To ensure success, it would be necessary to ban not only error but ignorance, for facts of which the agent is unaware can lead to failure, even if all the agent's beliefs are correct. Of course this will not be true if those beliefs include one to the effect that there are no circumstances of which the agent is unaware that would thwart the plan being adopted. But including such beliefs threatens the principle with triviality in the same way that including among the beliefs stipulated to be true the belief that the plan will succeed would threaten it. So it is not clear how the vague pragmatist pronouncement that "the truth is what works" can be made more precise, without either falsifying or trivializing it. Fortunately, for the purposes of this chapter it is not necessary to resolve this issue.

### 3. *Phenomenalism*

Nonetheless, pragmatism is normally identified with the claim that the measure of the correctness of the stance undertaken by a truth-attributor is the success of the actions it guides. But the explanatory role played by this most notorious of the pragmatists' tenets ought to be understood in the light of the larger strategy for relating the concepts of truth and belief that it subserves. From a methodological point of view, perhaps the most interesting feature of the pragmatic approach is its commitment to phenomenism about truth.<sup>14</sup> Only in the context of a phenomenalist explanatory strategy can commitments of the first three sorts be seen as illuminating the notion of truth. For what they really supply is a theory of *taking-true*. It is in the overarching commitment to the effect that once one understands what it is to take or treat something as true, one will have understood as well the concept of truth that the phenomenism of this strategy consists.

The force-redundancy approach to truth emphasizes the practical equivalence of taking something to be true and believing it, so another way of putting the point is the following: instead of starting with a metaphysical account of truth, such as that of the correspondence theorists, in opposition to which the pragmatists defined themselves, and employing that in one's account of beliefs, which are then conceived as representations that could be true, that is, have the property previously defined, the pragmatists go the other way around. They offer an account of believing or taking-true, characterized by the three sorts of commitments already canvassed, that does not appeal to any notion of truth. Being true is then to be understood as being *properly taken-true* (believed). It is this idea that is built on here, jettisoning the details of the classical pragmatist account of belief or taking-true, and substituting for it the account of assertion and doxastic discursive commitment introduced in Chapter 3.

From this point of view, what is of most interest about the classical pragmatist stories is not stereotypical pragmatism but the dual commitment to a *normative* account of claiming or believing that does not lean on a supposedly explanatory antecedent notion of truth, and the suggestion that

truth can then be understood phenomenalistcally, in terms of features of these independently characterized *takings*-true. The sort of explanatory strategies here called 'phenomenalist' in a broad sense treat the subject matter about which one adopts a phenomenalistc view as *supervening* on something else, in a way whose paradigm is provided by classical sensationalist phenomenalistc about physical objects. The slogan of this narrower class of paradigmatically phenomenalist views is, "To be is to be perceived." The characteristic shift of explanatory attention enforced by these approaches is from what is represented to representings of it. The representeds are explained in terms of the representings, instead of the other way around. Talk ostensibly about objects and their objective properties is understood as a code for talk about representings that are interrelated in complicated but regular ways. What the naive conservatism implicit in unreflective practice understands as objects and properties independent of our perceptual takings of them now becomes radically and explicitly construed as structures of or constructions out of those takings. Attributed existence, independence, and exhibition of properties are all to be seen as features of attributings of them.

The general structure exhibited by this sort of account is that the facts about *having* physical properties are taken to supervene on the facts about *seeming* to have such properties. Or in the vocabulary to be preferred here, the facts about what things *are Ks*, for a specified sortal *K*, supervene on the facts about what things are *taken* to be *Ks*. According to such an explanatory strategy, one must offer first an independent account of the takings—one that does not appeal in any way to what it is to be a *K* in order to explain what it is to take something to be one. Thus classical phenomenalistc concerning physical properties such as *red* found itself obliged to account for states of the attributing subject in which things *look-red* or *seem-red* without invoking the redness that is attributed in such takings. Once that obligation is satisfied, it can further be claimed that there are no facts about what things are red, or what it is for things to be red, over and above all the (possible) facts about what things look or seem red.

Classical subjective phenomenalistc regarding physical objects and properties notoriously failed in both of these component explanatory tasks. Cartesian mental acts seemed ideal candidates for the takings in question. This ontological category had been given an epistemic definition in terms of the privileged access (in the sense of transparency and incorrigibility) subjects have to the class of takings that includes perceptual seemings. That something could not *seem* red to a subject who did not by virtue of that very taking *know* that it seemed red, and that something could not merely *seem* to seem red without *really* seeming red, made this class of takings appear well suited to provide the independently characterized base of a supervenience relation. Their special epistemic status seemed to guarantee for these subjective takings or attributings the possibility of a characterization independent of what they take or attribute. For one knows all about these states

just by having or being in them, apart from any relation to anything but the knowing subject and the known mental state.

#### 4. 'Looks' Talk and the Errors of Subjective Phenomenalism

But this way of understanding 'looks' is a mistake. As various authors have shown (the locus classicus is Sellars's "Empiricism and the Philosophy of Mind"), 'looks'-talk does not form an autonomous stratum of language: a game one could play, though one played no other. When one understands properly how the 'seems' operator functions, one sees that the incorrigibility of such claims essentially arises from their *withholding* of the endorsements involved in unqualified claims about how things actually are. As was indicated already in Chapter 4, noninferential reporting practices involve two distinguishable components: a reliable differential responsive disposition and the capacity to exercise that disposition by *endorsing* a claim. It is this second component that puts the response into the inferentially articulated space of applications of *concepts*, by bringing it into the game of giving and asking for reasons. Sellars's account of 'looks' or 'seems' is that these locutions are introduced *after* such a practice is under way, as a way of dealing with systematic sources of perceptual error that one becomes aware of through having to withdraw unreflective noninferential claims on the basis of their incompatibility with commitments one is otherwise entitled to (noninferentially, inferentially, or by testimony). One can then *express* the responsive disposition to call something red—a disposition to endorse a claim (undertake a doxastic commitment) if one did not know better (have reason to think that commitment is liable to be flawed)—without actually endorsing the claim one is acknowledging a certain temptation to endorse. This is how one begins to become reflective about the implicit appropriate circumstances of reporting. Thus in saying how things *look*, one is withholding an endorsement and so is naturally incorrigible.

In arguing for this diagnosis, Sellars points to two sorts of uses of 'seems' or 'looks' that it makes sense of and that are unintelligible if one construes them as minimal, incorrigible-because-unconceptualized reports on immediate experience. One is *generic* looks claims: the hen looks to have a number of spots, but there is no particular number of spots she looks to have; the polygon looks many-sided, but it does not look to be 998-sided, and does not look to be 999-sided, and so on. For Sellars, these cases are matters of what commitment one is prepared to endorse and defend. One can be willing to endorse the claim that an object is red without being willing to go further and endorse the claim that it is scarlet—and similarly for all the other shades of red. One withholds endorsement from the more specific claim, while endorsing the generic one. On traditional accounts, one must either jettison the transparency and incorrigibility of our knowledge of our own experiences or sense data or else treat these things as unlike ordinary objects in some-

times being *merely* generic. Such strenuous metaphysical exertions are obviated by Sellars's analysis.

The other use he points to is the difference between:

It looks as though there is a tree over there.

There is something that looks to be a tree over there.

There is a tree over there.

The different scope of the 'looks' operator corresponds to the different scope of the commitments undertaken and withheld. In the first case, one endorses nothing about the claim, merely evincing the noninferential disposition to apply concepts one in the event refuses to apply. In the second case, one endorses the existence of something over there and evinces one's temptation to call it a tree, while resisting that temptation and not endorsing that characterization. In the third case, one endorses the whole claim. One is 'incorrigible' exactly as far as one withholds endorsement. But of course one cannot withhold endorsement unless one can grant it, and to do that one must be in the game of making ordinary corrigible noninferential reports. The very incorrigibility that recommended 'seems' statements as a basis in terms of which everything epistemically less certain could be understood turns out to be an expression of the parasitic relation that these withholdings of endorsement have to the risky practices of endorsement from which they derive their meaning, by contrast to which they exhibit their special status. Whatever may be their role in the order of *justification*, in the order of *understanding* 'seems-red' presupposes 'is-red'.

It may be noted in passing that Sellars's analysis applies to first-person uses of 'looks' (or more generally 'seems'). His endorsement-withholding diagnosis is confirmed by consideration of third-person, or attributing, uses. For in that case an ambiguity arises: the claim "It (merely) seems to *S* that *p*" may be used either to attribute *S*'s withholding of an endorsement (the state *S* would express in propria persona by uttering "It seems to me that *p*") or to attribute the undertaking of a commitment to *S* (the state *S* would express by claiming that *p*) while expressing the *attributor's* withholding of endorsement of the claim that *p*. In this sense, 'seems' is used to express a hybrid deontic state dual to that expressed by 'true': the first corresponds to "*S* *incorrectly* believes that *p*," and the second to "*S* *correctly* believes that *p*"—in the sense of 'correct' that has to do with commitment, rather than entitlement.

Although Sellars does not say so, an entirely parallel analysis applies on the side of noninferential practical discursive *exits*. Just as a doxastic commitment whose success one is not prepared to endorse can be denominated a mere *seeming*, so a practical commitment whose success one is not prepared to endorse can be denominated a *trying*. Just as while it can seem to one that something is red without its actually being red, but it cannot merely seem to one that it seems to one to be red without its actually seeming so,

so while one can try to lift the weight without actually lifting it, one cannot merely try to try to lift it, without actually trying to do so. Neither seemings nor tryings iterate nontrivially. Just as the first point can be put (with substantial danger of misleading in a Cartesian direction) by saying that the distinction between appearance and reality does not apply to appearances, so the second can be put (subject to the same danger) by saying that the distinction between willing and succeeding does not apply to willings. The danger is that one will be tempted to think of seemings as cognitions about which one cannot be in error, and of tryings as actions one cannot fail to accomplish. The mind can then be thought of as consisting in these two sorts of representings that cannot fall short of what is represented (which in turn can be understood only by the coincidence in this case of representing and represented): states such that just by having them one counts on the one hand as knowing something and on the other as doing something. On this rendering, the mind extends just as far as its total cognitive and practical dominion; its limits are scribed by the inapplicability of the concepts of error and failure.

The cost of treating these degenerate cases as representational paradigms is to render unintelligible in the ordinary fallible cases the relation between doxastic or practical representings and the represented states of affairs known or brought about by them. These temptations are best avoided by correctly diagnosing the source of the noniterability of the 'seems' or 'tries' operator, which is the phenomenon that originally motivates this disastrous metaphysics of the mental. For then grasp of what is expressed by both 'seems' and 'tries' talk is seen to depend on grasp of what is expressed in ordinary fallible 'is' talk; one cannot withhold endorsements one cannot undertake or attribute, and a further disavowal of an endorsement once disavowed is without effect. So understood, neither the cognitive infallibility of seemings nor the practical infallibility of tryings is eligible to serve a foundational role. A subject conceived as contracted to these activities alone cannot be coherently thought of as grasping or accomplishing anything, hence not as a subject at all.

Because in the order of understanding, grasp of what is expressed by concepts of the form 'seems-*K*' presupposes grasp of what is expressed by corresponding concepts of the form 'is-*K*', the classical phenomenalist basis of takings as subjectively certified seemings cannot be secured with the autonomy from the properties taken to be exhibited that is requisite for the subsequent framing of phenomenalist supervenience explanations. Those explanations have troubles, however, apart from those regarding their basis. Generic phenomenalist has been characterized here in terms of supervenience. The sense intended is that one vocabulary supervenes on another just in case there could not be two situations in which true claims (that is, facts) formulable in the supervening vocabulary differed, while the true claims formulable in the vocabulary supervened on do not differ. More neutrally (but in the end equivalently) put, once it is settled what one is committed to as

expressed in the one vocabulary, then it is settled what one is committed to as expressed in the other.

Classical subjective phenomenism about physical objects and properties typically makes stronger, reductionist claims, involving further commitments beyond supervenience. These regard the equivalence of sentences (or in the most committive cases, individual terms and predicates) in physical object talk to sentences (or terms and predicates) constructible in the language of takings-as-seemings. Again, the sort of equivalence in question might vary from the extreme of definitional or translational equivalence down to mere coextensiveness. In none of these forms are phenomenalist claims of this reductionist variety plausible today (as Quine argues in "Epistemology Naturalized").<sup>15</sup> Attempts to work out these reductive phenomenalist strategies have shown that the conditions under which there are reliable connections between how things seem perceptually and how they are can themselves be stated only in terms of how things are. The inference from things seeming red to their being red depends on there being *in fact* no filters, strange lights, retina-altering drugs, and so on. That there not seem to be such is far from sufficient.<sup>16</sup>

These explanatory failures of phenomenism in the narrow sense ought not to be taken to impugn the prospects of phenomenalist strategies in the broad sense. For those difficulties arise from the way its general phenomenalist commitments are specialized: in applying to perceivable physical properties, in offering an account of the relevant sort of takings as incorrigible subjective perceptual seemings, and in insisting on reduction rather than just supervenience as the relation between them. Phenomenalism in general is a structure that antirealist accounts of many different subject matters may exhibit. It elaborates one way of taking seriously what Dummett calls the issue of "recognition transcendence." To detail a specific version of this sort, one must specify three things: what it is that one is taking a phenomenalist approach to (for example, physical objects, mental activity, semantic properties, the past, and so forth), how one conceives the takings or attributings on which talk of such things is taken to supervene, and how in particular the supervenience relation is conceived. Corresponding to each specific phenomenalist claim of this sort will be a class of claims that qualify as *realist* in the sense of denying the phenomenalist's "nothing but" account of the subject matter in question. For the classical pragmatist the facts about what is true supervene on the facts about taking-true, that is, on the actual action-guiding roles of beliefs. In order to appreciate the significance of the pragmatists' phenomenalist strategy, one must first consider the development of the basic idea that truth locutions are force-indicating, rather than content-specifying. The subsequent trajectory of this idea will be reconstructed, before returning to the issue of phenomenalism about truth. In the end, the phenomenalism of the classical pragmatists about truth will appear as a special case of the way in which normative statuses have been taken to be instituted by

normative attitudes, so that talk of commitments can be traded in for talk of undertaking and attributing commitments.

Before going on to see what is wrong with pragmatic phenomenism about truth, it is worthwhile to recall briefly what can make it attractive. The account of knowledge claims offered in Chapter 3 depends on this approach. It can now be seen that that account should be seen as propounding a certain kind of phenomenism about knowledge, which depends on a corresponding phenomenism about truth. Its primary focus is not on knowledge itself but on *attributions* of knowledge, attitudes toward that status. The pragmatist must ask, What are we *doing* when we say that someone knows something? According to a phenomenalist reconstruction of the classic justified-true-belief account of knowledge (developed above in 4.1.2), in taking someone to know something, one first of all *attributes* a *commitment*, that is, takes someone to believe. One further *attributes entitlement* to that commitment, that is, takes the committed subject to be justified.

What, then, is the function of the truth condition on knowledge? Conventionally, treating taking the claim that the subject is committed to as true is understood as attributing some property to it, characterizing it or describing it. But it has already been pointed out that the pragmatist's account of taking the claim to be true is as *acknowledging* or *undertaking* a commitment to it. The truth condition does not qualify the entitled commitment that is attributed but simply indicates that the attributor of knowledge must endorse it. This is a deontic attitude that differs in its *social perspective*. Attributions of knowledge have the central linguistic status that they do because in them commitment to a claim is *both* attributed and undertaken. This phenomenalist distinction of social perspective, between the act of attributing and the act of undertaking a commitment, is what is *mistaken* for the attribution of a descriptive property (for which an otiose metaphysics then appears to be required).

A pragmatic phenomenalist account of knowledge will accordingly investigate the social and normative significance of acts of attributing knowledge. The account of taking-true being considered is what makes possible such a way of thinking about knowledge claims. The account of asserting introduced in Chapter 3 can be extended to such an account of knowledge, via the notion of knowledge *claims*, only if something can be made of the phenomenalist strategy of classical pragmatism about truth. There are some serious obstacles, however, to pursuing this strategy.

### 5. Embedded Uses of 'True' and the Failure of the Pragmatists' Analysis

On the pragmatic line being considered, it is the practical significance or force of asserting that defines taking-true, and this sense of

taking-true accounts for our use of 'true'. In spite of all that there is to recommend such a hypothesis, this conjunctive thesis cannot be correct as it stands. A familiar point of Frege's shows the inadequacy of the basic pragmatic claim. Truth talk cannot be given a purely pragmatic rendering because not all uses of '... is true' have assertoric or judgmental force. The force-based approach can at most account for a subset of our uses of truth locutions. Frege drew attention to the use of sentences as components of other sentences. Assertion of a sentence containing another sentence as a component is not in general assertion of the embedded sentence. That is, the embedded sentence does not occur with assertional force, does not express something the assertor of the containing sentence is thereby committed to.

As the antecedent of a conditional, 'It is true that *p*', for instance, cannot have the significance of a taking-true if that is understood as the expressing of assertional force. In this sense one does not take-true the claim that *p* in asserting, 'If it is true that *p*, then it is true that *q*'. Of course this is just the point that was urged in Chapter 3 against traditional accounts that identify assertion as a kind of predication or representation. In those cases as in this one, the embedded uses show that what is at issue has to do with the content expressed, rather than the force attached to the speech act in which that content is expressed.

The pragmatic approach, then, offers an account only of the freestanding uses of sentences formed with '... is true', not the embedded ones. This is the same rock on which, as Geach has shown,<sup>17</sup> performative accounts of the use of 'good' have foundered. It is precisely because one cannot embed, say, questions and imperatives as antecedents of well-formed conditionals (in which they would occur without their characteristic force) that their significance as askings and commandings is associated with their force—and so is not to be understood as a feature of the descriptive content they express. If the essence of calling something good consisted in doing something rather than saying something, then it should not be possible to say things like, "If that is good, then one ought to do it." That one can sensibly say things like this shows that 'good' has descriptive content that survives the stripping away by embedding of the force associated with freestanding describings.

So an embedding test can be treated as criterial for broadly descriptive occurrences of expressions. According to this test, 'It is true that *p*' has nonperformative uses that the pragmatists' approach does not account for. And it is not open to the pragmatist simply to distinguish two senses of truth claims, one freestanding and the other embedded, and proceed from ambiguity. For on such a line one would be equivocating in inferring from the freestanding 'It is true that *p*' and the conditional 'If it is true that *p*, then it is true that *q*', in which it occurs embedded, that it is true that *q*, by detachment. So the pragmatic theory must be rejected and the phenomena it points to otherwise explained.

This sort of objection surfaces in many forms. Those who incorrectly take

Dewey to have offered an analysis of 'true', rather than a candidate replacement notion, must thereby treat his assertibilism as the assertion of an equivalence of content between the sentences 'It is true that *p*' and the explicit performative 'I (hereby) claim that *p*'. The most such a made-up thinker would be entitled to claim is that the force of the freestanding utterance of these sentences is the same. The stronger theory is refuted by noticing that 'It is true that *p*' and 'I claim that *p*' behave differently as embedded components. For instance, they are not intersubstitutable as the antecedents of conditionals, saving the inferential role of the resulting compound.

So an account such as is often attributed to Dewey is subject (as Putnam has pointed out in different terms)<sup>18</sup> to a version of Moore's naturalistic fallacy argument. Not everyone who is committed to the conditional 'If it is true that *p*, then it is true that *p*' is committed also to the conditional 'If I claim that *p*, then it is true that *p*'. If we like, we can put this point by saying that there is nothing self-contradictory about the claim 'It is possible that I claim that *p*, and it is not true that *p*'.<sup>19</sup> The naturalistic fallacy point is thus just another way of putting the objection from embedding.

### III. FROM PRAGMATISM TO PROSENTENCES

#### 1. *Redundancy*

Pointing to the sentential embedded uses of '... is true' shows the inadequacy of the pragmatists' attempt to make do with a notion of taking-true as asserting. Analyzing and identifying uses of truth locutions by means of *redundancy of force* (that is, by a formal property of the pragmatic significance of acts of asserting freestanding truth claims) is not a sufficient explanatory strategy. It is not that freestanding force redundancy is not a central phenomenon of truth talk. But not all uses of truth locutions take this form. More is required for an account of the use of 'true' than can be provided simply by an account of taking-true as asserting or undertaking an assertional commitment. The pragmatic account cannot for this reason be the whole truth.

Rather than simply discarding that approach, it is possible to amend it so as to retain the pragmatic account for the freestanding uses to which it properly applies. For there is a more general redundancy view that has the force redundancy of freestanding truth-takings as a consequence. Embedded uses can be explained by a notion of *redundancy of content* according to which (apart from niceties having to do with type/token ambiguities) even in embedded contexts 'It is true that *p*' is equivalent to *p*. For even their embedded occurrences are equivalent as antecedents of conditionals, in the sense that anyone who is committed to 'If it is true that *p*, then *q*' is thereby committed to 'If *p* then *q*' and vice versa. Furthermore, intersubstitutability

of ‘it is true that *p*’ and *p* in *all* occurrences, embedded or not, is sufficient to yield force redundancy in freestanding uses as a consequence. If two asserted contents are the same, then the significance of asserting them in the same pragmatic context should be the same. On such a content redundancy view, the pragmatists have simply mistaken a part for a whole.

Redundancy views such as Ramsey’s accordingly provide a generalization of the pragmatist’s point, one that permits an answer to the otherwise decisive refutation offered by the embedding objection. Accounts that generalize to the intersubstitutability of ‘Snow is white’ and ‘It is true that snow is white’ are clearly on the right track. They show what is needed to supplement the pragmatists’ account in order to deal with some embedded occurrences. But they do not yet account for all the contexts in which the taking-true locution ‘... is true’ occurs. Such simple redundancy accounts will not offer a correct reading of sentences like ‘Goldbach’s conjecture is true’. For this sentence is not interchangeable with ‘Goldbach’s conjecture’. For instance, the former, but not the latter, appears as the antecedent of well-formed and significant conditionals. So content redundancy, while relaxing the limitations constraining the original pragmatic account, will not apply correctly in all the contexts in which truth locutions occur.

## 2. *Disquotation*

Such cases show that the content redundancy view must in turn be revised to include the operation of some sort of disquotation or unnominalizing operator. In the cases to which the simple content redundancy theory applies, the additional operation will be transparent. But in the case of sentences such as ‘Goldbach’s conjecture is true’, the claim with respect to which the truth-taking is content redundant must be determined by a two-stage process. First, a sentence nominalization is discerned. This may be a description such as ‘Goldbach’s conjecture’, a quote-name such as ‘Snow is white’, a ‘that’-clause sortal such as ‘the claim that snow is white’, or other sort of nominalization. Next, a sentence is produced that is nominalized by the locution picked out in the first stage. This is a sentence expressing Goldbach’s conjecture, named by the quote-name, one which says *that* snow is white, and so on. It is this sentence that is then treated by theory as intersubstitutable with the truth-attributing sentence, whether occurring embedded or freestanding.

A content redundancy account with disquotation or unnominalization is more satisfactory and deals with more cases than a simple content redundancy account does—just as content redundancy accounts represent improvements of theories acknowledging only redundancy of force. But even disquotational views will not account for all the uses of ‘... is true’ that might be important. They will not deal correctly, for instance, with occurrences such as ‘Everything the policeman said is true’, in which a *quantified*

sentence nominalization is employed. For here what is nominalized is a whole *set* of sentences, and there need in general be no single sentence that is equivalent to all of them. A further refinement of content redundancy accounts is required if they are to be able to deal with this range of cases.

### 3. Anaphora and Prosentences

The most sophisticated version of the redundancy theory, one capable of handling quantificational truth idioms, is the remarkable anaphoric analysis undertaken by Grover, Camp, and Belnap in their essay "A Prosential Theory of Truth."<sup>20</sup> For the original redundancy and disquotational theories, each use of '*...* is true' is associated with some sentence on which it is redundant or with which it shares its content. Whatever else this may mean, it at least includes a commitment that the intersubstitution of the sentence containing 'true' and its nonsemantic equivalent, in some privileged range of contexts, preserves assertional and inferential commitments.

The difficulty in extending this intersubstitutional account to the quantificational case is that there the use of the sentence containing 'true' is determined not by a single sentence but by a whole *set* of sentences, those expressing whatever the policeman has said. Of course disquotation or unnominalization may produce sets of sentences as well, as more than one sentence may express Goldbach's conjecture. But in such cases the sentences must all share a content or be redundant on each other—that is, must be intersubstitutable with each other in the relevant contexts—whereas there is no requirement that any two sentences that express things the policeman has said be in any other way equivalent. So what is it that the sentence containing 'true' shares its content with, or is redundant upon in the sense of intersubstitutability? What is distinctive of the anaphoric development of redundancy theories is its use of the model of *pronouns* to show how, in spite of this difficulty, the quantificational cases can be treated both as redundant in the same way nonquantificational cases are and as deriving their content from a whole *set* of nonintersubstitutable sentences.

It has been noticed that pronouns serve two sorts of purposes.<sup>21</sup> In the *lazy* use, as in 'If Mary wants to arrive on time, she should leave now', they are replaceable by their antecedents, merely avoiding repetition. In the *quantificational* use of pronouns, as in 'Any positive integer is such that if it is even, adding it to one yields an odd number', such replacement clearly would change the sense. 'If any positive number is even, adding any positive number to one yields an odd number' is not a consequence one becomes committed to by undertaking the original claim. In such cases, the semantic role of the pronoun is determined by a set of admissible substituends, which is in turn fixed by the grammatical antecedent (here 'any positive number'). Asserting the original sentence commits one to each of the results of replacing the

pronoun ‘it’ in some occurrence by some admissible substituend, that is, some expression that refers to a positive number.<sup>22</sup>

The prosentential theory of truth is what results if one decides to treat ‘... is true’ as a syncategorematic fragment of prosentences and then understands this new category by semantic analogy to other proforms, in particular to pronouns functioning as just described. So ‘Snow is white is true’ is read as a prosentence of laziness, having the same semantic content as its anaphoric antecedent, perhaps the token of ‘Snow is white’ that it contains. The prosentence differs from its antecedent in explicitly acknowledging its dependence upon an antecedent—as ‘She stopped’ differs from ‘Mary stopped’ when the pronoun has some token of the type ‘Mary’ as its antecedent. Otherwise, the lazy uses are purely redundant.<sup>23</sup> The advance on earlier conceptions lies in the availability on this model of *quantificational* uses of prosentences containing ‘true’. ‘Everything the policeman said is true’ is construed as containing a quantificational prosentence, which picks up from its anaphoric antecedent a set of admissible substituends (things the policeman said). Expanding the claim in the usual way, to ‘For anything one can say, if the policeman said it, then it is true’, explicitly exhibits ‘it is true’ as the quantificationally dependent prosentence. Each quantificational instance of this quantificational claim can be understood in terms of the lazy functioning of prosentences, and the quantificational claim is related to those instances in the usual conjunctive way.

By analogy to pronouns, prosentences are defined by four conditions<sup>24</sup>:

1. They occupy all grammatical positions that can be occupied by declarative sentences, whether freestanding or embedded.
2. They are generic, in that *any* declarative sentence can be the antecedent of some prosentence.<sup>25</sup>
3. They can be used anaphorically either in the lazy way or in the quantificational way.
4. In each use, a prosentence will have an anaphoric antecedent that determines a class of admissible sentential substituends for the prosentence (in the lazy case, a singleton). This class of substituends determines the significance of the prosentence associated with it.

There are many philosophical virtues to explicating each occurrence of ‘true’ as marking the use of a prosentence in this sense. Quite varied uses, including embedded ones, of expressions involving ‘true’ in English are accounted for by means of a unified model. That model is in turn explicated by appeal to the familiar and closely analogous pronominal anaphoric reference relation. Not only is the semantics of such uses explained, but their pragmatic features are as well—namely, acknowledgment of an antecedent and the use of truth locutions to endorse or adopt someone else’s claim. Tarski’s biconditionals are appropriately underwritten, so the necessary condition of adequacy for theories of truth that he establishes is satisfied.

A feature dear to the hearts of the prosententialists is the metaphysical

parsimony of the theory. For what in the past were explained as attributions of a special and mysterious *property* (truth) to equally mysterious bearers of truth (propositions) are exhibited instead as uses of grammatical proforms anaphorically referring only to the sentence tokenings that are their antecedents. A further virtue of the prosentential account is that anaphora is a relation between *tokenings*. Consequently the use of tokenings of types such as 'That is true' as a response to a tokening of 'I am hungry' is construed correctly—just as 'he' can have 'I' as its antecedent without thereby referring to whoever uttered 'he'. An incautiously stated content redundancy theory would get these indexical cases wrong. Finally, the uses of 'true' falling under the elegant, anaphorically unified treatment include quantificational ones such as 'Everything the oracle says is true', which are recalcitrant to more primitive redundancy and disquotational approaches.

The classical pragmatists' insistence that in calling something true, one is not describing it is respected. For one does not describe a cat when one refers to it pronominally by means of an 'it'. This point is further broadened to accommodate embedded uses where the account of the describing alternative as endorsing does not (as emerged above) apply. 'True' functions anaphorically and not descriptively even in such cases. And anaphoric inheritance of content explains equally why freestanding or force-bearing uses of 'It is true that *p*' have the pragmatic significance of endorsements of the claim that *p*. The prosentential account shows how the pragmatists' insights can be preserved, while accounting for the uses of 'true' that cause difficulties for their original formulation. It is accordingly a way of working out the content redundancy rescue strategy.

#### 4. Prosentence-Forming Operators

The treatment of quantificational prosentences represents an advance over previous redundancy theories. As the theory is originally presented, however, the treatment of lazy prosentences in some ways retreats from the ground gained by disquotational developments of redundancy theories. The explanatory costs associated with the original theory arise because it treats *most* occurrences of 'true' as quantificational. Thus the official version of 'The first sentence Bismarck uttered in 1865 is true' construes it as a quantified conditional of the form 'For any sentence, if it is the first sentence Bismarck uttered in 1865, then it is true', in which 'it is true' is a prosentence of quantification.

One of the strengths of the prosentential account is its capacity to use the logical structure of quantification to explain the use of complicated sentences such as 'Something John said is either true or has been said by George'. There should be no quarrel with the author's treatment of these sentences that "wear their quantifiers on their sleeves." And it is clear that any sentence that has the surface form of a predication of truth of some sentence nominalization can be construed as a conditional propositional quantifica-

tion. But it is not clear that it is a good idea to assimilate what look like straightforward predication of truth to this quantificational model. To do so is to reject the disquotational treatment of these lazy prosentences, which has no greater ontological commitments and stays closer to the apparent form of such sentences. Otherwise almost all sentences involving ‘true’ must be seen as radically misleading as to their underlying logical form. The account of truth talk should bear the weight of such divergence of logical from grammatical form only if no similarly adequate account can be constructed that lacks this feature. It would be preferable to follow the treatment of sentence nominalizations suggested by disquotational generalizations of redundancy theories.

In fact there is no barrier to doing so. The original motivations of the prosential account carry over directly to a disquotational or unnominalizing variant. According to such an account, ‘The first sentence Bismarck uttered in 1865’ is a sentence nominalization, a term that picks out a sentence tokening. In this case it describes the sentence, but it could be a quote-name, demonstrative, ‘that’-clause sortal, or any sort of nominalization. Its function is just to pick out the antecedent on which the whole prosentence formed using ‘true’ is anaphorically dependent, and from which it accordingly inherits its content. Ontological commitment is just to sentence tokens and to anaphoric dependence, which prosentialists require in any case.

A brief rehearsal of the considerations leading the authors of the prosential theory to do things otherwise will show that their reasons ought not to count against the adoption of a disquotational variant of the prosential account. They say: “This account differs radically from the standard one since on (what we have called) the subject-predicate account ‘that’ in ‘that is true’ is always treated separately as referring by itself to some bearer of truth, whether it be a sentence, proposition, or statement. On our account cross-referencing—without separate reference of ‘that’—happens between the *whole* expression ‘that is true’ and its antecedent.”<sup>26</sup> Another way to put this point is that where the classical account takes a subpart of the sentence as a referring term and takes ‘... is true’ as a predicate that forms a sentence from that term by characterizing its referent, according to the prosential theory the only expression standing in a referential relation is the whole sentence, which refers anaphorically to an antecedent. There are accordingly two innovations put forth concerning reference in sentences like ‘that is true’. The sentence is seen as an anaphoric proform and ‘that’ is no longer seen as a referring term. “Reference can involve either (or both) anaphoric reference or independent reference, and since people have not seriously considered the former, the possibility that the relation between ‘that is true’ and its antecedent may be that of anaphoric reference has not occurred to them. In ignoring anaphoric reference philosophers have assumed that the reference involved in ‘that is true’ is, through ‘that’, like that between a pronoun (say ‘she’ used

independently) and its referent (say Mary). Once this picture dominates, the need for bearers of truth begins to be felt, and it is then but a small step to the claim that in using ‘is true’ we are characterizing those entities” (emphasis added).<sup>27</sup>

But why should one have to choose between, on the one hand, treating the whole expression ‘that is true’ as a prosentence anaphorically referring to a sentence tokening from which it inherits its content and, on the other hand, treating ‘that’ as a referring expression (in particular a sentence nominalization) that picks out the tokening on which the whole prosentence depends? Instead of seeing ‘. . . is true’ as a syncategorematic fragment of a semantically atomic generic prosentence ‘that is true’, one can see it as a *prosentence-forming operator*. It applies to a term that is a sentence nominalization or that refers to or picks out a sentence tokening. It yields a prosentence that has that tokening as its anaphoric antecedent. To take such a line is not to fall back into a subject-predicate picture, for there is all the difference in the world between a prosentence-forming operator and the predicates that form ordinary sentences. Nor does it involve commitment to bearers of truth, apart from the tokenings that play the role of sentential antecedents, which no anaphoric account can do without.

#### IV. REFERENCE AND ANAPHORICALLY INDIRECT DESCRIPTIONS

##### 1. *From ‘True’ to ‘Refers’*

There is a further reason to prefer the account that treats ‘. . . is true’ as a prosentence-forming operator as here recommended, rather than as a fragment of the single prosentence, ‘that is true’ (functioning almost always quantificationally), as the original theory has it. Conceived in the former way, the treatment of ‘true’ has an exact parallel in the treatment of ‘refers’. ‘Refers’ can be understood as a pronoun-forming operator. Its basic employment is in the construction of what may be called *anaphorically indirect definite descriptions*. These are expressions such as ‘the one Kissinger referred to [represented, described, talked about] as “almost a third-rate intellect”’, understood as a pronoun whose anaphoric antecedent is some utterance by Kissinger. A full-fledged pronominal or anaphoric theory of ‘refers’ talk can be generated first by showing how other uses of ‘refers’ and its cognates can be paraphrased so that ‘refers’ appears only inside indirect descriptions and then by explaining the use of these descriptions as pronouns formed by applying the ‘refers’ operator to some antecedent-specifying locution.<sup>28</sup>

Treating ‘true’ as an operator that applies to a sentence nominalization and produces a prosentence anaphorically dependent upon the nominalized sentence token, and ‘refers’ as an operator that applies to an expression picking out a term tokening and produces a pronoun anaphorically dependent

upon it then permits a single theory form to explain the use of all legitimate semantic talk about truth *and* reference in purely anaphoric terms. Discussion of how anaphoric dependence is itself to be understood must be postponed until Chapter 7, when all the raw materials will have been assembled. In what follows, then, a construction is presented that is intended to account for the use of expressions containing 'refers' and its cognates in natural languages, in a fashion strictly analogous to that just rehearsed for 'true'. To do so, however, it will be necessary to look a little more closely at the phenomenon of pronominal anaphora than was done in introducing prosentences, even though the official account of the sort of commitment inheritance that underlies that phenomenon cannot yet be considered.

Reference comes in two flavors: word-world, or extralinguistic reference of the sort invoked when it is said that the phrase 'the author of *Dreams of a Spirit-Seer*' refers to a certain actual individual (namely Kant), and word-world intralinguistic or anaphoric reference instanced by pronouns such as 'he' in "Wittgenstein admired Frege's work, so he traveled to Jena to talk with the great man." Intralinguistic reference of this sort has not been of much interest to philosophers (as opposed to linguists), for it has seemed natural for semantic purposes to assimilate pronouns to bound variables, and so to expect to explain anaphoric reference as grammatically guaranteed coreference. This coreference is in turn to be understood in terms of the primary, word-world sense of reference to the same extralinguistic item. It will be argued in Chapters 7 and 8 that this is a mistake. For one thing, anaphoric mechanisms are more fundamental than, and are presupposed by, deictic mechanisms—one cannot have a language with indexicals such as demonstratives but without expressions functioning anaphorically, though the converse is possible. For another, the explicitly representational locutions by means of which we grasp and express the distinction between what our thought and talk is *about* and what we think and say about it all depend on anaphoric mechanisms.

The point of this section is to show how an analysis in terms of anaphoric mechanisms can provide the resources for a purely *intralinguistic* account of the use of the English sentences by means of which philosophers make assertions about *extralinguistic* referential relations. More specifically, although we can and must distinguish between our words and what the words refer to or have as their referents, the truth of claims about what we are referring to by various utterances is not to be understood in terms of a *relation* of reference between expressions and the objects we use them to talk about. Following Sellars,<sup>29</sup> it will be argued that 'refers' not be semantically interpreted by or as a relation and, a fortiori, not a word-world relation. Instead, 'refers' will be explained as a complex anaphoric pronoun-forming operator, by analogy (in the category of terms) to the analysis of '... is true' offered above. To show this, the anaphoric roles that expressions can play are botanized. This leads to the specification of a new part of speech—*anaphori-*

*cally indirect descriptions.* Next, a formal test is offered for identifying expressions that play the anaphoric role of indirect descriptions, and 'refer' is explained as an operator that forms such descriptions. A paraphrase strategy is then offered by means of which reference claims ostensibly in other forms can be wrestled into forms in which 'refer' appears only inside indirect descriptions.

## 2. Anaphoric Chains

In his seminal article "Reference and Context," Charles Chastain suggests a novel approach to the understanding of singular term reference.<sup>30</sup> The basic concept he employs is that of an *anaphoric chain*, a notion best approached at this point in terms of examples. Consider the discourse:

#A man in a brown suit approached me on the street yesterday and offered to buy my briefcase. When I declined to sell it, the man doubled his offer. Since he wanted the case so badly, I sold it to him.#

Two anaphoric chains are intertwined here, one corresponding to the buyer, and one to the briefcase:

A man in a brown suit . . . the man . . . he . . . him

and

my briefcase . . . it . . . the case . . . it.

The phenomenon may be indicated presystematically by saying that the reference of later elements in such chains (here 'it' and 'the man') is secured only by the relations these elements stand in to the singular terms that initiate the chains in which they appear. This is the word-word (token-to-token)<sup>31</sup> relation of *anaphoric reference* or *anaphoric dependence*. The presence of an anaphoric chain in a discourse signals that not all its singular terms have reference independently. Rather, some elements are related to their referents only in a derivative manner, in virtue of their anaphoric links to other expressions.

Examining the kinds of expressions that can initiate and continue such chains enables Chastain to make two important, related observations. The first concerns the significance of indefinite descriptions. Since Russell's discussions early in the century, indefinite descriptions have been treated as though they were not singular referring expressions at all, but rather to be understood by means of a quantificational paraphrase. The presence of an indefinite description often does signal existential quantification rather than singular reference as the proper semantic construal, but Chastain points out that the role of indefinite descriptions in anaphoric chains indicates that these expressions can also have a purely referential function. As in the

example above, an indefinite description can initiate an anaphoric chain, which may then be continued by pronouns or definite descriptions. And it seems clear that, in the context in which it occurs above, 'a man' purports to refer to a unique individual, namely the man in the brown suit who approached me on the street yesterday and eventually purchased my briefcase.<sup>32</sup>

This observation leads to Chastain's second point, which is that the reason that apparently nonquantificational uses of indefinite descriptions have not been thought of as straightforwardly referential is that they do not behave enough like proper names, the paradigm of singular terms. Except under deviant circumstances, if a proper name is used somewhere in a discourse invoking a particular referent, then other tokens of that same type which appear elsewhere in the discourse will be coreferential with it, in a sense that can be explained in terms of intersubstitution.<sup>33</sup> In

#Leibniz has been called a pluralist, and he has been called a monist, but no one has ever thought of that philosopher as a materialist.#

the sense is not altered if all the other elements of the anaphoric chain are replaced by the initiating expression to which they anaphorically refer and on which they anaphorically depend. An inelegant redundancy is the only cost of replacing 'he' and 'that philosopher' by 'Leibniz'. In the case of an anaphoric chain initiated by an indefinite description, on the other hand, such a substitution of terms alters the sense of the sentences in which the substituted terms appear.

Consider:

#A Republican senator threatened to filibuster the Wilderness bill. The senator's staff persuaded him that this action was unwise, so he left the chamber.#

The anaphoric chain of interest here is

A Republican senator . . . The senator . . . him . . . he.

The sense of the discourse is completely altered if the initiating expression is substituted for each of the terms that anaphorically depends on it:

#A Republican senator threatened to filibuster the Wilderness bill. A Republican senator's staff persuaded a Republican senator that this action was unwise, so a Republican senator left the chamber.#

In this passage the indefinite descriptions, although all of the same lexical type, do not purport to corefer. Each initiates a distinct anaphoric chain, and these chains may or may not involve the same individual. To continue such a chain requires either the use of a pronoun, which always continues an

existing chain, or the use of a definite description, which can either initiate or continue a chain. The fact that a chain beginning 'a SORTAL . . .' cannot be continued by repeating the initiating phrase, but can be continued with a definite description of the form 'the SORTAL . . .', is called by linguists the requirement of a "definitization transformation." One may conclude from such special requirements either that indefinite descriptions cannot function as singular referring expressions or that not all singular referring expressions must behave like ideal proper names. Chastain's suggestion is that the second alternative is worth exploring.

### 3. Anaphoric Roles of Tokens

Singular term tokens can play various roles in anaphoric chains. Such a token may initiate an anaphoric chain, as 'A Republican senator' does in the first example above. Or it may continue an existing chain and so depend for its referent on an anaphoric antecedent, as 'the senator' does in that example. Besides being distinguished as anaphoric initiators and dependents, tokens can be sorted according to two distinctions regarding the term types they instantiate. Chastain's considerations concerning substitution show that in dealing with anaphoric chains, one cannot in general assume that cotypical term tokens are coreferential, even in the absence of overtly indexical elements. Those expressions that (according to substitution tests such as that involved in comparing the last two examples) do not vary in reference from token to token within the type may be called *invariant under intratype or cotypical substitution*. An example would be proper names, as conceived and idealized by the tradition. Expressions that are referentially variable from token to token within the type may be described as not cotypically intersubstitutable. Pronouns would be a paradigm.<sup>34</sup>

The third distinction it will be useful to make is that between lexically complex expressions and those that are lexically simple, though perhaps grammatically complex—that is, between phrases that are nouns and words that are nouns. Consider, for instance, two varieties of cotypically nonintersubstitutable anaphoric dependents: dependent uses of definite descriptions and personal pronouns. The lexically simple pronoun 'he' is limited, in the information it can give about its anaphoric antecedent and the chain of which it is a part, to a small number of dimensions such as gender and number, specified in advance by the grammar of the language. Lexically complex anaphoric dependents, in contrast, can use the full descriptive resources of the language to give anaphoric information. This open-endedness permits dependents such as 'the senator mentioned above, who opposed the Wilderness bill and was dissuaded by his staff from expressing his stand'. The same contrast of lexical complexity applies to anaphoric initiators, so the indefinite description 'a Republican senator' can be compared with the proper name 'Leibniz' in the examples above.

Deploying these three independent functional distinctions—between anaphoric initiating tokens and dependent tokens, between cotypically intersubstitutable and cotypically nonintersubstitutable types, and between lexically complex and lexically simple types—yields eight roles that tokens can be thought of as playing in anaphoric chains. So among anaphoric initiators that are invariant under cotypical intersubstitution there are those that are lexically simple, such as proper names like ‘Leibniz’ as used above, and those that are lexically complex, such as ‘the first U.S. president’. Among the cotypically nonintersubstitutable anaphoric initiators there are again the lexically simple such as ‘this’, and lexically complex indefinite descriptions, such as ‘a Republican senator’ in the example above. Among the anaphoric dependents that are not cotypically intersubstitutable one can similarly distinguish lexically simple pronouns such as ‘it’ from lexically complex dependent uses of definite descriptions, such as ‘the man’ in the very first example. Finally, among the anaphoric dependents that are invariant under cotypical intersubstitution one can distinguish some uses of lexically simple proper names, as in

#I met a man I'll call ‘Binkley’. Binkley is a mechanic.#

from lexically complex dependents, which will be called *indirect definite descriptions*.

#### 4. Anaphorically Indirect Descriptions

With the exception of the last category mentioned, this tripartite division just rearranges familiar facts about the linguistic behavior of standard kinds of singular terms. The categorization was presented, however, to introduce the notion of indirect definite descriptions—a kind of singular term whose existence has not generally been recognized. Indirect definite descriptions are accordingly characterized as lexically complex anaphoric dependents that are invariant under cotypical intersubstitution. From this specification it follows that expressions in this category are complex pronouns, as are ordinary anaphorically dependent definite descriptions. Unlike such descriptions, however, all cotypical tokens of expressions in this category are guaranteed to be coreferential with each other, since they all anaphorically depend upon and hence corefer with a single common antecedent token, and so with each other.

The idea is that an indirect definite description is a pronoun that actually contains a description specifying the term occurrence that is its anaphoric antecedent. Cotypical tokens of an indirect definite description type contain the same description and so (except in special cases) specify the same antecedent. One immediate expressive advantage of a language containing locutions of this sort would be that identities employing anaphorically indirect descriptions could be used to assert that two term *tokens* (or tokenings) were

coreferential, even if the tokens were of cotypically nonintersubstitutable types (such as indefinite descriptions or pronouns), for which, as Chastain showed, standard substitutional accounts of coreference fail, since they presuppose invariance under cotypical intersubstitution.

A useful picture of the functioning of these expressions (the picture that motivates calling them *indirect* descriptions) is offered by the indirect addressing function offered in most basic computer architectures. Ordinarily, the central processor uses addresses to pick out values, just as we use descriptions to pick out objects. But, in indirect mode, the CPU when given an address as an input does not return the value stored at that address as its output. Instead it treats that value as another address and returns the value stored in that second address as its output. Indirect descriptions are to be understood by analogy to this two-step process. First, a token is specified, perhaps by being described as to type and spatiotemporal location. But the token thus picked out is not the referent of the whole indirect description. For next, an indirectness operator is applied to that token specification to produce the indirect description, which only *anaphorically* refers to the specified token and so, as a whole, refers not to that token but rather to whatever that token (its anaphoric antecedent) refers to—just as with ordinary pronouns. The flexibility of the von Neumann computer architecture is in large part due to its capacity to treat the same expression both as datum (that is, as a value) and as instruction (the address at which a value can be found). Indirect descriptions exploit the analogous use/mention amphibiousness made possible by anaphora. The claim to be defended is that the expressive dividend that traditional *semantic* vocabulary pays in a language to which it is added consists in its performing this function.

To become entitled to claim that there actually are expressions in natural languages that should be understood as playing the anaphoric role just abstractly described, and to see what indirect descriptions have to do with specifically *semantic* vocabulary, it is necessary to look at some examples. Consider a discourse in which Joe says:

#I should have known better than to let the mechanic Binkley work on my car. That airhead misadjusted the valves.#

Suppose that Jim, forgetting the name Joe used, later says:

#For car repair, don't go to the mechanic Joe referred to as 'that airhead'.#

How should this latter remark be understood? In particular, how should one understand the singular term:

(t) the mechanic Joe referred to as 'that airhead'?

Clearly this term refers to Binkley, Joe's hapless mechanic. But how is this

reference secured? The most obvious way to interpret such a singular term is as a straightforward definite description, by analogy to

the mechanic who worked on Joe's car and misadjusted the valves.

In both cases some purportedly unique feature of Binkley is used to single him out—his relation to Joe either in being referred to by him in a certain way or in having abused his car in a certain way.

But the anaphoric category of complex cotypically nonintersubstitutable dependents and the brief discussion of anaphorically indirect definite descriptions suggest that an alternative analysis might be more illuminating. For the term  $\{t\}$  can be thought of as being an anaphoric dependent, having Joe's original token of 'that airhead' as its anaphoric antecedent. If tokens of the type of  $\{t\}$  are anaphorically dependent on the original token of 'that airhead', then they are coreferential with it and hence refer to Binkley the mechanic. On this account,  $\{t\}$  should be thought of as referring to Binkley in the way that a token of 'he' would, if Jim could arrange to ensure that the antecedent of that token of 'he' were Joe's tokening of 'that airhead'. Pronouns, as simply cotypically nonintersubstitutable anaphoric dependents, can take such antecedents if the antecedent and dependent tokens are sufficiently close to each other in time, space, or audience attention. But for distant antecedents, one may not simply rely on the meager resources grammar gives us to work backward from a simple dependent token such as 'he', which even with contextual supplementation can give us only so much information about its antecedent. Here, according to the current suggestion, is where indirect definite descriptions enter. For these locutions are grammatically complex, like ordinary definite descriptions, and enable the use of the full descriptive resources of the language to specify the antecedent token to which they are anaphorically linked.

In the example, the antecedent token is specified as that token whereby Joe referred to someone as 'that airhead', the token directly picked out by the phrase "Joe's utterance of 'that airhead'." Knowing what the individual was referred to as specifies the *type* of the antecedent tokening. Invoking Joe locates the particular token of that type that is in question. The presence of 'refer' marks the indirect-addressing feature, by which it is specified that the referent of the whole description is to be understood not to be the term token picked out as anaphoric antecedent but rather, as with simple pronouns, the referent of that antecedent token. Indirect definite descriptions such as  $\{t\}$  should be understood as complex pronouns (anaphoric dependents), and 'refers' and its cognates should be understood as complex anaphoric pronoun-forming operators.<sup>35</sup>

The 'refer' cognates consist of all the sorts of expressions that would normally be thought of as being used to assert semantic word-world rela-

tions. So the following examples ought to be understood according to the model of indirect or anaphoric descriptions: 'the philosopher John mentioned yesterday', 'the restaurant he talked about at the committee meeting', 'the difficulty discussed above', 'the person denoted by the second name on the list', 'the criminal described by the police in the morning paper', and 'the cluster of buildings Russell called "as like as monkeys can make" (to those at Cambridge)'.<sup>36</sup>

### 5. Iterability

Although there are important differences among these examples, all of them could be paraphrased so as explicitly to use some form of 'refer'. But even this rough characterization is of use only insofar as it is possible to say what is special about the functioning of 'refer' that would enable one, for instance, to tell whether some alien language possessed an expression playing an analogous role. Putting the question more generally, even if it turns out that one can properly account for the behavior of expressions like those in the examples according to the indirect-addressing model of anaphoric descriptions, how could one explain and justify enforcing such a radical distinction between the analyses of descriptions as apparently analogous to (t) as

- (u) the one Joe startled (insulted, deafened) by his remark about airheads?

Does not the most intuitive reading of (t) assimilate it to (u), treating both as ordinary definite descriptions of a man who in each case happens to be picked out by his relation to some utterance of Joe's? What difference between these cases *makes* the difference in virtue of which (t) should be treated as an indirect description, which essentially involves an anaphoric link, whereas (u) should be treated as an ordinary description, which uses a relation to an utterance to pick out an object? What is the crucial difference between being referred to by a certain token and being startled, insulted, or deafened by it?

The clearest manifestation of the difference in question concerns the *iteration* of pronoun-forming operators. Because the relation '... is an anaphoric dependent of ...' is transitive, any operator that takes a term token and produces an expression that anaphorically depends on it should be iterable without change of resulting reference (assessed, as always, substitutionally). For tokens of the complex pronoun formed by applying the indirectness operator to (a token of) the result of applying that operator to an original initiating token should simply continue the anaphoric chain—dependents of dependents having the same original or ancestral antecedent.

Consider such iteration as applied to (t) and (u). If the description-forming

operators that produced these are iterated, with suitable variation of speakers, the results are:

(t') the one John referred to as "the one Joe referred to as 'that airhead'"

(u') the one John startled by his remark about the one Joe startled by his remark about airheads.

If descriptions formed in the appropriate way from 'refers' are anaphorically indirect descriptions, then, in virtue of the transparent iterability of anaphoric dependence, (t') ought to be coreferential (intersubstitutable) with (t) and, hence, with Joe's original tokening of 'that airhead'. And so they are, issues of speaker's reference aside.<sup>37</sup>

But though (t) and (t') are coreferential *de jure*, the superficially analogous (u) and (u') would be coreferential only by accident and under special circumstances.<sup>38</sup> These considerations can be formulated as the *iteration condition* [IC] below, which is a necessary condition for understanding an operator PF as a pronoun-forming operator. Subject to the conventions that a term designation surrounded by angle brackets forms a designation  $\langle \text{term} \rangle$  of the type of that term, and that such a type designation surrounded by subscripted slashes forms a designation  $/\langle \text{term} \rangle_i$  of a particular token(ing) of that type, the iteration condition can be expressed:

$$(\text{IC}) \quad \text{PF}(/ \langle \text{term} \rangle_i) = \text{PF}(/ \langle \text{PF}(/ \langle \text{term} \rangle_i) \rangle_j),$$

where the identity sign marks the intersubstitutability of these expressions.<sup>39</sup> It is clear that nothing can be thought of as a pronoun-forming operator unless it meets this condition, which embodies the transitivity of anaphora. The strategy here will be to exploit such an *iterability* requirement (suitably qualified) as a sufficient condition for identifying operators that form indirect descriptions (which have been explained as lexically complex anaphoric dependents that are invariant under cotypical intersubstitution).

Enough weight will be placed on the strategy of transforming the iterability condition from a necessary into a sufficient condition for interpreting a syntactically relational expression as a complex pronoun-forming operator to make it worth stating precisely. Consider a construction that on the surface has the form:

(v) the x [REF {x, /⟨term⟩\_i}].

The overall expression appears to be a definite description that picks out a thing x by means of its relation REF to a token /⟨term⟩\_i of type ⟨term⟩. An example would be

the man who was frightened when Bernadette uttered a token of type 'BOO'.

The iterability test says to consider terms of the form of  $\{v\}$  along with those of form

$\{v'\} \text{ the } y [\text{REF } \{y, /(\text{the } x [\text{REF } \{x, /(\text{term})/i]\})/;\}],$

as in

the man who was frightened when Bernadette uttered a token of type 'the man who was frightened when Bernadette uttered a token of type "BOO"',

as well as  $\{t'\}$  and  $\{u'\}$ .

The claim is that the syntactically relational expression REF here should be understood *not* as standing for a relation (as in an ordinary definite description) but as an anaphoric operator forming indirect descriptions, if and only if the following three conditions are met:<sup>40</sup>

1. *If  $\{v\}$  is a proper description (that is, in fact picks out one and only one object) and if  $\{v'\}$  is a proper description, then they corefer (are intersubstitutable).*
2. *If the token by relation to which the individual in  $\{v'\}$  is picked out were not of the same type as the expression  $\{v\}$ , then  $\{v\}$  and  $\{v'\}$  would not in general corefer.*<sup>41</sup>
3. Accepting the identity statement by means of which the coreference of  $\{v\}$  and  $\{v'\}$  is asserted is not accidental, in the sense that it is a condition of being taken to understand the expressions involved.

The first condition is required because expressions of type  $\{v\}$  need not always pick out unique objects, even when the expression of type  $\{v\}$  does. There need be no one who is the man Bernadette frightened by uttering a token of type 'the man who was frightened when Bernadette uttered a token of type "BOO"' (either because there is no such man or because there are too many), and similarly for genuinely anaphoric cases. Iterability is a relevant test in the (in general counterfactual) situations where the appropriate individuals exist. The second condition is required in order to rule out cases where the same individual is picked out no matter what token one looks at—the case where one and only one man is frightened, but he is frightened by whatever Bernadette says. The type of the intermediate antecedent of an anaphorically dependent expression is obviously essential to its having the reference that it has, so this condition represents a natural constraint. The third condition is required in order to rule out grammatically accidental coreference of  $\{v\}$  and  $\{v'\}$ , as might happen in a psychologically homogeneous population with the relation 'is the first object one is reminded of on hearing the expression  $\langle \text{term} \rangle$ '.<sup>42</sup> Together these three conditions ensure that any expression REF that satisfies them may appropriately be understood as forming expressions of type  $\{v\}$ , which ought to be understood as really having the form of complex anaphoric dependents, that is, as indirect definite descrip-

tions like *{t}*, rather than as ordinary definite descriptions like *{u}*. Frege placed great theoretical weight on the intersubstitutability of the terms '*t*' and 'the Bedeutung (or referent) of "*t*"', and this same essential redundancy feature of referring lies at the center of the present account.<sup>43</sup>

### *6. Other Uses of 'Refers'*

The account so far has described the anaphoric category of indirect descriptions as a form of pronoun, has offered a formal test discriminating operators that generate expression types of this category, and has pointed out that 'refers' as it appears in contexts such as

the one Joe referred to as 'that airhead'

can be understood as such a complex pronoun-forming operator. But there are other important uses of 'refers' and its cognates. The most fundamental of these are *Tarskian contexts*, such as

'Rabbits' refers to (denotes) rabbits;

*denials of reference*, such as

(The expression) 'the present king of France' does not refer (or refers to no one);

*mere reference claims*, such as

During his talk the speaker referred to Napoleon;

and *referential predication*, such as

The speaker talked about shadowy figures from the intelligence community.

The strategy is to approach these locutions in two stages. First, each such usage is paraphrased into a form in which the 'refers' cognate appears *only* inside an indirect description. Then that description is explained as functioning as a complex pronoun, according to the story already told. The present concern is thus with the paraphrase in terms of indirect descriptions.

The generalization of the token-based account of indirect descriptions required for Tarskian contexts is really a simplification to a special case. For statements about what a term refers to or denotes presuppose that the term type in question is invariant under cotypical intersubstitutions—that is, that all cotypical tokens corefer. So the Tarskian claim

The term 'Leibniz' denotes Leibniz

can be parsed as an identity involving an indirect description:<sup>44</sup>

The one denoted by the term 'Leibniz' is (=) Leibniz.

Given the presupposition of invariance under cotypical intersubstitution of the type corresponding to 'Leibniz' in the original claim, this indirect description is equivalent to

the one denoted by any token of 'Leibniz',

which may be straightforwardly understood as a complex pronoun, anaphorically dependent on an antecedent that may be any token of the specified type (for instance the one that appears on the other side of the identity sign). So the felt triviality of such reference claims is explained. Of course, reference claims involving expressions that are invariant under cotypical intersubstitution need not be epistemically trivial if different languages or different term types are involved, as in

In our world (the expression) 'the first postmaster general' refers to Benjamin Franklin (or: the inventor of bifocals),

which can be understood as using anaphoric relations to claim

In our world, the one referred to as 'the first postmaster general' is (=) Benjamin Franklin,

which is not a trivial assertion.

If Tarskian truth conditions are set up using denotation claims for semantic categories besides terms, these can be accommodated as well by this scheme.

'Red' refers to (denotes) red things

is to be read as

The ones referred to as (denoted by) 'red' are red things.

Also,

'Magnetic' applies to (has in its extension) magnetic things

is to be read as

The ones 'magnetic' applies to are magnetic things.

It was pointed out above that the prosentential account of '... is true' underwrites the Tarski biconditionals—that is, claims of the form

'Snow is white' is true if and only if snow is white.

That all the biconditionals of this form be generated is, of course, the primary criterion of adequacy Tarski imposes on candidate conceptions of truth (or better, from the present point of view, conceptions of what is expressed by '... is true'). In his own formal theory of truth, he shows how these biconditionals can be recursively generated (in a first-order formal language) on the

basis of a finite number of primitive denotation relations expressed by statements of the form

'Benjamin Franklin' denotes Benjamin Franklin

and

'Red' denotes red things.

The account of 'refers' and 'denotes' as proform-forming operators underwrites all of these basis clauses of the Tarskian recursion in just the same way that the account of '. . . is true' underwrites the biconditionals that result. Furthermore, the basis clauses of the recursion do not have to be simply stipulated, one by one, in the way that Field finds theoretically objectionable.<sup>45</sup> Rather, they are generated in a principled way from an underlying anaphoric account of what reference claims express. That account deals gracefully with extensions of the language being considered and generalizes to deal with arbitrary languages. The anaphoric approach accordingly makes intelligible what McDowell has called "modest" Tarskian theories of truth. All that is presupposed is that one understands the various subsentential parts of speech involved (which is the subject of Chapter 6) and what anaphora is (which is the subject of Chapter 7). The present discussion proceeds subject to these promissory notes.

This account of Tarskian contexts in which 'refers' and its cognates appear respects the different modal status of

The term 'Leibniz' denotes Leibniz,

which is only contingently true, and

Leibniz is Leibniz,

which is necessarily true. For the possibility that the first claim is not true can be understood in terms of its paraphrase as the existence of a possible world *w* such that

The one referred to as 'Leibniz' in *w* is not Leibniz,

that is, is not the one we refer to in our own world as 'Leibniz'. The explicit relativization of the indirect description to a possible world simply specifies which world its antecedent tokens are to be found in. The candidate antecedents of

the one referred to as 'Leibniz' in *w*

are tokenings of the type 'Leibniz' that are uttered in *w*. The anaphoric approach accordingly has room for what has been thought of as the contingency of word-world semantic relations, although it is not based on such relations.

Coreference claims represent a simple variation on Tarskian contexts and

can be interpreted in much the same way. To say that the expressions  $\langle \text{type}_1 \rangle$  and  $\langle \text{type}_2 \rangle$  corefer is just to assert an identity between the corresponding indirect definite descriptions, that is, to say that the one referred to by (tokens of)  $\langle \text{type}_1 \rangle$  is the one referred to by (tokens of)  $\langle \text{type}_2 \rangle$ . Asserting such an identity is licensing as (assertional) commitment-preserving the intersubstitution of expressions of those types. (The expressive role characteristic of identity locutions, making explicit substitution-inferential commitments, is discussed in Chapter 6.) In the case of cotypically nonintersubstitutable expressions such as demonstratives and pronouns, asserting an identity authorizes substitution of anaphoric dependents of the token on the left for anaphoric dependents of the token on the right of the identity, and vice versa. (The way in which anaphoric chains of such tokenings can play the same role in substitution inferences that classes of cotypical tokenings do for expressions that are suitably invariant is discussed in Chapter 7.)

It may be worth noticing that if attention is restricted to term types that are invariant under cotypical intersubstitution, the iteration test introduced above can be simplified correspondingly, by omission of token specifications. The necessary condition for operators PF to form complex pronouns that are invariant under cotypical intersubstitution is then:

$$\text{PF}(\langle \text{PF}(\langle \text{type} \rangle) \rangle) = \text{PF}(\langle \text{type} \rangle),$$

and the corresponding condition suggested as sufficient for REF to be an indirect description-forming operator is:

$$\text{the } y[\text{REF}\{y, \langle \text{the } x[\text{REF}\{x, \langle \text{type} \rangle\}] \rangle\}] = \text{the } x[\text{REF}\{x, \langle \text{type} \rangle\}].$$

Simple negations of statements of reference, as in

(The expression) 'the shortest man in the room' does not refer to John,

raise no new issues, for the underlying identity that is being negated has already been explained. But claims that an expression does not refer to *anything* deserve special mention. The obvious way of extending to these cases the previous strategy of paraphrasing what look like assertions of reference relations as identities involving indirect descriptions is to quantify into the identity and read the result as a negative existential statement. That is, statements of the form

$\langle \text{type} \rangle$  does not refer

are to be read as

The one referred to as  $\langle \text{type} \rangle$  does not exist,

where this last is to be understood in the same way as ordinary negative existentials, such as

The present king of France does not exist.

(Just how existential and negative existential statements ought to be understood is discussed officially in Chapter 7.) That the indirect definite descriptions involved in denials of referentiality are anaphoric dependents makes no difference to the reading of the negative existentials, any more than it causes difficulty in understanding remarks like

I would be comforted by the benevolence of a supreme being,  
except that such a being does not exist,

in which 'such a being' is an anaphoric dependent.

Statements like

During his talk the speaker referred to Napoleon

say that reference has taken place but give no information about what the referring tokens or types were. Such remarks may be understood as asserting that there is *some* term tokening *t* in the speaker's discourse such that the item referred to by *t* is Napoleon. Statements like

The speaker talked about shadowy figures from the intelligence community

are similar, except that a predication rather than an identity is what is asserted of the items referred to or talked about. This sentence says that there were tokenings *t*, *t'* (and perhaps more) such that the items referred to as (or talked about by the use of) *t* and *t'* are (have the property of being) shadowy figures from the intelligence community. These are predication involving pronouns, intrinsically no more mysterious than sentences like

They are confused.

Common nouns can be formed from indirect descriptions just as they can be from ordinary direct descriptions, and the present account extends straightforwardly to these expressions, as in

All the animals the speaker mentioned tonight were quadrupeds.

This example indicates as well how *generalizations* about reference are to be approached anaphorically. Endorsement of this claim commits one to all the substitution instances of the form

If *t* is an animal the speaker mentioned tonight, then *t* is a quadruped.

The antecedent of each such conditional is a referential predication, equivalent to

There is a term token /s/ such that the speaker uttered /s/ tonight and the item referred to by /s/ is an animal and the one referred to by /s/ is [=] *t*,

a kind of claim that has already been given an interpretation. As long as one knows what to make of each of the substitution instances to which a universally quantified claim undertakes commitment, one knows what to make of the universal generalization itself.

A full discussion of such cases requires accounts of anaphora and quantification that are not yet on the table. (The substitutional significance of such quantificational claims is discussed more fully in Chapter 6, and the account of anaphora is to be found in Chapter 7.) The complications arise in part from the recognition that in the general case the term substituend *t*, which is repeated in the specification above of the form of each sentential substitution instance of the quantification, would need to be replaced by two (not necessarily cotypical) term tokens, one of which is anaphorically dependent on the other. The present point is that although an account is not yet being offered of quantification in general, it is clear from the example that no new difficulties are added by the presence of anaphorically indirect descriptions in the quantificational substitution instances, so that generalizations about reference can be understood if any sort of generalization can.

Consideration of generalizations about what is referred to by various expressions makes salient another issue, which can be dealt with only in passing here, namely the susceptibility of an anaphoric account of reference claims to the formulation of semantic paradoxes. In the presumably analogous case of truth, a naive substitutional understanding of quantification into truth claims commits one to interpreting paradoxical sentences such as the Liar. Of course generalization is not the only way in which such paradoxical expressions can arise, nor is the possibility of semantic paradox restricted to the category of sentences. It is possible to use 'refer' to formulate empirically paradoxical term tokens, such as

(w) the square root of 2 that is the result of multiplying -1 by the one referred to by the term token marked 'w',

where 'square root of 2' is a sortal comprising the positive and negative square roots and 'one' is understood as a prosortal anaphorically dependent upon it. Interpreting such tokens as anaphorically indirect descriptions focuses attention on grounding conditions for anaphoric inheritance—a large and important topic. In "Inheritors and Paradox," Dorothy Grover elaborates an anaphoric approach to semantic paradoxes for the closely analogous anaphoric treatment of '... is true' discussed above.<sup>46</sup> Grover finds that the natural condition on anaphoric grounding yields an interpretation coinciding

in general with the sentences that Kripke assigns a semantic value to at the minimal fixed point, the interpretation he takes to provide the most natural model for the intuitive concept of truth. Her remarks can be applied to the present construction by means of the crucial analogy between the pronominal account of reference and the prosentential account of truth.

## V. THE FUNCTION OF TRADITIONAL SEMANTIC VOCABULARY IS EXPRESSIVE, NOT EXPLANATORY

### 1. Summary

The conclusions of the discussion of truth in Sections II and III may be summed up as follows. The pragmatists' approach to truth introduces a bold phenomenalist strategy—to take as immediate explanatory target the practical proprieties of *taking-true* and to understand the concept of truth as consisting in the use that is made of a class of expressions, rather than starting with a *property* of truth and then seeing what it is to express a concept used to attribute that property. Their implementation of this strategy is flawed in its exclusive attention to *taking-true* as a variety of *force* or pragmatic significance—as a doing, specifically an asserting of something. For 'true' is used in other contexts, for instance, embedded in the antecedent of a conditional; the semantic content that it expresses is accordingly not exhausted by its freestanding assertional uses. Content-redundancy theories can incorporate the insights of these force-redundancy accounts, and in their most sophisticated (anaphoric) form they account for the wider variety of uses of 'true'. Indeed, starting with an analogous pronominal account of 'refers' and 'denotes', it is possible to generate Tarski-wise the truth equivalences that jointly express the content redundancy of '... is true'. The strategy of the classical pragmatists has been vindicated at least this far: It is possible to account for truth talk without invoking a property of truth that such talk must be understood as answering to.

The project pursued in Section IV was to make it plausible that the use of 'refers' and cognate locutions in natural languages can be understood by first paraphrasing contexts in which they occur into a form in which they appear only inside indirect descriptions and then understanding their role in those paraphrases as operators taking token (or type) specifications and forming from them lexically complex pronouns invariant under cotypical intersubstitution, whose anaphoric antecedents are the specified tokens (or tokens of the specified type). This account of 'refers' as a pronoun-forming operator is evidently parallel to the account of '... is true' as a prosentence-forming operator. Each theory explains the use of a bit of traditional semantic vocabulary in terms of the formation of *anaphoric* proforms. Indirect descriptions formed from 'refer' both *mention* a term expression (in picking our anaphoric antecedents) and *use* that expression. The effect of applying an indirect

description-forming operator to a mentioned term is that of turning the mentioned occurrence into a used occurrence. Thought of in this way, 'refers' is an *anaphoric disquotation operator* in the same sense that 'true' is.<sup>47</sup>

It should be acknowledged that this anaphoric account of the use of traditional semantic vocabulary does not underwrite all of the idioms that have pressed those expressions into service. There is one sort of truth and reference talk that is not recoverable on a prosentential and pronominal rendering of 'true' and 'refers'. Talk in which the substantive 'truth' appears in a way not easily eliminable in favor of 'true' will receive no construal by such theories. Yet philosophers do say such things as "Truth is one, but beliefs are many" and "Truth is a property definable in the language of some eventual physics," which are outside the scope of the account of 'true' offered here.<sup>48</sup> Similarly, although accounts are offered of what someone referred to by an utterance and of what the reference of the utterance was, nothing is said about the *relation* of reference. The anaphoric approach does not say how to understand sentences such as "Reference is a physical, causal relation." The reason is clear enough. On the anaphoric account, although '... is true' has the surface syntactic form of a predicate, and '... refers to ...' the surface syntactic form of a relational locution, the grammatical and semantic roles these expressions play are not those of predicative and relational locutions. Their grammar is quite different; they are operators forming anaphoric dependents—namely prosentences and anaphorically indirect descriptions.

Philosophers have misconstrued ordinary talk using 'true' and 'refers' on the basis of a mistaken grammatical analogy to predicates and relational expressions. On the basis of this mistaken analogy (though the mistake is understandable, given the surface forms), they have hypostatized a property of truth and a relation of reference as the semantic correlates of the apparently predicative and relational expressions. Competing theories of the nature of this odd semantic predicate and relation have then been forthcoming. Such a search is of a piece with the search for the objects corresponding to each expression that plays the surface syntactic role of a singular term—for instance quantificational expressions such as 'someone' and 'everyone', or the 'it' in "It is raining." A more careful look at the (substitution-inferential) use of these expressions shows that the initial analogy to singular terms is misleading; a more careful look at the (anaphoric) use of 'true' and 'refers' similarly shows that the initially tempting assimilation of them to expressions of properties and relations is misleading.

One who endorses the anaphoric account of what is expressed by 'true' and 'refers' must accordingly eschew the reifying move to a truth property and a reference relation. A line is accordingly implicitly drawn by this approach between ordinary truth and reference talk and various specifically philosophical extensions of it based on theoretical conclusions that have been drawn from a mistaken understanding of what such talk expresses.

Ordinary remarks about what is true and what is false and about what some expression refers to are perfectly in order as they stand; the anaphoric account explains how they should be understood. But truth and reference are philosophers' fictions, generated by grammatical misunderstandings. It is no defect in the anaphoric account not to generate readings of the fundamentally confused remarks that result. Taking a claim to be true must be understood in the first instance as adopting a normative attitude—that is, endorsing the claim and so acknowledging a commitment. This normative attitude is presupposed by the possibility of ascribing an objective property and is not to be explained in terms of it.

Chapter 8 (Section VI) discusses in deontic scorekeeping terms what it is to ascribe objective properties, or more generally to make objectively representational claims—claims subject to objective assessments of correctness, depending on how things are with what is represented by them, regardless of the attitudes or endorsements of anyone. The expressive power of ‘true’ ensures that where an objective property is ascribed to something, the resulting claim can correctly be said to be objectively true or false. Properly understood, however, no property of truth (objective or otherwise) is being invoked by such a remark. One who asserts “The claim that all integers are the sum of at most nineteen fourth powers is objectively true” ascribes an objective property to the integers, but not to the claim. Navigating in the idiom that distinguishes these (endorsing the first specification of the ascription and not the second) is a somewhat delicate matter. The discussion below of the concept of a *fact* may be helpful in acquiring this skill.

Parallel remarks may be made about a supposed word-world relation of reference. The present account distinguishes sharply between expressions and their referents—where that latter expression is understood as shorthand for ‘what is referred to by those expressions’. For very different uses are associated with the expressions ‘the expression “Leibniz”’ and ‘the referent of the expression “Leibniz”’. In particular, although doxastic commitments are preserved by intersubstitution of the latter with the singular term ‘Leibniz’, they are not preserved by intersubstitution of the former with ‘Leibniz’. That is, Leibniz is (=) the one referred to by the expression ‘Leibniz’, but Leibniz is not the expression ‘Leibniz’. These remarks will not officially be intelligible until the substitution-inferential role of singular terms has been explained (in Chapter 6), and the relation between anaphoric dependents and their antecedents has been explained in terms of inheritance of substitution-inferential role (in Chapter 7). For present purposes it suffices to say that treating tokens of the type ⟨the referent of the expression ‘Leibniz’⟩ as anaphorically dependent on tokens of the type ⟨Leibniz⟩ commits one to the propriety of the intersubstitution that would be made explicit by endorsement of the assertible identity

Leibniz is (=) the referent of the expression ‘Leibniz’

and does not commit one to the propriety of the intersubstitution that would be made explicit by endorsement of the assertible identity

Leibniz is (=) the expression 'Leibniz'.

So although anaphora is an *intralinguistic* (or word-word) relation, adopting an anaphoric account of 'refers' as a proform-forming operator does not entail conflating linguistic items with extralinguistic items. No doubt, as with any other two items in the causal order, there are many relations that can correctly be said to obtain between a term tokening and what it refers to. But the present considerations show that talk about referring and referents provides no reason whatever to conclude that some one of these could be singled out as *the reference relation*—that unique semantically significant word-world relation in virtue of which the nonexpression is the referent of the expression. Various word-world relations play important explanatory roles in theoretical semantic projects, but to think of any one of these as what is referred to as "the reference relation" is to be bewitched by surface syntactic form.

In order to see what is and is not being claimed for this analysis of 'true' and 'refers' and to see the significance of the replacement of an account of a truth *property* and a reference *relation* by anaphoric accounts of what those traditional semantic locutions express, it is helpful to redescribe and clarify the explanatory role that the anaphoric analysis is supposed to play. Suppose that from some language-in-use such as English one extrudes all the sentences that contain specifically *semantic vocabulary*, such as 'true' and 'refers' and other words used in the same ways.<sup>49</sup> The anaphoric analysis presented in this chapter permits the extension of an account of the use of this *nonsense* fragment of the language to an account of the use of the whole language. Put otherwise, it explains how to add the expressive power provided by traditional semantic vocabulary to a set of linguistic practices that does not have such locutions. In particular, once the account (in Chapters 3 and 4) of discursive practices, the pragmatic significances they institute, and the semantic contents they confer has been extended (in Chapter 6) to incorporate the substitution-inferential commitments governing the use of subsentential expressions, and (in Chapter 7) to the anaphoric inheritance of such commitments, the considerations advanced in this chapter then suffice to incorporate 'true' and 'refers' into the deontic scorekeeping model of linguistic practice.

## 2. Semantic Deflationism

This is a *deflationary* account of the role of traditional semantic vocabulary, paradigmatically 'true' and 'refers'. Three deflationary consequences of the anaphoric approach to such vocabulary are particularly noteworthy. First, it is denied that there is a *property* of truth or a *relation* of

reference. Second, it is denied that claims expressed using traditional semantic vocabulary make it possible for us to state specifically *semantic facts*, in the way that claims expressed using the vocabulary of physics, say, make it possible for us to state specifically *physical facts*. Third, it is denied that the notion of *truth conditions* can be appealed to in explaining (as opposed to expressing) the sort of *propositional contents* expressed by declarative sentences—and similarly that the notion of *association with a referent* can be appealed to in explaining the sort of semantic contribution the occurrence of a singular term makes to the contents of sentences in which it appears.

It will help in clarifying the status and significance of these denials to consider them in connection with a significant structural difficulty that Boghossian has diagnosed as afflicting many forms of deflationism.<sup>50</sup> He points out that corresponding to each of the three sorts of claims just rehearsed there is a danger of deflationism undercutting itself and lapsing into incoherence. The general worry is that the force of deflationist claims depends on the contrast between predicates (such as ‘. . . has a mass of more than ten grams’) that do, and those (such as ‘. . . is true’) that do not, correspond to properties—and declarative sentences (such as “Snow is white”) that do, and those (such as “It is true that snow is white”) that do not, state facts. But such contrasts seem to presuppose a robust correspondence theory of the contents of some predicates and claims—at least those the semantic deflationist finds unproblematic, paradigmatically those of natural science. Deflationary approaches to semantics seem to be saying: physical predicates correspond to physical properties and relations, but semantic predicates do not correspond to semantic properties and relations; physical claims have truth conditions and if true correspond to physical facts, but semantic claims do not have truth conditions and so cannot correspond to semantic facts. Yet at the same time, the deflationists want to deny that content can be explained in terms of truth conditions and correspondence to facts, properties, and objects.

So, it is claimed, deflationary approaches are conceptually unstable. Consistently following out the rejection of robust correspondence theories of content requires treating using an expression as a predicate as all there is to expressing a property, and using a declarative sentence to make a true claim to be all there is to stating a fact. So on a deflationary construal, one is forbidden to deny that the predicate ‘. . . is true’ denotes a property, or that the claim “It is true that snow is white” states a semantic fact. Yet it is the essence of deflationism to deny these claims. So the very intuitions that deflationism seeks to develop and defend surreptitiously presuppose exactly the sort of robust truth-conditional correspondence theory of content they are concerned to reject.

These considerations present a serious challenge to many ways of pursuing the program of semantic deflationism, but the approach that has been presented here is immune to them. The first argument depends on treating

'... is true' as a predicate. If it is, then since that expression is used to make claims and state facts, it must, on deflationary accounts, be taken to express a property. But the essence of the anaphoric versions of semantic deflationism is precisely to take issue with this grammatical presupposition. According to the account endorsed here, '... is true' expresses a prosentence-forming operator. Its syntax and grammar are quite distinct from those of predicates, to which it bears only the sort of surface similarity that quantificational expressions bear to genuine singular terms. In particular, proform-forming operators such as those formed using 'true' and 'refers' are syntactically distinguished from superficially similar predicates and relational expressions by the iteration condition. The part of speech '... is true' is assimilated to by these theories does not have a directly denotational semantics but inherits its significance anaphorically, by an entirely distinct mechanism.

So when it is claimed here that '... is true' does not express a property, this means that it is not even of the right grammatical form to do so—any more than 'no one' is of the right form to pick out an individual, although there are some features of its use that could mislead one on this point. Furthermore, this claim is not made ad hoc, to avoid the sort of theoretical circularity Boghossian points out, but is motivated by ground-level considerations having to do with the use of 'true' and 'refers' that, it is claimed, cannot otherwise adequately be represented. Thus from this point of view, the argument to the effect that: "the denial that a given predicate refers to or expresses a property only makes sense on a robust construal of predicate reference . . . But if this is correct, the denial . . . that the truth predicate refers to a property must itself be understood as framed in terms of a robust notion of reference"<sup>51</sup> depends upon the ultimately incorrect presupposition that truth is properly rendered as a predicate. Given this, the second claim does not follow from the first. It can be granted that denying of a predicate that it expresses a property presupposes a robust conception, without being committed thereby to anything about how '... is true' ought to be understood.

### 3. Facts Are True Claims

What, then, about the sentential level, at which claims are made and facts stated? Is the semantic "nonfactualism" (Boghossian's term) of the deflationist incoherent? In assessing this claim, it is important to distinguish two different ways in which one might adopt a nonfactualist attitude toward what is claimed by "It is true that snow is white." According to the usage endorsed here, facts are just true claims.<sup>52</sup> That is, phenomenally, to call something a fact is just to take it to be true. 'Claims' here has the semantic sense of what is claimed, rather than the pragmatic sense of the claiming of it—a matter of content, not of force or deontic attitude. Thus to

say that facts are just true claims does not commit one to treating the facts as somehow dependent on our claimings; it does not, for instance, have the consequence that had there never been any claimers, there would have been no facts. (There are no possible situations in which there would have been no facts. A situation or set of circumstances just is one sort of set of facts.)

This notion of facts as true claims is meant to contrast with a view of them as what *makes* claims true—at least where that latter formula is conceived as potentially of *explanatory* use, rather than as providing an *expressive* equivalent. To say that it is the *fact* that *p* that *makes* it *true* that *p* (or that it is true that *p* because of the fact that *p*) is to provide an *explanation* only in the misleading sense in which what *makes* it the case that the Greeks defeated the Persians at Plataea is that the Persians were defeated by the Greeks at Plataea—the sense in which the Greeks defeated the Persians at Plataea *because* the Persians were defeated by the Greeks at Plataea.

In a certain sense, facts are what make *claimings* true. But *claimings* are true at all only in a derivative sense: We say “What you say is true,” not “Your saying of it is true.” Your saying can be speaking *truly*, but that is just saying something (making a claim) that is true. Talk of facts as what *makes* claims true is confused if it is thought of as relating two distinct things—a true claim and the fact in virtue of which it is true—in such a way that the former might be *explained* by appeal to the latter. Rather, “The claim that *p* is true” and “It is a fact that *p*” are two equivalent ways of saying the same thing—expressing the *same* content, and so (if the claim they both express is true) stating the *same* fact.

Truth claims can be true, so some of them state facts. (For the reasons rehearsed above, this does not entail that there is a property of truth.) So far this does not sound “nonfactualist.” Yet it is denied that there are any specifically *semantic* facts. “It is true that snow is white” expresses just the same fact that “Snow is white” expresses. The former uses anaphoric mechanisms to do so; its expression of that fact can thus involve presuppositions that are not involved in the latter expression (as becomes obvious if one considers other forms of truth claim), but these need not be considered part of the fact that is stated, any more than the differences involved in interpreting “Snow is white” and “La neige est blanche” or, under the right circumstances, “John is confused” and “He is confused” mean that these do not express the same fact.

Physical claims do state specifically physical facts because they essentially employ specifically physical predicates and so invoke specifically physical properties and relations. Semantic claims do not state specifically semantic facts because their use of specifically semantic vocabulary does not invoke specifically semantic properties and relations. This sounds like it ought to be called “nonfactualism” about truth talk. The distinction that must be kept in mind is that between claiming that “Snow is white is true”

states a fact (which deflationists had better not deny, for the reasons Boghossian points out) and claiming that it states a special *kind* of fact, namely a semantic fact. The 'deflating' part of deflationism can consist in its denial of this latter claim. Mastering the vocabulary of physics (or for that matter, etiquette) gives us expressive access to a range of facts we cannot otherwise express. Mastering semantic vocabulary just gives us a new way (useful for other reasons, having to do with communication and generalization) of getting at a range of nonsemantic facts we already had access to. This is just the point of redundancy deflationism. "It is true that snow is white" is a semantic expression of a nonsemantic fact.

In this connection it is worth making a third point about Boghossian's arguments against semantic deflationism. He thinks that the moral we should draw is that "we really cannot make sense of the suggestion that our thoughts and utterances do not possess robust truth conditions."<sup>53</sup> If contents must be explained as truth conditions, then an argument that such truth conditions cannot be understood in a deflationary way amounts to an argument that they, and so contents, must be construed robustly. But one could equally well conclude that one ought not to explain propositional contentfulness in terms of truth conditions. As Dummett argued long ago, anyone who holds to a deflationary theory of truth is precluded from explaining propositional contents in terms of truth conditions. For redundancy theories of 'true' presuppose the contentfulness of the nonsemantic sentences on which semantic claims are redundant, in order to explain how 'true' ought to be used. It would be circular to presuppose such contents in an account of truth, if the contents are themselves to be construed in terms of what then must be an antecedently intelligible notion of what it is to be true. If one can make robust antecedent sense of truth, then one can appeal to it to explain contents without circularity (a big 'if'). If, however, one is a deflationist about truth, then it is necessary to look elsewhere for the basic concepts one appeals to in explaining contentfulness. That is the line that has been pursued in this work.

One can say of anything that has a propositional content that it has truth conditions. According to the relaxed deflationary view, this characterization is just a harmless compliment paid to things whose contents can be expressed in declarative sentences or by the corresponding 'that' clauses (picked out in turn by their special role in the practice of making assertions). But it is one thing to say that whatever is contentful will, in consequence, have truth conditions. It is quite another to think that one could use the possession of truth conditions as part of an *explanation* of propositional contentfulness. As Dummett recognized, this latter strategy is forbidden to deflationists, on grounds of circularity. Deflationists ought to acknowledge the general possibility of expressing semantic content truth-conditionally, while denying the possibility of explaining semantic content truth-conditionally.

#### 4. Correspondence, Constraint, and Representation

From the perspective provided by this way of talking, then, correspondence theories of truth are unsatisfactory because they are unenlightening, rather than because they are false. True claims do correspond to facts, and understanding claims does require grasp of what the facts must be for those claims to be true. For when the 'ing'/'ed' ambiguity is resolved, these theses take one of two forms. If 'claim' is understood as what is claimed, true claimable contents just *are* facts; the relation of 'correspondence' is just that of identity. For that reason, grasp of such contents can be identified with grasp of what the facts must be for them to be true. But the basic question is what one must be able to *do* in practice in order to count as grasping or understanding an assertible (hence propositional) content. Paraphrases in terms of corresponding facts serving as truth conditions provide no independent explanatory grip on the issue—only an alternate way to express it.

If 'claim' is understood as the act of claiming, rather than the content claimed, however, true claim(ing)s can be said to correspond to facts in a stronger sense. They *express* those facts; they are the acts of making explicit, in virtue of whose significance as acknowledgments of inferentially articulated commitments anything at all can be understood as a claimable—and hence, if true, as a fact. In exactly the same sense, false claims express their claimable contents. In either case, scorekeeping mastery of the significance of claimings depends on one's grasp of the claimed contents. This is, trivially, grasp of what (claimable contents) must be true, what the facts must be, if the claiming is to be a true-claiming—a claiming of a true claimable content.

The important thing to get clear about is what it is for an act of claiming to express a claimable content, that is, the activity of making something explicit. Once the expressive role of 'true' and 'fact' is properly understood, it becomes apparent that their use presupposes a notion of propositional content (hence of propositionally contentful acts and states); so what such traditional semantic vocabulary expresses is not in principle available to *explain* the nature of propositional contentfulness. Their parasitic expressive role precludes their playing a fundamental semantic explanatory role.<sup>54</sup> By contrast (as Chapter 3 shows) it is possible to explain the practical significance of acts of claiming, and so to approach the propositional contents they express, without appealing to notions of truth conditions or fact. The use of expressions such as 'true' and 'fact' can then (as this chapter shows) be explained in terms of these same social practices of giving and asking for reasons.

On neither of these construals of claims is there room for a robust correspondence between facts and claims. What the facts are does not depend on what claimings we actually effect. But the worry may remain that a semantic idiom that identifies facts with true claims (via the identification of taking to be a fact with taking to be true, that is, with acknowledging a doxastic

commitment) must inevitably ‘lose the world’—trading its solidity for a froth of words. A threatening idealism of linguistic practice seems to be implicit in such an identification.

But this is a misplaced concern. What must not be lost is an appreciation of the way in which our discursive practice is empirically and practically *constrained*. It is not up to us which claims are true (that is, what the facts are). It is in a sense up to us which noises and marks express which claims, and hence, in a more attenuated sense, which express true claims. But empirical and practical constraint on our arbitrary whim is a pervasive feature of our discursive practice. Words form a distinct and largely independent realm within the world—in the sense not only that the nonlinguistic facts could be largely what they are even if the specifically linguistic facts (thought of as a class of facts about words) were quite different, but also in the sense that the words—as noises, marks, and so on—could be largely what they are, even if the nonlinguistic facts were quite different. But discursive *practices* as here conceived do not stand apart from the rest of the world in this way. The nonlinguistic facts could be largely what they are, even if our discursive practices were quite different (or absent entirely), for what claims are true does not depend on anyone’s claiming of them. But our discursive practices could not be what they are if the nonlinguistic facts were different.

For those practices are not things, like words conceived as marks and noises, that are specifiable independently of the objects they deal with and the facts they make it possible to express. Discursive practices essentially involve to-ing and fro-ing with environing objects in perception and action. The conceptual proprieties implicit in those practices incorporate both empirical and practical dimensions. All our concepts are what they are in part because of their inferential links to others that have noninferential circumstances or consequences of application—concepts, that is, whose proper use is not specifiable apart from consideration of the facts and objects that responsively bring about or are brought about by their application. The normative structure of authority and responsibility exhibited by assessments and attributions of reliability in perception and action is causally conditioned.

This sort of causal contribution to the norms implicit in discursive practice means that even though it is the practices of a linguistic community that make their words express the concepts they do, the members of the community may be understood to have undertaken commitments by using those words that outrun their capacity to recognize those commitments. Earthlings and twin-earthlings may apply the same phonetic and orthographic sign design ‘water’ to samples of clear, tasteless, odorless, thirst-quenching liquids on their respective planets (and may make corresponding inferential moves with them to and from other sign-designs) and still be understood (by us, who are describing the case) to be applying different *concepts* thereby, if the noninferential circumstances of appropriate application of their concept involves the presence of XYZ rather than H<sub>2</sub>O. This can be so even if neither

earthlings nor twin-earthlings can be trained reliably to discriminate XYZ from H<sub>2</sub>O perceptually.

Practitioners are not in general omniscient about the commitments implicit in their own concepts. For the interpreter who is making sense of their practices—and who is able (not necessarily perceptually, but conceptually) to distinguish H<sub>2</sub>O and XYZ—can understand transported earthlings as mistaking for water the XYZ they look at, as *inappropriately applying* the concept they express with their word ‘water’ to that unearthly stuff.<sup>55</sup> As with assessments of reliability, truth, and knowledge generally, the ‘externalist’ element in attributions of commitments implicit in conceptual contents reflects the social difference in perspective between the scorekeeper and those whose normative statuses are at issue. One can (according to an interpreter or scorekeeper) have bound oneself by one’s practice, in part because of the things one was actually dealing with, in such a way that using a particular word is correct in one circumstance and incorrect in another—even when the individual so bound cannot tell the situations apart.

Discursive practices incorporate actual things. They are solid—as one might say, *corporeal*: they involve actual bodies, including both our own and the others (animate and inanimate) we have practical and empirical dealings with. They must not be thought of as hollow, waiting to be filled up by things; they are not thin and abstract, but as concrete as the practice of driving nails with a hammer. (They are our means of access to what is abstract—among other things—not its product.) According to such a construal of practices, it is wrong to contrast discursive practice with a world of facts and things *outside* it, modeled on the contrast between words and the things they refer to. It is wrong to think of facts and the objects they involve as constraining linguistic practice from the outside—not because they do not constrain it but because of the mistaken picture of facts and objects as outside it. What determinate practices a community has depends on what the facts are and on what objects they are actually practically involved with, to begin with, through perception and action. The way the world is, constrains proprieties of inferential, doxastic, and practical commitment in a straightforward way from *within* those practices.<sup>56</sup> So if I perceive a liquid as tasting sour, infer that it is an acid, infer further that it will therefore turn litmus paper red, and, intending to match a red pigment sample, accordingly dip litmus paper in the liquid, I may nonetheless subsequently acquire perceptually a commitment to the result being a blue, rather than a red, piece of paper, and hence an acknowledgment of my practical failure. In this way I can find myself with incompatible commitments (which need to be sorted out if I am to remain entitled to any of my commitments in the vicinity). The possibility of incompatible commitments arising from the cycle of perception, inference, action, and perception reflects the way the normative structure of perception and action incorporates elements of the causal order.

As a result, empirical and practical constraints get built into what commitments (including inferential commitments) one can sustain entitlement to.

Thus a demotion of semantic categories of correspondence relative to those of expression does not involve ‘loss of the world’ in the sense that our discursive practice is then conceived as unconstrained by how things actually are. It does involve giving up the picture of how things are as contrasting with what we can say and think. Facts are (the contents of) true claims and thoughts. As Wittgenstein says: “When we say, and *mean*, that such-and-such is the case, we—and our meaning—do not stop anywhere short of the fact; but we mean: this-is-so.”<sup>57</sup> What is lost is only the bifurcation that makes knowledge seem to require the bridging of a *gap* that opens up between sayable and thinkable contents—thought of as existing self-contained on their side of the epistemic crevasse—and the worldly facts, existing on their side.<sup>58</sup> What the picture of facts as true claims loses is only “the little rift within the lute, / that by and by will make the music mute, / and ever widening slowly silence all.”<sup>59</sup>

The world is everything that is the case, a constellation of facts. But as the author of these words hastened to point out, those facts are structured and interconnected by the objects they are facts about; they are articulated by the properties and relations the obtaining of which is what we state when we state a fact (claim when we make a claim). To make a claim is to say that things are thus and so—that is, to talk about objects, and to say how they are propertied and related. Propositional contents (and hence facts) cannot be properly understood without understanding their representational dimension—what it means for them to be *about* objects and their properties and relations. The next three chapters accordingly show how to move from an account of the *expression* of facts to an account of the *representation* of objects and properties.

# 6

---

## Substitution: What Are Singular Terms, and Why Are There Any?

---

I start out from judgments and their contents, and not from concepts . . . I only allow the formation of concepts to proceed from judgments. If, that is, you imagine the 2 in the content of judgment  $2^4 = 16$  to be replaceable by something else, by  $(-2)$  or by 3 say, which may be indicated by putting an x in place of the 2:  $x^4 = 16$ , the content of possible judgment is thus split into a constant and a variable part. The former, regarded in its own right but holding a place open for the latter, gives the concept '4<sup>th</sup> root of 16' or 'the individual 2 falls under the concept "4<sup>th</sup> root of 16"' or "belongs to the class of 4th roots of 16''. But we may also just as well say '4 is a logarithm of 16 to the base 2'. Here 4 is being treated as replaceable and so we get the concept 'logarithm of 16 to the base 2':  $2^x = 16$  . . .

And so, instead of putting a judgment together out of an individual as subject and an already previously formed concept as predicate, we do the opposite and arrive at a concept by splitting up the content of a possible judgment.

FREGE, "Boole's Logical Calculus and the *Begriffsschrift*"

---

### I. MULTIVALUED LOGIC AND MATERIAL INFERENCE

#### 1. *Three Challenges for Inferential Approaches to Semantics*

The theoretical structure being explored here is animated by commitments both to a *deontic pragmatics* and to an *inferential semantics*. The first means that the states to be investigated, the original bearers of intentional contents, are to be understood normatively—more particularly as species of *commitments* and *entitlements*. The phenomenalist account of deontic statuses such as commitment, in terms of scorekeeping with the socially complementary deontic attitudes of *attributing* and *undertaking*, is offered as a way to begin filling in such an approach to pragmatics. The second theoretical commitment means that the *contents* that determine, in context, the deontic *significance* of adopting or altering a deontic status, or of performing a contentful act, are to be understood as broadly *inferential* roles. The content must, in context, fix the circumstances in which one would be entitled to adopt or undertake a commitment with that content and must fix the appropriate consequences of undertaking such a commitment. Employing an expression with that content then involves endorsing

the inferential commitment from those circumstances of entitlement to those consequences of commitment. The description of the game of asserting, of the inferentially articulated practices that confer assertible, that is, propositional, contents on states, acts, and utterances in virtue of their roles in that game, and the account of logical vocabulary as distinguished by its expressive task of making explicit as assertible contents precisely the inferential commitments that determine those roles, are offered as a way to begin filling in such an approach to semantics. This chapter continues the inquiry into inferential notions of semantic content.

There are three topics that would seem to pose special explanatory difficulties for attempts to understand semantic content in terms of properties of inference. First, the functional involvements that could plausibly be taken to be responsible for the conferral of such contents relate conceptually contentful deontic states not only to each other but also to the nondiscursive environment. Perception and action, as entries to and exits from the discursive realm, are governed by practical proprieties every bit as important as, and irreducible to, those governing purely inferential moves within that realm. So significant have the entries and exits seemed that each has been taken, by some empiricists and by some pragmatists in turn, to be the sole source of content for intentional states—to the exclusion not only of inferential articulation but of each other. How can a broadly *inferential* approach incorporate the aspects of semantic content conferred by the *noninferential* aspects of such entries and exits?

Second, the notion of content as inferential role seems naturally adapted to account only for *propositional* content, for it is only commitments with contents of this category that can play the role of premise and conclusion in inferences. But the sentences that express propositions typically have significant parts that are not sentences, which do not express propositions, and so which cannot serve as inferential premises and conclusions. Yet these subsentential expressions certainly ought to be said to be contentful, in virtue of what Dummett calls the “contribution” they make to the propositional contents expressed by sentences in which they occur. How can a broadly inferential approach to semantic content be extended from the grammatical category of sentences, the only sort of expression directly involved in inference, to various subsentential categories such as singular terms and predicates? For in the absence of contents corresponding to these categories, it would not be possible to understand important sorts of inferences, paradigmatically those codified explicitly by the use of identity and quantificational logical locutions.

Third, when the semantic theorist seeks to express conceptual contents explicitly, and so to reason about them—for instance when a question has arisen concerning how a certain remark should be interpreted—the semantic vocabulary employed includes, not only the logical locutions that have been construed as making *inferential* relations explicit, but also *representational*

locutions that should be understood as making referential relations explicit. Such locutions make it possible to say what someone is talking about, what is being referred to, what a belief is of or about, or what would make it true. How can a broadly inferential approach to semantic content account for the representational features of content that are expressed explicitly by means of such locutions?

The first of these *prima facie* difficulties has already been addressed (see Chapter 4). Although entry and exit moves are not themselves inferential moves, neither the noninferential acknowledgments of doxastic commitments that proximally terminate perceptual entries (as distinct from mere differential responses) nor the acknowledgments of practical commitments that noninferentially initiate actions (as distinct from other performances) can be understood apart from the role they play in the game of giving and asking for reasons, most directly as premises for cognitive reasonings, and as conclusions of practical reasonings, respectively. Perceptual reports are to be distinguished from mere reliable differential responses generally by their liability to demands for justification and their utility in providing justifications for other claims. Actions are to be distinguished from behavioral performances generally by their responsibility to assessment and deliberation concerning the inferentially articulated responsibilities they incur and discharge. So not only do perceiving that a content is true and acting so as to make it true involve endorsement of the inferential propriety of the move from the circumstances in which one is entitled to produce such a performance to the consequences one becomes committed to thereby, but those circumstances (of action) or consequences (of perception) themselves are inferentially significant.

The third of the cited challenges to an inferential approach to semantic content concerns its explanatory adequacy to the phenomena that make representational approaches to semantic content attractive and unavoidable. This is the most important and difficult issue. The general strategy for responding to it that is pursued here is to attempt to explain, in terms of the inferentially articulated social scorekeeping practices that institute discursive deontic statuses, what is expressed by the central sorts of representational semantic locutions. Where this can be done, the result is an account of what the theorist is *saying* when making claims about what represents what. Chapter 5 began this discussion by explaining the use of ‘true’ and ‘refers’ or ‘denotes’ (and so one crucial sense of ‘represents’) in terms of *anaphoric* links between expression tokenings. Chapter 8 completes the official treatment of representational locutions by specifying in discursive scorekeeping terms what it is to use locutions to make propositional-attitude ascriptions *de re*. This is the trope that makes it possible to specify what we are talking or thinking *of* or *about*, what objects our beliefs are directed at. This is the essential use in virtue of which expressions are properly interpreted as expressing attributions of ofness or aboutness in the intentional or semantic sense.

Both the discussion of what it is for a belief or claim to be of or about an object, or to be true of an object, and the discussion of 'refers', however, require that the inferentialist account of conceptual content be extended to subsentential expressions, paradigmatically singular terms and predicates. So the treatment of what is expressed by the central, explicitly representational locutions requires that the second challenge to the inferentialist order of semantic explanation be addressed. This should come as no surprise. For although some semantic thinkers (Davidson and Stalnaker are recent examples) conceive representational relations as obtaining in the first place between propositionally contentful intentional states and facts or states of affairs, they are in a distinct minority. Most representationalists have not taken the pragmatic priority of the propositional to entail a corresponding priority in the semantic order of explanation of conceptual contents. The more common position holds that the notion of representation is to be understood, to begin with, in terms of the representation of *objects*, particular things, and their properties and relations. According to this way of thinking, the basic representational bonds—in terms of which, for instance, the capacity for propositional representation, the capacity to represent possible states of affairs, is to be accounted for—are taken to be those linking represented objects to object-representings and represented properties to property-representings. If something like this turned out to be correct, adequate explanations of the function of attributions of representational purport and success could not be conducted entirely at the level of propositional contents.

There is an interaction between one's choice of semantic primitives (inference or representation) and one's choice of grammatical categorial primitives (sentences, or terms and predicates). The interaction is motivational rather than strictly conceptual, though—it is not that commitment to one semantic order of explanation entails commitment to a particular categorial order of explanation, or vice versa. Leibniz, who may serve as a paradigm for pre-Kantian inferentialists generally, begins his account with concepts standing in essentially inferential relations of inclusion to one another. Propositional contents are reached only by suitably combining these independently contentful items. So semantic inferentialism can coexist with a bottom-up categorial strategy.<sup>1</sup>

Conversely, semantic representationalism is compatible with a top-down categorial strategy, which takes the fundamental sort of content to be propositional. Representing states of affairs, purporting to represent facts, need not be thought of as semantically decomposable. If the propositions represented are thought of, for instance, as sets of possible worlds, there would seem to be no necessity to continue by explaining the capacity to represent these things in terms of more primitive capacities to represent objects or properties.<sup>2</sup> Talk of objects and object-representings and properties and property-representings would then proceed in terms of role in propositions and proposition-representings (as it does for Kant).

The bottom-up categorial strategy is obliged to explain propositional se-

mantic contents at some point, however, for these are the contents expressed by sentences, the only expressions with which, as Wittgenstein says, one can make a move in the language-game. Failure to ascend to an account of such contents would disqualify a theory as a semantic theory, for it would sever that theory from any account of the use of linguistic expressions or the significance of beliefs. It is precisely the role it plays in explaining the proprieties of the use of linguistic expressions or the possession of intentional states such as belief that qualifies something associated with those expressions as a semantic content. Dually, the top-down categorial strategy is obliged to explain the subpropositional contents expressed by subsentential expressions such as singular terms and predicates. Failure to descend to an account of such contents would doom a theory to explanatory inadequacy, for it would then be able to make no sense of the connection between saying something (expressing a proposition) and talking about something (characterizing an object). While these two ought not to be identified at the outset, the latter phenomenon is too central to our understanding of what we are doing when we talk and think simply to be ignored. Unless it accounts for the possibility of representing particular objects, a semantic theory will not address the concerns that many have taken to define its topic. The relation between these categorial strategies may be compared in this regard to that between inferentialist and representationalist commitments to fundamental semantic concepts. Each reductive order of explanation must account for the notions treated as primitive by the other, or independent accounts must be offered of each sort of primitive, together with a theory that specifies how they collaborate.

## *2. Freestanding and Ingredient Contents*

The conclusion is that any account of the representational character of propositionally contentful states, acts, and utterances is obliged to offer a reading of singular reference (the representation of particular objects) and of property-representation. For the link between belief and particular objects is a sort of paradigm of representational directedness. Offering such a reading requires looking at subpropositional contents and the way in which one expression can occur as a semantically significant component in another. The only sort of contentful expressions that have been officially discussed so far are sentences (and a very special sort of sentential operator). So it will be well to begin by considering the concept of sentential embedding in general: how the content of one sentence can contribute to the content of a compound sentence in which it is embedded as a semantically significant component. Starting with the special case in which the only subsentential components considered are themselves sentences has the advantage that this grammatical category can already be specified and understood in terms of another aspect of its use, its directly inferential significance in expressing assertional commitments. Thus a sense can be given to the question, What is the relation

between the sort of content relevant to this fundamental assertional and inferential use of sentential expressions and their derivative use as (in general) unasserted components of assertible sentential expressions? With a grip on this relation it will be possible to move on to consider the contents of expressions whose *only* use is as unasserted components of sentences, paradigmatically singular terms and predicates.

The primary job of a concept of semantic content, it has been emphasized, is to account for the pragmatic significance of the states, performances, and expressions that are understood as exhibiting such contents. The more specific theoretical commitments that have been forwarded so far are intended to fill in notions of content and significance that can satisfy this basic principle. These subordinate endorsements include the practical and normative understanding of those significances in terms of deontic states, the social-phenomenalist understanding of those deontic states in terms of discursive scorekeeping by adoption of socially perspectival deontic attitudes, the idea that the sort of practice or use to begin with is linguistic, the idea that linguistic practice is distinguished by its government of assertional performances, the idea that assertional uses are essentially inferentially articulated, and the idea that inferential involvements correspond to propositional contents.

Frege builds a basic structure of semantics and pragmatics into his system from the beginning, distinguishing accounts of the significance of judging, under the heading "theories of force," from accounts of the contents judged, under the heading "theories of content." As part of his specification of the task of the theory of content, Frege recognizes that expressions can be contentful not only in the sense that a certain force can be attached to their utterance but also in the sense that their occurrence expresses something about the content, in the first sense, of sentences in which they appear.<sup>3</sup> As Dummett puts the distinction:

In speaking of sentences themselves there are two different ways in which we may regard them; and these may give rise to two distinct notions of [content]. On the one hand, we may think of sentences as complete utterances by means of which, when a specific kind of force is attached, a linguistic act may be effected: in this connection, we require that notion of [content] in terms of which the particular kind of force may be explained. On the other hand, sentences may also occur as constituent parts of other sentences, and, in this connection, may have a semantic role in helping to determine the [content] of the whole sentence: so here we shall be concerned with whatever notion of [content] is required to explain how the [content] of a complex sentence is determined from that of its components. There is no *a priori* reason why the two notions of [content] should coincide.<sup>4</sup>

It is this second notion, and its relation to the first, that is the current topic. The technical terms Dummett introduces to capture the two dimensions

of sentential content that Frege discerned are “freestanding sense” and “ingredient sense.” Each of these indicates an explanatory role that the notion of content as truth conditions has been thought to play: settling, in context, what the assertor of a freestanding (unembedded) sentence with that content thereby becomes committed to, and settling, in sentential context, the free-standing content of a compound sentence in which it is an (embedded) ingredient. Understanding these relations is particularly important from the point of view of a strategy, such as the present one, that seeks to work backward from notions of commitment and inference to notions such as truth conditions and representation. How should the notion of ingredient content be understood, and what does it have to do with talk about truth?

To begin with, it may be pointed out that in the passage above, where the bracketed word ‘content’ has been inserted, Dummett writes “truth-value.” He is discussing Frege, and in the semantics of Frege’s extensional logic, the concept of truth-value plays both sorts of role. Truth is what matters for the force of assertions of freestanding sentences. For it is what is preserved by good inferences, in particular the inferences that are good in virtue of their logical form—the ones Frege is codifying. Furthermore, possession, by a logically compound sentence, of the property preserved by logically good inferences is determined by the truth-values of its component sentences. When the same formal apparatus is maintained as much as possible—consistent with letting different notions play these two roles—the result is classical multivalued logic.

### *3. Multivalued Logic*

The standard way of presenting these semantic ideas is as part of a bottom-up compositional definition of logical connectives, and of the validity of compound sentences formed by their use. The semantics is provided by a generalization of truth tables, defined not over Frege’s two (truth) values but over many, perhaps an infinite number. Corresponding to each  $n$ -ary syntactic compounding device is a function mapping  $n$ -tuples of values assigned to component sentences onto the value assigned to the compound sentence in which they are components. These functions are most easily visualized in the form of the familiar sort of table:

!	[*1]	[2]	[3]
[*1]	1	2	3
[2]	1	3	3
[3]	2	3	1

According to this table, for instance, an interpretation that assigns  $p$  the value [2] and  $q$  the value [3] must assign  $p \vee q$  the value [3].

Since the original role played by the notion of truth-value is being bifur-

cated, it is best not to beg questions by continuing to employ it for one or the other of these notions. The values in the set  $\{[1], [2], [3]\}$  may be called *multivalues*. One or more of the multivalues is distinguished or *designated* (indicated by the '\*' attached to the multivalue  $[1]$  in the table). A compound formula is valid in virtue of its form in case it is assigned a designated value no matter what multivalues are assigned to its component sentences.

Designatedness here indicates whatever the force-relevant notion is, for instance truth or, more generally, what is preserved by good inferences (which might in another context be some sort of commitment or entitlement). According to such a scheme, an interpretation assigns each sentence two sorts of value: as designated or not, and as having a certain multivalue.<sup>5</sup> The designatedness value includes everything that matters for the pragmatic significance of the freestanding uses of the sentence (to which assertional force can be attached) as far as it is represented by this formal apparatus. Differences between sentences that are assigned the same designatedness value (in the example, designated or not designated) are significant at all only insofar as they affect the designatedness of compounds containing them. The two undesignated multivalues in the example differ in that substituting one for the other changes not only the multivalues but the designatedness of some compounds containing them.

The standard, bottom-up direction of explanation exploits this apparatus to move from an antecedent set of multivalues that can be associated with sentences, and from functions antecedently associated with compounding devices, via a notion of designatedness, to attributions of formal validity. The same apparatus, however, can be exploited in the service of the converse, top-down direction of explanation. Then the move is from antecedently understood attributions of material designatedness to assignments of multivalues to sentences and of functions to compounding devices. The essential principle is that if two sentences have the same multivalue, then substituting one for the other never changes the designatedness of any compound sentence in which they can appear as components. This is what is meant by saying that the multivalues express the contribution a sentence makes to the designatedness value of compounds containing it.

Since any sentence can be regarded as a degenerate compound containing itself, it follows from this principle that two sentences with the same multivalue must have the same designatedness value. This is what justifies the usual procedure—embodied both in the standard tabular way of setting out semantic definitions of connectives in multivalued logic and in the definitions that generalize it to semantic matrices—of treating multivalues, rather than sentences, as what take designatedness values. Turning the basic principle around, two sentences can be treated as having the same multivalue just in case substituting one for the other never changes the designatedness value of a compound sentence in which one appears as a component. In this way, sentences are assimilated into co-multivalue classes—and so taken as

sharing their ingredient contents—accordingly as their intersubstitution as components of compound sentences preserves the designatedness values of those compounds. Lindenbaum can be understood as employing an extreme form of this strategy in his mechanism for constructing, from the set of theorems of a logic meeting certain general conditions, a matrix of multivalues and compounding functions defined over them that would validate just those theorems, by identifying multivalues with equivalence classes of logically interderivable sentences, and designatedness with theoremhood. There is no guarantee that this procedure will not end up with an infinite number of small equivalence classes (as it does in the standard Lindenbaum algebra for the propositional calculus).

How finely the ingredient contents are individuated by this substitutional test depends on the expressive power of the language, specifically on what sentential embedding contexts and embedded sentences are discerned in it. Strictly speaking, substitutional assimilation according to multivalues or ingredient contents is always relativized to a class of embedded sentence occurrences, and so to a class of sentential embedding contexts. It is not implausible that in natural languages, for any two lexically distinct sentences, there is some context in which substitution of one for the other can affect the assertional designatedness of the compound sentence resulting from such substitution. For instance, '*S* now thinks (or wishes) that *p*' is quite discriminating. This fact need not rob the substitutional form of analysis of its usefulness, for relative to various restricted classes of contexts, important assimilations are brought about nonetheless. Indeed, the partial ordering on sentential contexts that is brought about by looking at proper-inclusion relations among the multivalue equivalence classes they generate can contain interesting information about the semantic relations between those compounding devices.

Two embedding contexts can generate the same multivalues (in case they sort possible embedded sentences into just the same equivalence classes), or one can cut finer than another. Suppose, though, that every sentential embedding context that is discerned yields a different way of carving up the embedded sentences into equivalence classes, in a crazy-quilt of overlapping classes exhibiting no substantial identities or inclusions. In that case there would seem to be no theoretical advantage to discerning the semantically significant occurrence of one sentence in another. Occurrence of a sentential expression as a lexical part or syntactical subunit of another sentence is neither necessary nor sufficient to make it appropriate to discern the semantically significant occurrence of one sentence in another. The theorist may discern such occurrences where there are no lexical sentences, as with an embedded expression such as 'Kant's claim about *Aufklärung* and responsibility', or may deny them where there are, as Quine would do with direct quotation of sentences uttered. Discerning subsentential structure is enlightening only insofar as the assimilation of embedded sentences shows how the

capacity to use the embedded sentences, together with the capacity to use some of the embedding sentences, could generalize to a capacity to use the compound sentences with arbitrary embedded components.

It is worth considering a somewhat different, but closely related, point that Dummett makes, in connection with an approach to linguistic theorizing he associates with Wittgenstein:

One way in which these passages from Wittgenstein may be taken is as rejecting the whole idea that there is any one key idea in the theory of meaning: the meaning of each sentence is to be explained by a direct characterization of all the different features of its use; there is no uniform means of deriving all the other features from any one of them. Such an account would have no use for any distinction between sense and force: while it could admit some rough classification of sentences, or particular utterances of sentences, according to the kinds of linguistic act effected by means of them, it could cheerfully regard the totality of such types of linguistic act as unsurveyable—as Wittgenstein does—and would not need to invoke the classification of linguistic acts in its accounts of the meanings of particular sentences . . . The difficulty with such a theory is to see how it could do justice to the way in which the meanings of sentences are determined by the meanings of the words which compose them. The great strength of a theory which admits something as the key concept for the theory of meaning—at least a theory which is as developed as that of Frege—is that it displays a plausible pattern for the determination of the meaning of a sentence by the meanings of the constituent words . . . If nothing is to be taken to be a key concept, then we are once more without any conception of what the meaning of a word, as opposed to that of a sentence, is taken to be.<sup>6</sup>

The idea is that understanding a word need consist only in understanding the contribution it makes to the sense or content of sentences containing it. From there, the speech-act theory is to explain how that content contributes to the force or significance of various sorts of performances involving it. Otherwise, understanding the word requires mastery of the contribution it makes to all of the different acts that can be performed by means of it.

What underlies the analogy

$$\begin{array}{l} \text{force : sense} \\ (\text{or pragmatic significance : semantic content}) \\ \quad \vdots \\ \text{sentential content : subsentential content} \end{array}$$

is the thought that each variety of significance a performance can have sorts sentences that may be uttered with that force or significance into content-equivalence classes accordingly as intersubstitution preserves it. Unless dif-

ferent kinds of force or significance sort sentences into content-equivalence classes in the same way, no theoretical advance is made by discerning contents in addition to significances. Just so with assimilation of subsentential components preserving the content (or significance) of sentential compounds.

It was said above that although Frege is the first to take seriously the requirement that some aspect of semantic content determines the contribution a contentful expression makes to compounds in which it occurs as a component, nonetheless in the semantics for his logic he employs one notion, truth-value, to play both freestanding and ingredient roles. He there codifies inferences that depend only on the logical form of the sentences involved, and not on their material or nonlogical content. For these purposes he finds that it is possible to treat truth- (or commitment-) preservation not only as necessary for goodness of inference but also as sufficient. Thus for the inferences codified by his classical conditional, not only is designatedness preserved by good inferences, but any inference that preserves designatedness is a good one. For this compounding device, the two-valued conditional, sameness of designatedness value (which does duty here for freestanding content) is sufficient for sameness of multivalue (which does duty here for ingredient content). In this sentential context, the force-relevant content determines the role of sentences as components as well.

Designatedness-functional contexts such as this may be said to embed *homogeneously* with respect to designatedness values, since those values are all that matter in determining the contribution made by an embedded sentence to the designatedness value of the whole (that is, they can serve as multivalues). This term is used to mark off one of the several distinct senses sometimes attached to the expression *extensional*. Whether or not a sentential context is homogeneous in this sense concerns the relation between designatedness values and multivalues. It is quite independent of any specific conceptions of what plays the role of the immediately pragmatically relevant freestanding content. That role could be played by an antecedent concept of truth values or (looking ahead) by a concept of truth conditions.

Truth is preserved by good inferences of a certain important class. That class can be thought of as corresponding to deductive inferences, provided the notion is broadened beyond the concern with formally good inferences that is traditionally tied up with the notion of deduction. (Here the principle that a good inference never leads from premises that are true to a conclusion that is not true is being thought of as only a necessary condition on the goodness of inferences.) The inferences in question are just the commitment-preserving ones, ('committive inferences', for short). The pragmatic force of freestanding utterances of the expressions that can take truth-values of the sort preserved by good inferences (that is, sentential expressions) is assertional commitment, overtly acknowledged by, and so appropriately attributed to, the utterer. If the inference in question is a good one of this sort,

then to be committed to the premises is to be committed to the conclusion. No further understanding of the notion of truth-value in its role as designatedness value is required in order to proceed to the assimilation of sentences into multivalue equivalence classes.

Start with any set of concomitantly attributed or undertaken commitments to claims, some of which are expressed by sentences that are sententially compound (in that substitution for sentences that occur as their components makes sense). Then assign to two sentences the same componential value or multivalue (relative to that set of commitments and that compounding vocabulary) just in case substituting one for the other never turns a sentence expressing a claim in the set of concomitant commitments into a sentence expressing a claim that is not in that set. It follows that if two sentences are componentially equivalent, then they have the same designatedness value—the commitments in question include either both of them or neither.

To recap: The sort of content that has been considered here previously is the broadly inferential content that determines the correct uses of freestanding sentential utterances, paradigmatically the significance of asserting them. Content understood in this way can be associated only with expressions whose freestanding use has a pragmatic significance. It is not available as an interpretation of the contribution made by the occurrence of essentially subsentential expressions, such as singular terms and predicates. Following Frege and Dummett, a further sort of sentential content, ‘ingredient’ content, has been discerned, corresponding to the role that sentences can play as components of compound sentences. Although freestanding content may play the role of ingredient content (in homogeneous contexts), in general the latter is not reducible to the former. Ingredient contents are a sort that can coherently be attributed to expressions functioning only as components of assertible sentences, although so far only the contents to be associated with sentential sentence components have been considered. In the usual synthetic use of contents as multivalues, to define logical connectives, one begins with contents of this sort and determines designatedness values and, eventually, formal validity by their means. But the same apparatus can be exploited analytically, to move down from a notion of formal validity (as Lindenbaum does) to the assimilation of sentences according to their componential roles or, as has just been seen, from a notion of material designatedness (for example as assertional commitment) to multivalue equivalence classes. The mechanism whereby a simple notion of ingredient content (multivalue) is extracted from a simple notion of freestanding content (designatedness value) is purely substitutional. Two expressions are assimilated as making the same contribution to compound sentences in which they occur relative to some property of freestanding sentences just in case substituting one for the other never changes an embedding sentence from one that has the property to one that does not. As concern shifts to material, rather than formal, issues,

validity ceases to be the key notion, and designatedness comes to the fore, as the topmost property with respect to which substitutional invariances are assessed. The route to the notion of semantically significant occurrences of subsentential expressions, then, goes through the notion of substitution.

## II. SUBSTITUTION, SENTENTIAL EMBEDDING, AND SEMANTIC ROLES

### 1. *Substitution and Subsentential Content*

Frege's notion of *substitution* is the key to appreciating the characteristic theoretical role played by concepts of semantic content. This point begins to emerge when it is noticed that in the story just told, the relations between designatedness and multivalue, on the one hand, and between validity and designatedness, on the other, are of the same general sort. The fundamental pragmatic status that a notion of content is to help keep track of is that of assertional or doxastic commitment. As the previous chapter argued, this is a notion sufficiently intimately tied to that of truth claim that whatever sort of content ends up accounting for the pragmatic significance associated with that status for that reason has credentials as explicating one important dimension of truth talk.

Given the general understanding of the relation between material and formal proprieties of practice that has been urged earlier, the concept of the formal logical validity of claims should be treated as derived from that of material assertional commitment. The means of derivation are straightforwardly *substitutional*: A (logically) valid claim is one, first, that is designated (to which one does or ought to undertake or attribute commitment) and one, second, that cannot be turned into an undesignedated claim by any substitutions restricted to a special class of vocabulary. In the case being considered, the nonlogical vocabulary consists just in the component sentences, from which the compound sentence is conceived as resulting upon the application of a logical sentential connective. If a sentential context is not valid, in that not all substitutions preserve designatedness, then it may be substitutionally homogeneous (designatedness-functional), provided that substitution within codesignatedness classes preserves designatedness. If not, then intersubstitution within multivalue classes, substitutionally heterogeneous with respect to designatedness, by definition will preserve designatedness. Valid claims are just those special sentential contexts with respect to which the multivalue substitutional equivalence class assimilates all sentences. Ordinary claims—which are not valid with respect to substitution for components generally, nor with respect to codesignated components—are valid with respect to substitution for component sentences by sentences sharing a multivalue. The same substitutional structure is responsible for moving up from material assertional commitments to assertional validity and moving down

from material assertional commitments to multivalues or ingredient contents.

If logically valid sentential contexts are just those that assimilate all sentences into one single multivalue equivalence class, what is their special interest? They are of interest because the way those valid contexts are compounded out of other, nonvalid ones has much to teach about those nonvalid contexts, which include the basic sentential connectives. Any metatheory that identifies a logic with the set of its theorems is committed to understanding the semantics of logical expressions only insofar as it is expressed by the capacity of those expressions to enter into assertorially valid combinations. As was pointed out in Chapter 2, Dummett correctly argues that this is an unduly restrictive view of the subject matter of logic.<sup>7</sup> He would identify that subject matter by reference not to the theorems characterizing a logic but to its derivability relation. For classical Boolean logic these two notions are equivalent, but in general the theorems need not settle the derivability relation. He shows how the apparatus of multivalues can be applied in the definition of valid inferences, and not just in the definition of valid claims. He is concerned with defining validity from antecedent sets of multivalues, that is with the synthetic rather than the analytic use of this substitutional machinery. And since his topic is logical validity, he is concerned only with formal, and not with material, inferences. But the point he makes carries over to the analysis of material inferences and the derivation of a notion of material ingredient content, where its real significance becomes apparent.

A move from material assertional commitment as designatedness to material inferential commitment as designatedness corresponds to the move Dummett recommends from formal assertional validity to formal inferential validity as the notion with respect to which substitutional equivalence is assessed. The suggestion is to look at inferential commitments and correctnesses of inference instead of, or as well as, looking at assertional commitments and correctnesses of claims. A condition on the individuation of sentential contents as inferential roles can be generated from the notion of goodness of inferences by considering two sequential applications of the methodology of substitution that generates multivalues from the designatedness of compound sentences. Extending Frege's usage, two claims can be said to have the same inferential content just in case substitution of a token of the one type for a token of the other never turns a good inference into one that is not good, no matter whether the sentence appears as a premise or as part of the conclusion of the inference.

This principle does not depend on the existence in the language in question of sentential operators producing compound sentences in which other sentences are embedded. An inference here can be thought of as a pair of sets: of premise claims and of conclusion claims. Inferences can be treated as themselves a sort of compound in which sentences can appear as embedded

components, and such inferences can be classified as ‘designated’ (good) or not. If attention is restricted to inferences involving only freestanding occurrences of sentences, the equivalence classes of claims defined by preservation of goodness of inference on intersubstitution within the class may be called ‘freestanding inferential contents’. Inferential contents so defined are generated just the way multivalues are, except that instead of looking at the designatedness of compounds such as conditionals as what must be preserved by substitution, one looks at the goodness of inferences. They are the products of the first application to the analysis of inferences of the substitutional methodology suggested by multivalues. The result is just what Frege defined as “*begriffliche Inhalt*” (conceptual content) at the beginning of the *Begriffsschrift*, as discussed in Chapter 2.

Component conceptual contents can then be defined by a second application of the substitutional methodology that introduces multivalues—this time to a case where what must be preserved on substitution is not the designatedness of compound sentences but their inferential contents, which were constructed substitutionally by the first application of the analogy with multivalues. The result is to put into play two concepts of broadly inferential content: contents as the inferential potentials of freestanding sentential utterance (including both their employment as premises and as conclusions of inferences), on the one hand, and contents as the contribution a sentence makes to the inferential content of compound sentences in which it appears as a component, on the other. This latter sort of content, which may be called the ‘component content’ of a sentence, arises from considering substitution within compounds, rather than within inferences. Putting the two definitions together, it follows that two sentences have the same component content if and only if substitution of one for the other as embedded components of any compound sentence never turns from good to not good an inference in which the compound sentence appears freestanding. Assimilating sentences accordingly as their intersubstitution in inferences preserves the material goodness of inferences yields freestanding content equivalence classes, and assimilating them accordingly as their intersubstitution in sentential compounds preserves freestanding content, yields component or ingredient content-equivalence classes. On the side of assertional commitments and proprieties, beginning with material designatedness of compound claims yields one level of further substitutional assimilation, namely multivalues. On the side of inferential commitments and proprieties, the substitutional machinery can be applied twice—once to yield a notion of freestanding inferential content, and once again to yield a notion of component inferential content. (It is irrelevant for this contrast that in either case the top-level material notion, whether assertional or inferential, can also be used substitutionally to define notions of formal validity.)

The first step in generating the inferential hierarchy of substitutional levels of content was made by noticing that Frege’s substitutional definition

of equivalence of conceptual contents from material goodnesses of inference is analogous in structure to the definition of equivalence of multivalues from truth-as-designatedness, that is, material goodnesses of claims. The main adjustment required for this analogy is that inferences must be treated as a kind of context in which sentences can appear embedded, as premises and conclusions, and which as a whole is assessable according to its correctness. The second application of the substitutional machinery is more closely analogous to the assertional designatedness-multivalue paradigm in that only substitution within compound sentences in which other sentences occur as components is envisaged. It is less closely analogous to the paradigm than the first step in that what is preserved as the test of assimilation (corresponding to multivalue equivalence in the assertional paradigm) is in the inferential case not an on/off property, designated or not designated, but possession of a certain freestanding inferential role or value—of which there are many, perhaps infinitely many. It is by way of preparation for this point that the initial account above of the relation between designatedness and multivalues speaks of intersubstitution within multivalue classes as preserving designatedness *value*, even where this just means preserving designatedness.<sup>8</sup> The substitutional conceptual machinery as such is indifferent as to whether what is preserved is membership in a single class (the designated ones), alike for all compound sentences whose components are being varied, or membership in whichever element of some partition the compound sentence whose components are being varied belongs to. That there is no technical difference does not mean, however, that there is no difference in the explanatory value of applying the technical machinery.

Dummett robustly acknowledges the requirement that a notion of semantic content qualifies as such only by its relevance to the pragmatic significance of acts, for which asserting serves as a prototype. He is concerned to argue that this requirement means that the substitutionally topmost level of interpretation, the level of designatedness, must be two-valued or on/off, since what must ultimately be settled is whether an assertion is or is not correct (assertible).<sup>9</sup> As a general point, this seems dubious—perhaps a normative pragmatics need not be founded on the application of the dichotomy correct/incorrect to performances such as assertions. The discursive score-keeping account offered here is substantially more complex, as not only are commitment and entitlement distinguished, but track is kept of which are undertaken and which attributed—all articulating various ways in which a claim or an inference can count as correct. Again, thinking of pragmatic status as what must be preserved upon intersubstitution of sentences sharing a semantic content may be too narrow a formal paradigm—perhaps semantic contents can determine the correctness of material inferences without having for that reason to be conceived as preserved by good inferences.

Whether or not this on/off requirement could be shown to apply to semantic interpretation generally, there is a sense in which it is satisfied by the

inferential hierarchy that is here laid alongside the assertional one that Dummett considers. The topmost notion there is the goodness of inference, which can be thought of as a yes/no, correct/incorrect, two-valued affair: of designatedness rather than 'designatedness values'. This is because, though the purely substitutional machinery does not require it, the topmost notion in these hierarchies is a pragmatic one, as Dummett urges; furthermore, the sort of pragmatics being pursued here is one of deontic status and social attitude (commitment and entitlement, attributing and undertaking), and these are conceived as either characterizing an individual or not. The three-leveled inferential hierarchy shows, though, that the assessments of correctness that generate this two-valuedness at the top need not be directed in the first instance at sentences. When one starts with inferences, sentences are assimilated into many inferential role-equivalence classes, not simply into those that are designated as correct and those that are not.

## 2. Two Concepts of Extensionality

The assertional interpretive hierarchy—of designatedness as substantive assertional commitment and multivalues as equivalence classes of component sentences intersubstitutable saving the designatedness of compounds—gave rise to a natural notion of extensionality for sentential contexts as consisting in componential homogeneity. In this sense a context is extensional if the multivalue equivalence relation need cut no finer than the codesignatedness classes. How does the componential notion of extensionality apply to the inferential interpretive hierarchy? Since sentences are not inferences, the relation between designated inferences and the inferential roles of freestanding sentences cannot be homogeneous. It cannot strictly be that all one needs to know about freestanding sentences in order to assimilate them in such a way that intersubstitution within the resulting classes will preserve goodness of inference is whether or not they are good inferences, for they are not inferences at all. It will be necessary to look elsewhere for an analog of this sort of extensionality at the top level of the inferential hierarchy. At the lower level, when what is at issue is the relation between the inferential roles of sentences and their componential roles, however, a notion of extensionality as homogeneity does apply, since it is sentences in both cases that are assigned such roles.

A sentential context in which sentences can appear embedded as components is extensional in the sense of being componentially homogeneous just in case substitution of one claim for another with the same inferential role never alters the inferential role of the compound sentence containing them. It is a criterion of adequacy on semantically explicitating vocabulary—which has been picked out here as deserving to be called specifically *logical* vocabulary in virtue of playing that expressive role—that it generate embedding contexts that are inferential-role-functional or homogeneous in this sense.

Thus the inferential role of a conditional claim is to be settled by the inferential roles of its antecedent and consequent. Not all vocabulary is like this. In some of its uses, for instance, the inferential role played by the important, pragmatically explicitating expression '*S* claims that *p*' in some speaker's mouth depends not on the inferential role played by *p* for that speaker but on the role it is taken to play for *S*.<sup>10</sup>

Componential homogeneity is a concept that has application only within a substitutional hierarchy. There sentences are assimilated at the lower level, as associated with one sort of semantic interpretant, in case substituting one for another does not alter the assimilation at a higher level. That assimilation corresponds to association of another sort of semantic interpretant with sentences that depend in some way upon the sentences substituted for. A different sort of reducibility can be conceived in terms of the relations between the two-leveled hierarchy of assertional semantic interpretation and the three-leveled hierarchy of inferential semantic interpretation. A particularly strong bond between the assertional and the inferential orders would be forged if the designatedness of sentences determined the designatedness of inferences involving those sentences. The topmost assertional level of interpretation (committed/not committed) would then assimilate sentences into inferential-role equivalence classes, and assertional multivalues would coincide with inferential component contents. Commitment to the goodness of inferences would be preserved by substitutions for premises and conclusions, provided those substitutions preserve assertional commitment.<sup>11</sup>

Notice that the concept of multivalue is not equivalent to that of component content unless never turning a designated claim into one that is not designated is sufficient for never turning a good inference into one that is not good. Since in any case preserving designatedness is a necessary condition of a good inference, sameness of component content will guarantee sameness of multivalue, but not in general vice versa. Multivalues capture the contribution that component sentences make to only the designatedness-functional inferences involving the compounds they are embedded in. Commitment to the goodness of an inference in this sense of goodness is what is expressed by the assertion of a classical two-valued truth functional, so-called material conditional. In connection with assessments of the formal correctness of certain kinds of logical inferences, treating preservation of assertional commitment as sufficient as well as necessary is not an entirely useless strategy, as Frege shows. But the principle it embodies is simply false if applied to genuinely material inferences, whose correctnesses constitute the possession of material content by the assertible sentences that appear as their premises and conclusions.

That inferential commitments should be determined (in this substitutional sense) by assertional commitments regarding their premises and conclusions—which is an interhierarchy rather than an intrahierarchy relation—is, however, another sense that has sometimes been associated

with the notion of extensionality. Dummett, in the chapter in which he discusses the distinction between truth-value as designatedness and truth-value as ingredient or multivalue, considers the tension in Frege between two conceptions of Bedeutung. In one sense this technical term is used just to mean something like ‘semantic role’, defined by substitutional assimilations. In another sense the relation between an expression and its interpreting Bedeutung is understood to be modeled on that between a name and its bearer. At this stage in the present exposition, names and bearers are not among the concepts it is officially permissible to pretend to understand. Sentences, and their pragmatic and semantic correlates—that is, assertional commitments and inferentially articulated propositional contents—are all that are onboard so far. But Bedeutung notoriously embraces not only particular objects as referred to by singular terms but also truth-values, taken by sentences. The analog at the categorial level of sentences, to the tension at the categorial level of terms between a substitutional notion of Bedeutung as semantic role and a representational notion of it modeled on the name/named relation, is the ‘tension’ between notions of content derived from the inferential hierarchy and those derived from the assertional.

For substitution within assertional codesignatedness classes to preserve inferential designatedness is only one way in which the assertional interpretants might determine the inferential ones. Another possibility is that two sentences might have the same freestanding inferential role in case they have the same multivalue or assertional component content. Whether or not this is so depends on the expressive resources of the language, on what sort of sentence-forming locutions it makes available. Where these resources include conditionals, since these codify inferential commitments as explicit assertional commitments, the assertional designatedness of conditionals will vary with substitution of antecedent for antecedent and consequent for consequent, unless the substituends share their freestanding inferential roles. Different kinds of conditionals may codify different classes of inferences, each of which defines a correlative substitutional notion of inferential role. Where the inferences corresponding to that role are expressible by conditionals in the language, assertional multivalues must cut as fine as freestanding inferential contents.

Of course, the assertional multivalues may partition the sentences into even smaller classes, as they will if the language permits compound sentential contexts interpretable as having the form ‘*S* believes that if . . . then *q*’. Whether or not the expressive resources of the language suffice to establish a general determination of freestanding inferential contents by assertional multivalues depends not only on what conditionals exist but on how they behave. If only finitary conjunction is available, for instance, assertional multivalues adequately represent freestanding inferential contents in general only if the language is compact. Again, if infinitary conjunction is available, it must be able to form ‘enough’ conjunctions, and so conditionals. In any

case, it is clear that the proper order of explanation runs from inferential role to assertional codifying locutions (such as conditionals), to a multivalue defined substitutionally with respect to assertions formed by the use of those locutions, not the other way around.

That is, one does not start with an intrasentential notion of assertional multivalue and then use that to define conditionals, and those to define inferences. Freestanding inferential contents must be defined as part of the same conceptual package as assertional designatedness. Inferential commitment must be considered along with assertional commitment. For it is its inferential role that determines what asserting a sentence commits and entitles one to, and what could commit or entitle one to it. Apart from such inferential involvements, an assertional commitment would be without content. The assertional hierarchy of interpretation should not be conceived as independent of and antecedent to the inferential one. It may be noticed, furthermore, that even if sameness of assertional multivalue ensured sameness of freestanding inferential content, it would not follow that it ensured sameness of inferential component content, unless sameness of freestanding inferential role were sufficient for sameness of inferential component content, that is, unless the inferential hierarchy were componentially homogeneous.

If for these reasons the material-inferential interpretive hierarchy should not be seen as derivative from the material-assertional interpretive hierarchy, what about the other way around? Within these substitutional hierarchies there is a definite sense to the claim that one sort of content 'determines' another. Assertional multivalues determine assertional designatedness in the sense that two sentences cannot have the same assertional multivalue and different designatedness values. It is this sense in which freestanding inferential contents determine inferential designatedness, and inferential component contents determine the freestanding ones. One cannot in the same sense ask whether inferential designatedness determines assertional designatedness, since there is no one sort of thing that can take both values. One can ask whether it is possible for two interlocutors to undertake or have attributed to them just the same inferential commitments but different assertional ones. Apart from commitment to conditionals, it would seem that this possibility ought to be allowed. Two scientists or two politicians might agree entirely about what would be true if certain conditions obtained but nevertheless have quite different beliefs about what conditions do in fact obtain. Though it is true that in classical two-valued logic, fixing the truth-values of all the conditionals (corresponding to inferential commitments, of a sort) determines the truth-values of all of the atomic propositions, this property is an embarrassment and provides further good reason to deny that the classic horseshoe means "if . . . then . . ."<sup>12</sup> Hypothetical commitments ought not to settle categorical commitments. Rather, inferential commitments determine assertional commitments only taken together with other assertional

commitments. They cannot do the job all on their own. When purely formal inferences are at issue, there is a purely formal sort of assertion (namely of theoremhood) for which such determination can be envisioned: given a formal derivability relation  $M$ , one can consider  $\{A : \Phi M A\}$ , that is, the set of claims derivable from the empty set of premises. Of course the same set can be defined with respect to a material-inferential relation, but the resulting set of claims is not the only one compatible with the inferential commitments that generate it.

### 3. Compositionality and Decompositionality

The primary lesson that should be drawn from this discussion is that there is an intimate relationship between the notion of *semantic content* and the concept of *substitution*. That concept is one of Frege's grand themes, exploited everywhere in the official definitions of logical and semantic concepts. He is methodologically quite self-conscious about the importance of his substitutional approach—it is the basis for his technical concept of function, which in his most metaphysical writings he takes as an explicit topic for philosophical inquiry. Frege's earliest semantic and logical work introduces the concept of conceptual content in terms of substitutional behavior with respect to a kind of pragmatic significance: two claims have the same content if substituting one for the other never turns a good inference into a bad one. Goodness of inference is a pragmatic matter—in Fregean terms, a matter of force; in this paradigmatic case, it is a matter of the force of reasons. In the terms being recommended here, it is a matter of normative force, of deontic status, and so of social practice and attitude. However the pragmatic end is conceived, the route from pragmatics to semantics is that of assimilating expressions according to invariance (of pragmatic significance of some sort) under substitution. This same substitutional path that leads from inference to sentential conceptual content leads as well from the possession of freestanding inferential content by compound sentences to the possession of component-inferential content by embedded ingredient sentences and, as will appear in the rest of this chapter, from sentential content to the content of subsentential expressions such as singular terms and predicates. The substitutional way of working out a top-down categorial explanatory strategy is already implicit in the substitutional form taken by Frege's inferentialist approach to propositional content.

This *decompositional* methodology is what lies behind what is often called Frege's 'principle of compositionality'. According to that principle, the semantic interpretant associated with a compound sentence such as a conditional should be a function of the semantic interpretants associated with its semantically significant components. In spite of the way it is usually interpreted or exploited, this principle by itself is neutral between bottom-up and top-down categorial explanatory strategies. As Frege's own substitutional

understanding of functions indicates, the principle operates as a constraint on what it is for one expression to count as a semantically significant component of another, regardless of whether the compound is conceived as built up in the first place by operations on antecedently specifiable components or, conversely, the components are conceived as substitutionally precipitated out of antecedently specifiable compounds.

Typically, discussions of the compositional constraint are framed, not in terms of the generic notion of a semantic interpretant, as above, but in terms of the specific notions of sense and reference that Frege introduced in 1891. Standard sketches of the explanatory roles characteristic of those two semantic conceptions center on the following leading ideas:

1. The referent of a sentence is its truth-value, what must be preserved by good inferences.
2. The sense of a sentence is the thought it expresses, what is grasped by someone who understands it.
3. The referents associated with compound expressions are functions of the referents of their components.
4. The senses associated with compound expressions are functions of the senses of their components.
5. The sense of a sentence, together with how things actually are, fixes its referent, that is, its truth-value.

Great care is required in specifying just what commitments one is attributing to Frege under headings such as these, but the point to be made here concerns only gross structure. Again, if subsentential expressions were currently at issue, further doctrines would need to be included, most notably that the referent of a proper name is the object that sentences it occurs in are about, or on which their truth depends.

At the crude level of description expressed by these five dicta, it may be noted that there is a natural mapping of the substitutionally generated assertional and inferential semantic hierarchies onto the Fregean scheme of sense and reference. Corresponding to (1), assertional designatedness plays the role of truth-value as what must be preserved by good inferences. Corresponding to (2), what one must grasp in order to understand a sentence is conceptual content, the *begriffliche Inhalt* of the *Begriffsschrift*. Two claims are defined as having the same conceptual content in case substituting one for the other never turns a good inference into a bad one. If substitutions within compound sentences such as conditionals, appearing as premises or conclusions, are included (as the subsequent practice of the *Begriffsschrift* in fact does), then the restricted compositional principle (4) will obtain as well, for the ingredient inferential contents of component sentences determine the ingredient inferential contents of the compounds they occur in quite generally. If the relevant substitutions are restricted to sentences playing freestanding roles as premise or conclusion, then what must be grasped is only freestand-

ing inferential content. In that case the restricted compositional principle (4) will obtain only for substitutionally homogeneous contexts of embedding, that is, where sharing an inferential role is sufficient for sharing a component inferential content. In any case, the corresponding restricted compositional principle on the side of reference, (3), will obtain only where assertional codesignatedness is sufficient for multivalue assimilation, that is, in assertionally homogeneous sentential contexts.

Frege is the first to investigate systematically the concept of the special sort of semantic value or content that an expression can have in virtue of its contribution to the semantic value or content of compound expressions in which it appears as a significant component. The latter notion arises in the context of a substitutional understanding of semantic contents in their relation to pragmatic significances generally. The distinction between freestanding and ingredient semantic contents is blurred, however, by the fact that both on the side of reference (which in one of its explanatory functions has been identified here with the hierarchy of assertional contents) and on the side of sense (which in one of its explanatory functions has been identified here with the hierarchy of inferential contents), Frege seeks to have *one* notion from each hierarchy play both sorts of role. (Attention is still restricted for the time being to sentential interpretants, so this means truth-values, on the side of reference, and thoughts, on the side of sense.) A less confining semantic metatheory answering to the same insights expressed in the compositional requirements (3) and (4) above is achieved if two different sorts of content—freestanding and ingredient—are distinguished, both within the (de)compositional hierarchy generated by the semantic interpretants immediately relevant to assertional force, and within that generated by the semantic interpretants immediately relevant to the inferential content that is asserted.

What about (5), the structural principle that sense determines reference? Does the inferential role (freestanding or ingredient) of a sentence determine its assertional designatedness value? In one sense, clearly not, insofar as inferential agreement among interlocutors need not entail assertional agreement. In another sense, though, it does. Since the inferential role of the sentence determines what an assertor is committed to by it, one interlocutor cannot licitly assign to two sentences the same inferential role and different assertional designatedness values. The most interesting sense of this question, however, is neither of these. It is rather whether inferential contents determine assertional commitments in a way analogous to that in which truth conditions are understood to determine truth-values in the familiar conceptions of sentential semantics that take their cue from the later Frege's willingness to specify the senses expressed by sentences in terms of their truth conditions. Characteristic of these conceptions is an understanding of truth conditions as supplying one sort of sentential content: meaning as intension, as defining a function, which given a world or set of facts, deter-

mines the truth-value of that sentence (relative to that world or those facts). Model theory and formal intensional semantics provide representationalist, bottom-up implementations of such an understanding.

The role being played by truth-values in this sort of story is that of assertional designatedness. In the rival scheme being developed here, that role is played by the notion of assertional commitment. The notion of free-standing inferential content, which is derived substitutionally from that of inferential commitment (or designatedness of inferences), is, like that of content as truth conditions, intended to specify the content of sentences in the sense of what it is that can be assertorially designated or true. The present question is whether and in what sense inferential contents can serve the function that truth conditions serve, of determining, together with the facts, truth-values-as-designatedness-values. Can the contents substitutionally extracted from the pragmatic status of inferential commitment be construed as related to those substitutionally extracted from the pragmatic status of assertional commitment as intension to extension?

There is a sense in which inferential contents, together with the facts, determine what is true (assertorially designated). Thus component contents may be seen as corresponding on the inferentially intensional side to multi-values on the inferentially extensional, that is assertorial, side. Component contents, which determine inferential contents, can thus be thought of as expressing the contribution sentences make to the truth conditions-as-intensions of compounds containing them. Further consideration of the issue must be postponed until Chapter 8, because it cannot be pursued while continuing to suppress the additional level of analysis at which the deontic status of assertional commitment (designation) is resolved into social attitudes, which are explicitly relativized as to attributor and attributee.

When loose talk of deontic status is replaced by careful talk about deontic social attitudes, the essential clues will be seen to be that facts are just true claims, and that taking-true is just asserting. For each interlocutor, the inferential contents associated with anyone's sentences, together with the facts, determine which of those sentences express truths. The facts consulted in each case are the claims the attributing interlocutor takes to be true (that is, endorses) or acknowledges assertional commitment to. If inferential commitments, and so inferential contents, were uniform across a community, the determined truths and the determining facts would in every case coincide (though unless assertional commitments were also uniform, they would vary from attributor to attributor). But this is an extraordinary and degenerate sort of case, one that bears the same relation to the fundamental situation of communication involving the practice of truth assessment that a community in which assertional commitments are universally shared bears to the fundamental situation of communication involving the practice of assertion. The full story of how the inferential content attached to an expression by an interlocutor affects the proprieties of truth attributions by others, and espe-

cially the role played by (what are treated as) the facts, of how not only sentential but subsentential sense determines reference, is presented in Chapter 8 as the centerpiece of an account of the significance of representational idioms for semantic theories.

#### 4. Summary

Dummett explicitly distinguishes the explanatory role played by freestanding and ingredient contents and recognizes Frege's accomplishment in conceiving of the latter. He also explains the relevance of multivalued logic to these two notions of semantic content by pointing out that the two notions of truth-value in play in such logics—what have been called here ‘designatedness’ and ‘multivalue’—ought to be understood just as versions of freestanding and ingredient content, respectively. These insights of Frege and Dummett have been applied and extended here by conjoining them with three further theoretical orientations. First, the substitutional apparatus that induces the distinction between levels of content is applied *analytically*, or in a top-down categorial direction, rather than *synthetically*, or in the bottom-up categorial direction of explanation that has dominated logic and semantics since Frege. Second, where standard treatments focus exclusively on the pragmatic goodness of *asserting*, to generate a top-level notion of truth-value as assertional designatedness, attention has been drawn here as well to the pragmatic goodness of *inferring*, to assign inferential roles to sentences, on the basis of which ingredient contents can then be defined substitutionally. Dummett recommends a move like this in understanding multivalued logics, under the heading of shifting from concern with logical validity to concern with logical derivability, from formally good claims to formally good inferences.

Finally, and perhaps most important, the substitutional mechanism that relates designatedness to multivalues is applied to *contents* rather than to *forms*. It provides a general semantic structure answering to material commitments and endorsements and the proprieties they induce, not merely to formal ones. On the assertional side, assertional commitments generally are considered at the top level, not just formally validated theorems. On the inferential side, material-inferential commitments are considered at the top level, not just formally valid inferences. Furthermore, this shift from the formal to the material is extended down the substitutional hierarchy—not just from pragmatic significance to freestanding sentential semantic content, but from such content to the ingredient contents that matter for the behavior of sentences as components in compound sentences. So the sentential compounding devices that can be considered are extended from purely formal vocabulary such as the conditional to any materially contentful sentential context in which other sentences can appear embedded as components (embedded as components in the sense that they can be substituted for). The

account that emerges adds another piece to the puzzle concerning the relations between logic and semantics, a piece that dovetails with the semantically expressive view of the distinctive task of logical vocabulary.

The problem with which this discussion began is generated by the fact that concepts of *semantic content* must, to deserve that appellation, play a role in determining the *pragmatic significance* of producing an utterance or adopting a state that exhibits such content. However, direct assertional and inferential significance attaches only to sentences, and furthermore only to sentences appearing freestanding, that is, as asserted or as premise or conclusion of an inference. Introducing the notion of substitution provides a model of a sort of indirectly assertional or inferential significance that the subsentential occurrence of an expression can have. This sort of content can be associated with sentences occurring as significant components in other sentences, rather than freestanding in an assertional or inferential way. Once this sort of ingredient content has been introduced into one's semantic theory, however, it becomes available to be associated also with expressions that (unlike sentences) can occur *only* as parts of assertible sentences. So the notion of *substitution* and substitutional content—which have been used here to cash out the notion of the ‘contribution’ the occurrence of a sentential component makes to the freestanding content of compound sentences it appears in—makes available a route into the assignment of broadly assertional and inferential contents to expressions of subsentential grammatical categories, such as singular terms and predicates, to which the concept of freestanding content does not apply. The rest of this chapter is devoted to considering the sort of semantic interpretant that can be substitutionally associated with such expressions.

A bonus arising out of this line of thought is that, while remaining entirely at the level of sentences, and while eschewing any appeal to notions of reference or representation (what Dummett calls “the semantic model of the name/bearer relation”), it has been shown how to make sense of the notion of extensionality of an embedding context. Dummett is wrong to say: “If the notion of reference were introduced in the first place simply as that of the semantic role of expressions of different kinds, without an appeal to the name/bearer relation as prototype, then, at the outset, we should have no inclination to distinguish intensional from extensional contexts, or to treat the former separately; on the contrary, there would be a natural presumption in favour of a uniform semantic treatment for all contexts.”<sup>13</sup> In fact, two different (though related) senses have been specified in which a context may be called ‘extensional’. One has to do with substitutional homogeneity, that is, the sufficiency of assimilations of sentences according to their freestanding role to serve as assimilations of sentences according to their ingredient role. The other has to do with the sufficiency of concepts of content extracted from the *assertional* substitutional hierarchy of freestanding and ingredient contents to do duty as concepts of content in the *inferential* substitutional

hierarchy of freestanding and ingredient contents. It is in this sense that an inference—thought of as a context in which sentences can appear as premises and conclusions—and hence also the conditionals that makes that sort of inference propositionally explicit) can be called ‘extensional’. The clues provided by these ways of conceiving extensionality will be exploited later in discussing the relations between *inferential* and *referential* semantics.

### III. SUBSENTENTIAL EXPRESSIONS

#### 1. Singular Terms

What conditions on the use of an expression are necessary and sufficient for it to be functioning as, or playing the role of, a singular term? What sort of expressive impoverishment is a language condemned to by not having anything playing that sort of role? The answers to these questions may seem straightforward, at least in the large. Singular terms are linguistic expressions that refer to, denote, or designate particular objects.<sup>14</sup> The point of having something playing this role in linguistic practice is to make it possible to talk about particular objects, which, together with their properties and relations, make up the world in which the practice is conducted.

The first of these claims may be accepted without accepting the order of explanation presupposed by the transition from the first claim to the second. To begin with, it may be questioned whether the concept *particular object* can be made intelligible without appeal to the concept *singular term*. Frege, for instance, implicitly denies this when in the *Grundlagen* he explains the ontological category of particular objects, to which he is concerned to argue numbers belong, in effect as comprising whatever can be referred to by using singular terms, to which linguistic category he argues numerals belong. Again, it may sensibly be doubted whether the concept of singular reference is itself sufficiently clear to serve as an unexplained explainer. If it is not, then the responses offered above provide not so much answers to the questions they address as recipes for turning a suitable theory of reference into such answers. Insofar as one is sanguine about the prospects for such a theory, this of course is no bad thing. But it is important to be clear about what such a theory must account for in order to be serviceable in this explanatory context.

It is not enough, for instance, to explain only *successful* reference. For put somewhat more carefully, the first answer forwarded above must be that singular terms are expressions that, in Quine’s useful phrase, “purport to refer to just one object.”<sup>15</sup> The qualification expressed in this slogan by the use of ‘purport’ has two different functions: to acknowledge the notorious possibility that a name or a definite description may *fail* in its referential bid, as ‘the most rapidly converging sequence’ (or ‘the square root of 2’, as opposed to ‘the positive square root of 2’) does, and to exclude *accidentally* singular

expressions such as 'natural satellite of the earth', which succeed at unique signification, though they do not profess it. Ruling out possibilities of failure would require either omniscience on the part of those speaking the language or unacceptable restrictions on the formation of definite descriptions from predicates.

Quine is suspicious of the full-blooded notions of representational purport implicit in intentional idioms, and the echoes in his phrase are a reminder of his desire to explain much of what they might be thought to explain by appeal to more austere linguistic analogs. For singular referential purport, in the sense he appeals to, need not be an intentional affair. As Quine is quick to point out, "Such talk of purport is only a picturesque way of alluding to the distinctive grammatical roles that singular . . . terms play in sentences." The real task is to specify this role. Explanatory ground is gained by appeal to the principle Quine states only in the presence of such an account. That story, however, would offer a direct answer to the question, What is a singular term? one that does not appeal to (but on the contrary can itself be used via Quine's principle to help explain) the dark and pregnant notion of referential or representational purport. It is such an account that the remainder of this chapter aims to provide.

A further reservation concerning the line of thought about singular terms just considered has to do with the part the concept of singular terms is envisaged as playing in an account of the use of a language as a whole. Of particular importance is the relative explanatory priority of the category of singular terms with respect to the category of sentences. Semantic theories typically do not treat expressions of all grammatical categories equally. It is not just that different sorts of semantic interpretants are assigned to sentences, say, than to singular terms. In addition, some of those assignments of interpretants are considered basic, while others are derived from them. These latter are expression kinds whose semantic interpretation proceeds by appeal to the semantic interpretation of other sorts of expressions. A familiar example is broadly Tarskian compositional theories, which appeal to primitive assignments (perhaps relative to an index, such as a context or a model) of particular objects to atomic singular terms and of sets of those objects to atomic predicates in order to generate interpretations for sentences compounded out of them (and, along the way, to compound terms and predicates).

A contrasting direction of explanation is exhibited by broadly Fregean functional-categorial grammars and their corresponding semantics. These are less restrictive, both syntactically and semantically, than the Tarskian ones, in that any categories can be chosen as basic, not just terms and predicates, and any sort of interpretants can be associated with items of those categories, not just objects and sets of objects. A general mechanism is provided whereby (an infinite number of) further grammatical categories can be derived from the basic ones, and categorially appropriate interpretants supplied for expressions of those categories.

Suppose, as is usual, that singular terms ( $T$ ) and sentences ( $S$ ) are chosen for the basic categories. Then the derived category of (single-place) predicates ( $T \rightarrow S$ ) is understood to consist of expressions that combine with a term to yield a sentence, as 'writes' combines with 'Frege' to give 'Frege writes'. Adverbs, such as 'carefully', are expressions that combine with predicates to produce further predicates. They are  $((T \rightarrow S) \rightarrow (T \rightarrow S))s$ , taking 'writes' into 'writes carefully', for instance.<sup>16</sup> Exactly corresponding to this infinite syntactic hierarchy of derived categories is a semantic one, which turns arbitrary assignments of kinds of semantic interpretant to expressions of the basic categories into assignments of kinds of interpretant to expressions of the derived ones.

The general rule is that the semantic interpretant of an item of derived category  $(X \rightarrow Y)$  is a function taking arguments of the kind semantically associated with the category  $X$  into values of the kind semantically associated with the category  $Y$ . So if singular terms were associated with objects, and sentences with sets of possible worlds, then predicates would be assigned functions from objects to sets of possible worlds, and adverbs would be assigned functions from functions of that kind to functions of that kind. The functional mechanism is completely indifferent as to the interpretation of the primitive categories—singular terms could as well be associated with recognizability conditions, and sentences with assertibility conditions. It would then be settled automatically that quantifiers, as  $((T \rightarrow S) \rightarrow S)s$ , must be semantically interpreted by functions taking functions from recognizability to assertibility conditions into assertibility conditions.<sup>17</sup>

## *2. Subsentential Expressions and Projecting the Use of Novel Sentences*

In these two schemes for deriving the interpretation of some categories from the interpretation of more basic ones, sentences appear either as a derived semantic category or as a basic category on a par with singular terms. But it has been argued (under the heading of the pragmatic, and therefore semantic, priority of the propositional) that sentences are more special than this—that expressions of other categories count as having semantic content at all only insofar as they contribute to the content of sentences. The pre-Kantian tradition took it for granted that the proper order of semantic explanation begins with a doctrine of concepts or terms, divided into singular and general, whose meaningfulness can be grasped independently of and prior to the meaningfulness of judgments. Appealing to this basic level of interpretation, a doctrine of judgments then explains the combination of concepts into judgments, and how the correctness of the resulting judgments depends on what is combined and how. Appealing to this derived interpretation of judgments, a doctrine of consequences finally explains the combination of judgments into inferences and how the correctness of inferences depends on what is combined and how. Kant rejects this. One of his

cardinal innovations is the claim that the fundamental unit of awareness or cognition, the minimum graspable, is the judgment. For him, interpretations of something as classified or classifier make sense only as remarks about its role in judgment. In the *Grundlagen* Frege follows this Kantian line in insisting that "only in the context of a proposition [Satz] does a name have any meaning."<sup>18</sup> Frege takes this position because it is only to the utterance of sentences that pragmatic force attaches, and the explanatory purpose of associating semantic content with expressions is to provide a systematic account of such force.

That is, a further presupposition of the direction of explanation embodied in the answers forwarded above is that by saying what some expression represents (or purports to represent), one has thereby said how it ought to be used. As those candidate answers acknowledge, the category of singular terms whose nature and utility is being inquired into here is a *semantic* category. Associating objects (concrete or abstract) with expressions amounts to *semantic* interpretation of the expressions only if that association figures in the right way in accounts of how it would be correct to *use* them. Semantic properties and relations of expressions are distinguished from other sorts by the role they play in explaining the circumstances under which it is correct to use those expressions to perform various speech acts, and the appropriate consequences of so using them. Syntactic theory is concerned only to formulate rules determining what expressions are well formed, that is, can appropriately be used to perform standard speech acts. So it may group 'something' and 'everyone' in with 'the longest sentence in the A edition of Kant's first *Kritik*' and 'Aristotle'. Semantic theory ought nonetheless to distinguish the first two expressions from genuine singular terms, in virtue of the very different sorts of contribution their occurrence makes to the pragmatic significance of an utterance in a particular context.

Since semantics must in this way answer to pragmatics, the category of sentences has a certain kind of explanatory priority over subsentential categories of expression, such as singular terms and predicates. For sentences are the kind of expression whose freestanding utterance (that is, whose utterance unembedded in the utterance of some larger expression containing it) has the pragmatic significance of performing a speech act. Declarative sentences are those whose utterance typically has the significance of an assertion, of making a claim. Accordingly, there is available a sort of answer to the question,

What are sentences, and why are there any?

that is not available for any subsentential expression—namely, sentences are expressions whose unembedded utterance performs a speech act such as making a claim, asking a question, or giving a command. Without expressions of this category there can be no speech acts of any kind, and hence no specifically linguistic practice.

From this point of view it is not obvious why there should be subsen-

tial expressions at all. For they cannot have the same sort of fundamental pragmatic role to play that sentences do. As a result, they not only cannot have the same *sort* of semantic content that sentences do, they cannot even have semantic content in the same *sense* that sentences do. Sentences have pragmatic priority, in that they are the category of expressions whose use constitutes linguistic practice. Accordingly, it is sentences whose proper deployment must be determined, in context, by associating semantic interpretants with them.

From this perspective, it is necessary to ask a question more general than that of the subtitle of this chapter:

What are subsentential expressions, and why are there any?

Given the pragmatic priority of sentences, why should other semantically significant categories be discerned at all? Sentences are assigned semantic contents as part of an explanation of what one is doing in asserting them, what one claims, what belief one avows thereby. But the utterance of an essentially subsentential expression, such as a singular term, is not the performance of this sort of speech act. It does not by itself make a move in the language game, does not alter the score of commitments and attitudes that it is appropriate for an audience to attribute to the speaker. Accordingly, such expressions cannot have semantic contents in the same sense in which sentences can. They can be taken to be semantically contentful only in a derivative sense, insofar as their occurrence as components in sentences contributes to the contents (in the basic, practice-relevant sense) of those sentences.

So if, with Davidson, one takes the semantic interpretation of linguistic expressions to be an aspect of the intentional interpretation of behavior—as signing truth conditions to sentences according to the beliefs they express, and assigning truth conditions to beliefs and desires so as to make possible the explanation and prediction of behavior as largely rational for one who has beliefs and desires with those contents—then one ought to follow him as well in taking the *only* constraint on an assignment of denotations to subsentential expressions to be that it makes the truth conditions come out right. That is, one ought not to take there to be some independent notion of primitive denotation for such expressions that constrains or even determines the assignment of truth conditions.<sup>19</sup> The Tarskian technical apparatus is indifferent to whether it is exploited philosophically in the compositional, bottom-up direction Tarski originally envisaged or in the decompositional, top-down direction Davidson recommends. If one starts with the interpretation of subsentential expressions, then the primacy of the category of sentences in the linguistic practice one aims ultimately to account for provides sufficient motivation for moving up, compositionally, to generate truth conditions. What needs explanation is not this move but the concept of primitive denotation that provides its starting point. By contrast, if, because of their

pragmatic priority, one begins rather with the semantic interpretation of sentences, what is the motivation for decomposing them so as to interpret subsentential expressions as well? Why recognize the semantically significant occurrence of expressions of any category other than sentences?

Frege begins one of his later essays with this response: "It is astonishing what language can do. With a few syllables it can express an incalculable number of thoughts, so that even a thought grasped by a human being for the very first time can be put into a form of words which will be understood by someone to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thought corresponding to parts of a sentence."<sup>20</sup> The ability to produce and understand an indefinite number of novel sentences is a striking and essential feature of linguistic practice. As Chomsky has since emphasized, such creativity is the rule rather than the exception. Almost every sentence uttered by an adult native speaker is being uttered for the first time—not just the first time for that speaker, but the first time in human history. This high proportion of sentential novelty appears in surveys of empirically recorded discourses and becomes evident on statistical grounds when one compares the number of sentences of, say, thirty or fewer words, with the number there has been time for English speakers to have uttered, even if we never did anything else.<sup>21</sup> "Please pass the salt" may get a lot of play, but it is exceptionally unlikely that a sentence chosen at random from this book, for instance, would ever have been inscribed or otherwise uttered elsewhere.

The point is often made that individual speakers in training are exposed to correct uses only of a relatively small finite number of sentences and must on that basis somehow acquire practical mastery, responsive and productive, of proprieties of practice governing an indefinitely larger number.<sup>22</sup> The need to explain the possibility of projecting proper uses for many sentences from those for a few is, however, not just a constraint on accounts of language learning by individuals. For what is of interest is not just how the trick (of acquiring practical linguistic competence) might be done, but equally what the trick consists in, what counts as doing it. As just remarked, the whole linguistic community, by the most diachronically inclusive standards of community membership, has only produced (as correct) or responded to (as correct) a set of sentences that is small relative to the set of sentences one who attributes to them a language is thereby obliged to take it they have somehow determined the correct uses for. The idea that there is a difference between correct and incorrect uses of sentences no one has yet used involves some sort of projection.

There are a number of ways in which the use of a smaller number of sentences might determine the use of a larger number. If the alternative populations of sentence uses one is seeking to project to are sufficiently constrained, a small sample may suffice to determine which population is being sampled—for instance if no two of the candidate populations have any

subsets larger than  $n$  members in common, then a sample of that size will suffice to distinguish them. On the syntactic side, Chomsky follows this strategy in proposing that the reason grammar is learnable is that the final grammars available to the human learner are so severely constrained that the 'evidence' provided to such a learner in the form of sample grammatical sentences will in this way pick one out. On the semantic side, if there were relatively few in some sense possible constellations of correct usage for indefinitely many sentences, then specification or practical mastery of the correct use of a relative handful of them might well determine (in theory or in practice) the use of the rest—as ethologists have taught us that the most elaborate canned behavioral routines can be triggered by the occurrence of a few ordinary events.

Still, before resorting to the heroic postulation of the sort of structure that could make projection comprehensible even in the absence of subsentential structure (where sentences are individuated the way numerals are, for instance), it would seem the better part of valor to follow Frege in taking seriously the fact that the sentences we are familiar with do, after all, have parts. A two-stage compositional strategy for the explanation of projection would take it that what is settled by proprieties of use governing the smaller, sample set of sentences, which is projected, is the correct use of the subsentential components into which they can be analyzed or decomposed. The correct use of these components is then to be understood as determining the correct use also of further combinations of them into novel sentences.<sup>23</sup> The linguistic community determines the correct use of some sentences, and thereby of the words they involve, and so determines the correct use of the rest of the sentences that can be expressed using those words. By learning to use a relatively small initial sample of sentences, the individual learns to use the words they involve and thereby can learn to use all the sentences that can be formed out of those words by recombining them.

The need to project a distinction between proper and improper use for novel sentences provides the broad outlines of an answer to the question, What are subsentential expressions for? or Why are there any subsentential expressions? But what *are* subsentential expressions, functionally? According to the two-stage explanatory scheme, there are two sorts of constraints on the correct use of subsentential expressions, corresponding to their decompositional and compositional roles respectively. Their correct use must be determined by the correct use of relatively small subset of the sentences in which they can appear as components, and their correct use must collectively determine the correct use of all the sentences in which they can appear as components. In the passage cited above, Frege points out that this semantic projection of what he calls "thoughts" depends upon the possibility of syntactically analyzing the sentential expressions of those thoughts into elements, which can then be recombined to form novel sentences, expressing novel thoughts.

The key to the solution Frege endorses is the notion of *substitution*. For the first, or decompositional stage, sentences are to be analyzed into subsentential components by being assimilated as substitutional variants of one another—that is, related by being substitutionally accessible one from another. Regarding two sentences as substitutional variants of one another is discerning in them applications of the same function, in Frege's sense. In the second, or recompositional stage, novel sentences (and their interpretations) are to be generated as applications of familiar functions to familiar substitutable expressions. Familiar sorts of substitutional variation of familiar classes of sentences result in a host of unfamiliar sentences. This substitutional clue to the nature of subsentential expressions and their interpretation is pursued in what follows.

#### IV. WHAT ARE SINGULAR TERMS?

##### 1. Syntax: *Substitution-Structural Roles*

What are singular terms? The question has been posed from the point of view of someone who understands (or is prepared to pretend to understand) already what it is to use an expression as a sentence but who admits to puzzlement concerning the distinctive contribution made by the occurrence of singular terms in such sentences. One way to get into this situation (that pursued in Chapter 3) is to begin with a pragmatics, an account of the significance of some fundamental kinds of speech act. A line can then be drawn around the *linguistic* by insisting that for the acts in question to qualify as *speech* acts, the fundamental kinds must include *asserting*. A general pragmatic theory then specifies for each speech act the circumstances in which, according to the practices of the linguistic community, one counts as entitled or obliged to perform it, and what difference that performance makes to what various interlocutors (the performers included) are thereby entitled or obliged to do. Assertional performances (and thereby specifically linguistic practices) are in turn picked out by *inferential* articulation—namely by the way in which the pragmatic circumstances and consequences of acts of asserting depend upon the inferential relations of ground and consequent among sentences. The category of sentences is then defined as comprising the expressions whose (freestanding or unembedded) utterance standardly has the significance of performing a speech act of one of the fundamental kinds. A pair of sentences<sup>24</sup> may be said to have the same pragmatic potential if across the whole variety of possible contexts their utterances would be speech acts with the same pragmatic significance (Fregean force).

Frege's notion of *substitution* can then be employed again to define subsentential categories of linguistic expression. Two subsentential expressions belong to the same syntactic or grammatical category just in case no well-

formed sentence (expression that can be used to perform one of the fundamental kinds of speech act) in which the one occurs can be turned into something that is not a sentence merely by substituting the other for it. Two subsentential expressions of the same grammatical category share a semantic content just in case substituting one for the other preserves the pragmatic potential of the sentences in which they occur. (Where sentences can occur embedded in other sentences, of course they too can be assigned semantic contents as well as pragmatic significances and potentials.) Then the intersubstitution of cocontentful subsentential expressions can be required to preserve the semantic contents of the sentences (and other expressions) they occur in, according to the structure laid out in the first two sections of this chapter. In this way, the notion of substitution allows both syntactic and semantic equivalence relations among expressions to be defined, beginning only with an account of force or pragmatic significance. The relations differ only in the substitutional invariants: expressions assimilated accordingly as well-formedness is preserved by intersubstitution share a syntactic category; those assimilated accordingly as pragmatic potential is preserved share a semantic content.

There are three sorts of roles that expression kinds can play with respect to this substitutional machinery. An expression can be substituted *for*, replacing or replaced by another expression, as a component of a compound expression. An expression can be substituted *in*, as compound expressions in which component expressions (which can be substituted for) occur. Finally, there is the substitutional frame or remainder: what is common to two substituted-in expressions that are substitutional variants of each other (corresponding to different substituted-for expressions). ' $q \rightarrow r$ ' results from ' $p \rightarrow r$ ', by substituting ' $p$ ' for ' $q$ '. (In this example the expressions that are substituted in, ' $p \rightarrow r$ ' and ' $q \rightarrow r$ ', are sentences, and so are the expressions substituted for, ' $p$ ' and ' $q$ '.)<sup>25</sup> The substitutional frame that is common to the two substitutional variants may be indicated by ' $\alpha \rightarrow r$ ', in which ' $\alpha$ ' marks a place where an appropriate substituted-for expression would appear.

Being substituted in, substituted for, or a substitutional frame are the *substitution-structural roles* that (sets of) expressions can play. The relation *being a substitutional variant of* obtains between substituted-in expressions, which must accordingly already have been discerned. Substitutional variation is indexed by pairs of expressions that are (substituting and) substituted for, which accordingly also must be antecedently distinguishable.<sup>26</sup> Substitution frames, by contrast, are not raw materials of the substitution process; they are its products. To discern the occurrence of a substitution frame—for instance ' $\alpha \rightarrow r$ ' in ' $p \rightarrow r$ '—is to conceive of ' $p \rightarrow r$ ' as paired with the set of all of its substitutional variants, such as ' $q \rightarrow r$ '. These are available only after a substitution relation has been instituted. For this reason, being substituted for and being substituted in may be said to be *basic* substitution-structural roles, while being a substitution frame is a (substitutionally) *derived* substitution-structural role.

Frege was the first to use distinctions such as these to characterize the roles of singular terms and predicates. Frege's idea is that predicates are the substitutional sentence frames formed when singular terms are substituted for in sentences.<sup>27</sup> One of the features that Strawson finds important in distinguishing singular terms from predicates follows immediately from this characterization in terms of substitution-structural role: namely that predicates do, and singular terms do not, have argument places and fixed adicities.<sup>28</sup> But it is clear that playing the substitution-structural roles of substituted for and frame with respect to substitutions in sentences is not by itself sufficient to permit the identification of expressions as singular terms and predicates, respectively. For, as in the schematic example of the previous paragraph, what is substituted for may be sentences, rather than singular terms, and the frames exhibited by substitutionally variant (sets of) sentences thereby become sentential connectives or operators, rather than predicates.<sup>29</sup> It will be seen, though, that the substitution-structural roles do provide important necessary conditions for being singular terms and predicates.

Why not think of predicates also as expressions that can be substituted for? If "Kant admired Rousseau" has "Rousseau admired Rousseau" as a substitutional variant when the category substituted for is terms, does it not also have "Kant was more punctual than Rousseau" as a substitutional variant when the category substituted for is predicates? Indeed, does not talk about predicates as a category of expression presuppose the possibility of such replacement of one predicate by another, given the substitutional definition of 'category' offered above? It does. Though either notion can be used to assimilate expressions accordingly as it preserves well-formedness of sentences, however, it is important to distinguish between *substituting* one expression (of a basic substitution-structural kind) for another and *replacing* one sentence frame with another. These differences are discussed in detail in Section V below. A few brief observations suffice here.

To begin with, it should not be forgotten that the frames on which the latter sort of replacement operates must themselves be understood as products of the former sort of substitution operation. Whatever items play the substitutionally derivative roles, for instance of sentence frames, can be counted as *expressions* only in an extended sense. They are more like *patterns* discernible in sentential expressions, or sets of such expressions, than like parts of them. A sentence frame is not a prior constituent of a sentence (though its occurrence may be marked orthographically that way) but a product of analyzing it, in particular by assimilating it to other sentences related to it as substitutional variants, when one or more of its actual constituents is substituted for. As a result, relative to such an analysis a sentence can exhibit many occurrences of expressions that can be substituted for, but only one frame resulting from such substitutions. A further difference, which is also a consequence of the substitutionally derivative status of sentence frames, is that replacing sentence frames, or more generally discerning substitutional variants in the second, wider sense, which involves replacement

of derived categories, requires matching argument places and keeping track of cross-referencing among them.<sup>30</sup> This has no analog in substitution for expressions of substitutionally basic categories. Thus although replacement of derivative expressions is sufficiently like substitution for basic expressions to define syntactic equivalence classes of expressions, they differ in ways that will later be seen to be important.

## *2. Semantics: Substitution-Inferential Significances*

Raising the issue of the inferential significance of the occurrence in a sentence of some kind of subsentential expression shifts concern from the syntactic consequences of substitutional relations to their specifically semantic significance. Syntactic substitutional categories are defined by specifying which substitutions preserve sentencehood—where being a sentence is understood as having a pragmatic significance of its own, in that its freestanding utterance standardly counts as performing a basic speech act, paradigmatically making an assertion (overtly and explicitly acknowledging a doxastic commitment). Semantic substitutional contents can be defined by specifying which substitutions preserve the basic feature or features of sentences, in terms of which the pragmatic theory explains the proprieties of their use—namely the significance of the various speech acts they can be used to perform. This might be truth, justification, assertional commitment, or entitlement to such commitment (or whatever), as discussed in Sections I and II.<sup>31</sup>

Inferences that relate substitutionally variant substituted-in sentences as premise and conclusion, whether or not their goodness consists in the preservation of some semantically relevant whatsit, are called *substitution inferences*. An example is the inference from

Benjamin Franklin invented bifocals

to

The first postmaster general of the United States invented bifocals.

The premised sentence is substituted *in*, and a singular term is substituted *for*, to yield the conclusion. Because Benjamin Franklin was the first postmaster general of the United States, the inference from the premise to its substitutional variant is truth preserving: in the appropriate context, commitment to the premise involves commitment to the conclusion.

The substitution inference above *materially involves* the particular singular terms that occur (and are substituted for) in it. The particular predicate is not materially involved. For it is possible to replace that predicate with others without affecting the correctness (in this case, status-preservingness) of the inference. Thus if “ $\alpha$  invented bifocals” is replaced by “ $\alpha$  walked,” the substitution inference from

Benjamin Franklin walked  
to

The first postmaster general of the United States walked

will be correct under the same assumptions as the original. The predicates of concern are complex predicates, not simple ones. As such, they cannot be substituted for in a strict sense—they are substitution frames, playing a derived substitution-structural role. But a suitable notion of replacement of one sentence frame by another can be defined in terms of substitution for expressions playing basic substitution-structural roles.

The idea of replacing substitutional frames permits, for instance, substitution instances quantified over in “Anyone who admires someone admires himself,” such as

Rousseau admires Montaigne, and Rousseau admires Rousseau  
to appear as *frame-variants* of

Rousseau writes about Montaigne, and Rousseau writes about  
Rousseau,

when “ $\alpha$  admires  $\beta$ , and  $\alpha$  admires  $\alpha$ ” is replaced by “ $\alpha$  writes about  $\beta$ , and  $\alpha$  writes about  $\alpha$ .” The notion of substitution inference may be broadened to include inferences whose conclusion results from the premise upon replacement of a substitutional frame or pattern it exhibits by another. That is, the conclusions of inferences to be called ‘substitution inferences’ may be either frame-variants or strict substitutional variants of the premises (corresponding to basic and derived substitutional variation). Examples would be the inference from “Benjamin Franklin walked” to “Benjamin Franklin moved,” and from “Benjamin Franklin walked” to “The inventor of bifocals walked,” respectively.

The substitution inferences (in this broad sense) in which singular terms are materially involved differ in their formal structure from the substitution inferences in which predicates are materially involved. This difference provides another way of distinguishing the characteristic role of singular terms from that of other subsentential expressions, paradigmatically predicates. The point is noted by Strawson, who observes that predicates, but not singular terms, stand in “one-way inferential involvements.” If the inference from “Benjamin Franklin walked” to “The inventor of bifocals walked” is a good one, then so is that from “The inventor of bifocals walked” to “Benjamin Franklin walked.” Substitutions for singular terms yield reversible inferences. But it does not follow that the inference from “Benjamin Franklin moved” to “Benjamin Franklin walked” is good one, just because the inference from “Benjamin Franklin walked” to “Benjamin Franklin moved” is a good one. Replacements of predicates need not yield reversible inferences. Substitution inferences materially involving singular terms are *de jure* sym-

metric, while all predicates are materially involved in some asymmetric substitution inferences (though they may be involved in some symmetric ones as well).

One way to think about this difference is that where the goodness of a substitution inference is defined by its preservation of some semantically relevant whatsis, reflexivity and transitivity of those inferences is guaranteed by the nature of the preservation relation. The stuttering inference from  $p$  to  $p$  preserves any status that  $p$  might be accorded, while if the inference from  $p$  to  $q$  preserves that status, and that from  $q$  to  $r$  preserves it, then so must that from  $p$  to  $r$ . The symmetry of the relation, however, is assured neither by its status as an inferential relation nor by its holding accordingly as some status of the premise is preserved or transmitted to the conclusion.<sup>32</sup> Predicate substitution inferences may be asymmetric, while singular-term substitution inferences are always symmetric.

So singular terms are grouped into equivalence classes by the good substitution inferences in which they are materially involved, while predicates are grouped into reflexive, transitive, *asymmetric* structures or families. That is to say that some predicates are simply inferentially weaker than others, in the sense that everything that follows from the applicability of the weaker one follows also from the applicability of the stronger one, but not vice versa. The criteria or circumstances of appropriate application of '... walks' form a proper subset of those of '... moves'. Singular terms, by contrast, are not materially involved in substitution inferences whose conclusions are inferentially weaker than their premises.<sup>33</sup> To introduce a singular term into a language one must specify not only criteria of application but also criteria of identity, specifying which expressions are intersubstitutable with it.

Each member of such an inferential interchangeability equivalence class provides, symmetrically and indifferently, both sufficient conditions for the appropriate application and appropriate necessary consequences of application, for each of the other expressions in the class.<sup>34</sup> So, when the material substitution-inferential commitments that govern the use of singular terms are made explicit as the contents of assertional commitments, they take the form of identity claims. Identity locutions permit the expression of claims that have the significance of intersubstitution licenses. Weakening inferences, the one-way inferential involvements that collectively constitute the asymmetric substitutional significance of the occurrence of predicate expressions, are made assertorially explicit by the use of quantified conditionals. Thus "Benjamin Franklin is (=) the inventor of bifocals" ( $t = t'$ ) and "Anything that walks, moves"  $[(x)(Px \rightarrow Qx)]$ .<sup>35</sup>

### 3. Simple Material Substitution-Inferential Commitments

The substitution inference from "The inventor of bifocals wrote about electricity" to "The first postmaster general of the United States wrote

about electricity" is a material inference. Since the inventor of bifocals is (=) the first postmaster general of the United States, it is a good inference (as is its converse). This last remark is worth unpacking a bit. Part of my associating the material content I do with the term 'the inventor of bifocals' consists in the commitment I undertake to the goodness of the substitution inferences that correspond to replacements of occurrences of that term by occurrences of 'the first postmaster general of the United States' (and vice versa). (This point is independent of the availability within the language of identity locutions permitting this substitution-inferential commitment to become explicit in the form of an assertional commitment.) That commitment has a general substitution-inferential significance, which is to say that the particular material inference endorsed above is correct as an instance of a general pattern. That same material-substitutional commitment regarding "the inventor of bifocals" and "the first postmaster general of the United States" governs also the propriety of the inference from "The inventor of bifocals was a printer" to "The first postmaster general of the United States was a printer," also that from "The inventor of bifocals spoke French" to "The first postmaster general of the United States spoke French," as well as a myriad of others. So one simple material substitution-inferential commitment regarding two expressions determines the correctness of a great many substitution inferences materially involving those expressions, across a great variety of substituted-in sentences and residual sentence frames, of which '... wrote about electricity' is only one example.<sup>36</sup>

Also, the substitution inferences to and from "The inventor of bifocals was a printer" are determined by all the simple material substitution-inferential commitments (SMSICs) that link the expression 'the inventor of bifocals' with another. Not all occurrences of those expressions, however, have their substitution-inferential significances determined in this way. For instance, that the inventor of bifocals is the first postmaster general of the United States does not settle the propriety of the substitution inference from

The current postmaster general of the United States believes  
that the first postmaster general of the United States was a  
printer

to

The current postmaster general of the United States believes  
that the inventor of bifocals was a printer.<sup>37</sup>

These observations motivate the discrimination of certain occurrences of an expression, in a syntactic sense of 'occurrence', as in addition semantically significant occurrences of it. A subsentential expression has a syntactic occurrence as a component of (is exhibited by) a sentence just in case it is replaceable by other expressions of its category (either in the original sense of being substituted for or in the secondhand sense appropriate to expressions

of substitutionally derived categories), saving sentencehood. (Syntactic categories are interreplaceability equivalence classes, since replacement is reversible and preservation of sentencehood symmetric.) For an occurrence of an expression in this syntactic sense to count also as having primary substitution-semantic occurrence in a sentence, the substitution inferences to and from that sentence, in which that expression is materially involved, must be governed (their proprieties determined) by the set of simple material substitution-inferential commitments that link that expression with another.<sup>38</sup>

How do SMSICs relating subsentential expressions settle the correctness of the substitution inferences in which the sentences exhibiting primary substitution-semantic occurrences of those expressions figure as premises and conclusions? According to a general pattern. A material substitution-inferential commitment regarding  $A$  and  $A'$  is a commitment to the effect that for any  $B$  such that  $AB$  is a sentence in which  $A$  has primary substitution-semantic occurrence, the inference from  $AB$  to  $A'B$  is good. Likewise, a material substitution-inferential commitment regarding  $B$  and  $B'$  is a commitment to the effect that for any  $A$  such that  $AB$  is a sentence in which  $B$  has primary substitution-semantic occurrence, the inference from  $AB$  to  $AB'$  is good. Five points may be noted concerning this structure relating substitutional commitments to substitutional inferences.

First, *all* of the substitution inferences in which a sentence such as  $AB$  figures as premise or as conclusion are determined according to this pattern by all of the SMSICs dealing with expressions having primary substitution-semantic occurrences in  $AB$  (which might, but need not, be just  $A$  and  $B$ ). Second, responsibility for those proprieties of substitution inferences to and from a sentence is apportioned between the various subsentential expressions having primary occurrences in it, with the SMSICs dealing with a particular expression responsible for the inferences in which that expression is materially involved. The content (determiner of material proprieties of inference) of each expression is represented by the set of SMSICs that relate it to other expressions. Only the collaboration of all of the SMSICs corresponding to subsentential expressions having primary occurrence in a sentence settles the correctness of the whole set of substitution inferences it appears in as premise or conclusion. Third, a consequence of this division of labor in the determination of the correctness of material inferences (assigning aspects of it to different sorts of expression) is that material-inferential roles are determined thereby for novel compounds of familiar components. So even if no one has ever encountered the sentence  $A'B'$ , the SMSICs cited above determine a commitment to the propriety of the inference from  $AB$  to  $A'B'$ . Other SMSICs already in place may in the same way license the inference from  $A'B'$  to  $A''B'$ , and so on. Accumulating the content (what determines material proprieties of inference) to be associated with subsentential expressions in the form of substitutional commitments regarding pairs of expressions, then, permits the projection of material proprieties of substitu-

tion inference involving a potentially large set of novel sentences from the proprieties involving relatively few familiar ones. Fourth, on this model it is clear how to understand additions to or alterations of content. For when I discover or decide (what would be expressed explicitly in the claim) that the inventor of bifocals is the inventor of lightning rods and thereby undertake a new simple material substitution-inferential commitment, the substitution-inferential potentials both of sentences in which these expressions have primary occurrence and of others substitutionally linked to them are altered in determinate and predictable ways. Fifth, for the same reason, it is easy to understand what is involved in introducing new subsentential vocabulary, as expressing novel contents. Such vocabulary will make exactly the same sort of contribution to the strictly inferential contents of sentences that the old vocabulary does, as soon as its use has been tied to that of the old vocabulary by suitable SMSICs.

The criteria of adequacy responded to by these five points jointly constitute the *point* of discerning semantically significant subsentential structure, once the pragmatic, and so semantic, priority of sentences is acknowledged. Against the background of this sort of understanding of the semantically significant decomposition of sentences into their components, the formal difference between the material-substitutional commitments governing singular terms and those governing predicates becomes particularly striking. The SMSICs that determine the material-inferential significance of the occurrence of singular terms are symmetric: a commitment to the correctness of the inference that results from substituting  $A'$  for  $A$  is also a commitment to the correctness of the inference that results from substituting  $A$  for  $A'$ . The set of SMSICs that determine the material-inferential significance of the occurrence of any predicate, by contrast, include asymmetric ones. From this point of view, what is special about singular terms is that the simple material substitution-inferential commitments relating pairs of terms partition the set of terms into equivalence classes. This is what it is for it to be (particular) objects that singular terms purport-to-refer-to. An equivalence class of inter-substitutable terms stands for an object. It follows from the substitutional definition of the object-specifying equivalence classes of terms that it makes no sense to talk of languages in which there is just one singular term (*pace* 'the Absolute' as Bradley and Royce tried to use that expression), nor of objects that can in principle only be referred to in one way (by one term). The SMSICs that confer material-inferential content on predicates, by contrast, do not segregate those expressions into equivalence classes, and so do not confer a content that purports to pick out an object. The asymmetric structure conferred on the material contents of predicates is quite different.

There are, then, two fundamental sorts of substitution-inferential significance that the occurrence of expressions of various subsentential categories might have: symmetric and asymmetric. The claim so far is that it is a necessary condition for identifying some subsentential expression-kind as

predicates that expressions of that kind be materially involved in some asymmetric substitution inferences, while it is a necessary condition for identifying some subsentential expression-kind as singular terms that expressions of that kind be materially involved only in symmetric substitution inferences. These paired necessary semantic conditions distinguishing singular terms from predicates in terms of substitution-inferential significance (SIS) may be laid alongside the paired necessary syntactic conditions distinguishing singular terms from predicates in terms of substitution-structural role (SSR). The suggestion then is that these individually necessary conditions, symmetric SIS and substituted-for SSR, are jointly sufficient to characterize the use of a kind of expression that distinguishes it as playing the role of singular terms. In the rest of this work, the expression 'singular term' is used to signify expressions that play this dual syntactic and semantic substitutional role. It is to whatever expressions play this role that the argument is addressed.

## V. WHY ARE THERE SINGULAR TERMS?

### *1. Four Alternative Subsentential Analyses*

Here is an answer to the question, What are singular terms? They are expressions that are substituted for, and whose occurrence is symmetrically inferentially significant. The question, Why are there any singular terms? can now be put more sharply. Why should the expressions that are substituted for be restricted to symmetric inferential significance? What function does this arrangement serve? What expressive necessity enforces this particular combination of roles?

It is clear enough why the use of a substitutional scalpel to dissect sentential contents into subsentential components requires distinguishing expressions substituted for from substitutional frames. But why should any sort of subsentential expression have a symmetric SIS? And if some sort for some reason must, why should it be what is substituted for rather than the corresponding substitutional frames? Why should not both be symmetrically significant? The argument developed in the rest of this chapter is an attempt to answer these questions, and so in the specified sense to answer the question, Why are there singular terms?

What are the alternatives? They are structured by the previous pair of distinctions, between two sorts of substitution-structural syntactic role and between two sorts of substitution-inferential semantic significance. The possibilities are:

- |  |  |
|--|--|
| (i) substituted for is symmetric;<br>(ii) substituted for is asymmetric; | substitutional frame is<br>symmetric<br><br>substitutional frame is<br>symmetric |
|--|--|

- |                                      |                                       |
|--------------------------------------|---------------------------------------|
| (iii) substituted for is asymmetric; | substitutional frame is<br>asymmetric |
| (iv) substituted for is symmetric;   | substitutional frame is<br>asymmetric |

The final arrangement (iv) is the one actualized in languages with singular terms. One way to ask why this combination of syntactic and semantic roles is favored is to ask what is wrong with the other ones. What rules out the combinations (i), (ii), and (iii)? What sort of consideration could? The strategy pursued here is to look at the constraints on the expressive power of a language that are imposed by each of those varieties of complex substitutional roles.

The first alternative is a good place to begin, for it is fairly easily eliminated from contention. The semantic point of discerning subsentential structure substitutionally is to codify an antecedent field of inferential proprieties concerning sentences by associating material contents with recombinable subsentential expressions so as to be able to derive those proprieties of inference and to project further ones, according to a general pattern of substitution-inferential significance of material-substitutional commitments. But the substituted-in sentences whose inferences are to be codified themselves stand in "one-way inferential involvements." The goodness of an inference may require that when the conclusion is substituted for the premise(s), some status (doxastic or assertional commitment, truth, and so on) is preserved. But the converse replacement need not preserve that status. Substitution inferences are not always reversible, saving correctness. Conclusions are often inferentially weaker than the premises from which they are inferred. A restriction to sentential contents conferrable by exclusively symmetrically valid material inferences is a restriction to sentential contents completely unrecognizable as such by us. But if both substituted-for expressions and the substitutional frames that are the patterns according to which they assimilate substituted-in sentences are significant only according to symmetric SMSICs, then asymmetric inferential relations involving substituted-in sentences can never be codified as substitution inferences materially involving subsentential expressions, and so licensed by the SMSICs regarding those expressions. Since the inferences to be codified include asymmetric ones, either the substituted-for expressions or the substitutional frames, or both, must be assigned asymmetric substitution-inferential significance.

The other two alternatives, (ii) and (iii), are alike in assigning the substituted-for expressions asymmetric substitution-inferential significance. If a good reason can be found for ruling out this combination of syntactic and semantic substitutional roles, then the employment of singular terms and their corresponding sentence frames will have been shown to be necessary. For if it can be shown that what is substituted for must have symmetric substitution-inferential significance, then since by the argument just offered the expressions playing some substitution-structural role must be asymmet-

ric (besides the substituted-in expressions), it will follow that the substitution frames must permit asymmetric substitution. Just this combination of roles has been put forward as characteristic of singular terms and predicates.

The first task addressed (in Section IV) was to answer the question, What are singular terms? The answer that has emerged is that they are expressions that on the syntactic side play the substitution-structural role of being substituted for, and on the semantic side have symmetric substitution-inferential significances. The second task is to answer the question, Why are there any singular terms? by presenting an explanation of why the inferential significance of the occurrence of expressions that are substituted for must be symmetric (and so segregate expressions materially into equivalence classes whose elements accordingly jointly purport to specify some one object). It takes the form of an argument that certain crucial sorts of expressive power would be lost in a language in which the significance of substituted-for expressions was permitted to be asymmetric.

## 2. The Argument

What is wrong with substituted-for expressions having asymmetric inferential significances? An asymmetric simple material substitution-inferential commitment linking substituted-for expressions *a* and *b* is a commitment to the goodness of all the inferences that are instances of a certain pattern. Where *Pa* is any sentence in which *a* has primary occurrence, the inference from *Pa* to *Pb* (the result of substituting *b* for *a* in *Pa*) is a good one, though perhaps its converse is not. The point of discerning primary occurrences of substituted-for expressions depends on these generalizations, for they provide the link that permits the projection of proprieties of substitution inference, based on associating particular substituted-fors with material contents in the form of determinate sets of simple substitution-inferential commitments relating their use to that of other substituted-fors. Whether the generalizations that animate asymmetrically significant substitutional commitments regarding substituted-fors make sense or not depends on the contents expressed by the sentences substituted in, and it is this fact that in the end turns out to mandate symmetric substitutional significances for what is substituted for. To see this, consider three ever more radical ways in which the generalizations associated with simple material substitution-inferential commitments might fail to obtain.

First, suppose there were sentences *Qa*, *Qb*, *Q'a*, *Q'b*, such that the inference from *Qa* to *Qb* is a good one, though the converse is not, and the inference from *Q'b* to *Q'a* is a good one, though the converse is not. Then although *Qb* results from *Qa* by substituting *b* for *a*, and correspondingly in the case of *Q'a*, so that *a* and *b* have syntactic occurrences in these sentences, *a* and *b* cannot be admitted as having primary substitution-semantic occurrence in these contexts. For there is no simple material substitution-infer-

tial commitment determining the substitution-inferential potential of those sentences. No symmetric SMSIC governs those inferences, since they are not reversible. But no asymmetric SMSIC governs them either, since such a generalization can at most license either the inference from  $Qa$  to  $Qb$  as a replacement of  $a$  by the inferentially weaker  $b$ , or the inference from  $Q'b$  to  $Q'a$  as a replacement of  $b$  by the inferentially weaker  $a$ , but not both. There need be nothing anomalous about such a situation.  $Q\alpha$  and  $Q'\alpha$  are inferentially complementary frames with respect to  $a$  and  $b$ . This precludes codifying substitution inferences involving those expressions in terms of SMSICs relating them, because the generalizations that would be appropriate to  $a$  and  $b$  with respect to  $Q\alpha$  are different from and incompatible with those that would be appropriate with respect to  $Q'\alpha$ . The only cost of not being able to discern semantically primary occurrences of  $a$  and  $b$  in these contexts is that some of the good substitution inferences involving them are not captured by the material contents associated with  $a$  and  $b$ . This is no different than what happens in other cases where one ought not to discern primary substitution-semantic occurrence, for instance in "S believes that  $Qa$ ."

For the next step, suppose that  $Q\alpha$  and  $Q'\alpha$  behave in this inferentially complementary fashion for every pair of substitutional variants in which they appear. That is, suppose there were a pair of predicates  $Q\alpha$ ,  $Q'\alpha$  such that for any substituted-for expressions  $x$  and  $y$ , if the inference from  $Qx$  to  $Qy$  is a good one, but not the converse, then the inference from  $Q'x$  to  $Q'y$  is not a good one. The presence of such a pair of predicates would block the possibility of the substitution-inferential generalization that would be required to give a substitutional commitment asymmetric significance, no matter what substituted-for expressions it involves. This situation would preclude discerning primary substitution-semantic occurrences of *any* substituted-for expression, in sentences exhibiting the substitution frames  $Q\alpha$  and  $Q'\alpha$ . Again, the only cost is that certain proprieties of substitution inference, the ones that involve those frames, will not be projectable based only on the material contents associated with subsentential expressions, crystallized in sets of relational SMSICs.

Finally, in order to see how one might argue against admitting asymmetrically significant substituted-for expressions, strengthen the supposition yet again and consider what happens if there is a general recipe for producing, given any frame  $Q\alpha$ , a frame  $Q'\alpha$  that is inferentially complementary to it, in the prior sense. That is, each  $Q'\alpha$  is to be so constructed that whenever the inference from  $Qx$  to  $Qy$  is good, but not vice versa (intuitively, because  $y$  is inferentially weaker than  $x$ , the way 'mammal' is inferentially weaker than 'dog'), the inference from  $Q'y$  to  $Q'x$  is good, but not vice versa, for any substituted-for expressions  $x$  and  $y$ . Such a situation precludes discerning any primary substitution-semantic occurrences of *any* substituted-for expressions. For there are no syntactic occurrences of any substituted-for expressions whose substitution-inferential significance is correctly captured by an

asymmetric SMSIC (the symmetric ones are not currently at issue). For an asymmetric substitution-inferential commitment relating  $a$  to  $b$  governs inferential proprieties via the generalization that for any frame  $P\alpha$ , the inference from  $Pa$  to  $Pb$  is a good one, though not in general the converse.

Under the hypothesis being considered, no matter what particular instance  $P\alpha$  is chosen, it is possible to construct or choose a complementary predicate  $P'\alpha$  for which only the complementary pattern of substitution-inferential proprieties obtains. In the presence of a recipe for producing for arbitrary substitution frames other frames that are inferentially complementary to them, then, no proprieties of substitution inference can be captured by asymmetric SMSICs, and so no primary substitution-semantic occurrences of substituted-for expressions corresponding to them. The upshot of this line of thought, then, is that the existence of asymmetrically significant substituted-for expressions is incompatible with the presence in the language of expressive resources sufficient to produce, for arbitrary sentence frames, inferentially complementary ones. To explain why substituted-for subsentential expressions have symmetric substitution-inferential significances, which on the current understanding is to explain why there are singular terms, then, it will suffice to explain what sort of expressive impoverishment a language is condemned to if it eschews the locutions that would permit the general formation of inferentially complementary sentence frames.

When it has been seen that the particular constellation of syntactic and semantic roles characteristic of singular terms is necessitated by the presence in the language of vocabulary meeting this condition, it becomes urgent to see what locutions make possible the production of arbitrary inferentially complementary frames, and how dispensable the role they play in linguistic practice might be. What locutions have this power? Examples are not far to seek. The one to focus on is the *conditional*. Because conditionals make inferential commitments explicit as the contents of assertional commitments, inferentially weakening the antecedent of a conditional inferentially strengthens the conditional. "Wulf is a dog" is inferentially stronger than "Wulf is a mammal" because everything that is a consequence of the latter is a consequence of the former, but not conversely. But the conditional "If Wulf is a mammal, then Wulf is a vertebrate" is inferentially stronger than "If Wulf is a dog, then Wulf is a vertebrate." For instance, everything that combines with the first conditional to yield the conclusion that Wulf is a vertebrate also combines with the second to yield that conclusion, but not conversely. Again, endorsing all the inferences from sentences exhibiting the frame " $\alpha$  is a dog" to the corresponding " $\alpha$  is a mammal" does not involve commitment to the goodness of the inferences from sentences exhibiting the frame "If  $\alpha$  is a dog, then  $\alpha$  belongs to an anciently domesticated species" to those exhibiting the frame "If  $\alpha$  is a mammal, then  $\alpha$  belongs to an anciently domesticated species." Instances of the first conditional are true claims expressing correct inferences, while instances of its substitutional variant are

false conditionals expressing incorrect inferences. Quite generally, let  $Qa$  be a particular sentence in which the substituted-for expression  $a$  has primary occurrence, and  $Qb$  be a substitutional variant of it, and let  $r$  be some other sentence. Then  $Qa \rightarrow r$  is a sentence in which  $a$  has primary occurrence, and the symbol  $Q'\alpha$  may be introduced for the sentence frame associated with its occurrence, writing the conditional above as  $Q'a$ . If  $a$  is inferentially stronger than  $b$ , asymmetrically, then the inference from  $Qa$  to  $Qb$  is good, but not its converse (Wulf is a dog, so Wulf is a mammal).<sup>39</sup> But if that is so, then the inference from  $Q'a$  to  $Q'b$  cannot be good, for inferentially weakening the antecedent of a conditional inferentially strengthens the conditional.

This last formulation suggests another example. Inferentially weakening a claim inside a negation inferentially strengthens the compound negation. If the substitution inference from  $Qa$  to  $Qb$  is good but the converse not, then the substitution inference from  $\sim Qa$  to  $\sim Qb$  cannot be good. Embedding as a negated component, like embedding as the antecedent of a conditional, reverses inferential polarities. The conclusion is that any language containing a conditional or negation thereby has the expressive resources to formulate, given any sentence frame, a sentence frame that behaves inferentially in a complementary fashion, thereby ruling out the generalizations that would correspond to asymmetric simple material substitution-inferential commitments governing the expressions that are substituted for in producing such frames.

### *3. The Importance of Logical Sentential Operators*

The conditional and negation are fundamental bits of logical vocabulary. Is it just a coincidence that it is logical sentence-compounding locutions that permit the systematic formation of inferentially inverting sentential contexts? The sentence  $q$  is inferentially weaker than the sentence  $p$  just in case everything that is a consequence of  $q$  is a consequence of  $p$ , but not vice versa (consequences are not preserved but pruned). It is an immediate consequence of this definition that inferentially weakening the premises of an inference can turn good inferences into bad ones. The defining job of the conditional is to codify inferences as claims (make it possible to express inferential commitments explicitly in the form of assertional commitments). It is essential to doing that job that embedded sentences that can play the role of premises and conclusions of inferences appear as components—antecedents and consequents—in the conditional. The contexts in which component sentences occur as antecedents accordingly must be inferentially inverting. Notice that this argument presupposes very little about the details of the use of the conditional involved. It is enough, for instance, if the conditional has the designated (semantic or pragmatic) status in case the inference it expresses preserves the designated status. As the defining job of the conditional is to codify inferences, that of negation is to codify incompatibilities. The negation of a claim is its inferentially minimal incompat-

ible— $\sim p$  is what is entailed by everything materially incompatible with  $p$ .<sup>40</sup> These underlying incompatibilities induce a notion of inferential weakening: “Wulf is a dog” incompatibility-entails, and so is inferentially stronger than, “Wulf is a mammal” because everything incompatible with “Wulf is a mammal” is incompatible with “Wulf is a dog” but not vice versa (incompatibilities pruned, not preserved). It follows that incompatibility-inferentially weakening a negated claim incompatibility-inferentially strengthens the negation. “It is not the case that Wulf is a mammal” is incompatibility-inferentially stronger than “It is not the case that Wulf is a dog,” just because “Wulf is a mammal” is incompatibility-inferentially weaker than “Wulf is a dog.” Thus negation also enables the formation of arbitrary inferential complements. It was argued in Chapter 2 that what makes both conditionals and negation, so understood, specifically logical vocabulary is that the material inferences and material inference-inducing incompatibilities, of which they permit the assertorily explicit expression, play a central role in conferring material contents on prelogical sentences. It is a direct result of this defining semantically expressive function that they form semantically inverting contexts.

Since it is the availability of such contexts that rules out asymmetrically significant substituted-for expressions, it follows that a language can have either the expressive power that goes with logical vocabulary or asymmetrically substitution-inferentially significant substituted-for expressions, but not both. It is leaving room for the possibility of logical locutions that enforces the discrimination of singular terms (and as a consequence, of predicates) rather than some other sorts of subsentential expression. This conclusion would not be surprising if the logical vocabulary in question were that employed to make explicit the substitution-inferential relations in virtue of which singular terms and predicates can be assigned distinct material contents. For, as has already appeared, the symmetrically significant SMSICs associated with singular terms can be made assertorily explicit with the use of identity locutions, and the asymmetrically significant SMSICs associated with predicates can be made explicit with the use of quantificational locutions (together with sentential logical vocabulary). But like all logical locutions, the use of these presupposes, and so ought not be appealed to in trying to explain, the material contents of the nonlogical expressions that are explicated. In short, the use of identity and quantificational locutions presupposes singular term and predicate use. So of course any language whose use is sufficient to confer on expressions the significance of such locutions must already have in play the symmetric and asymmetric SMSICs associated with nonlogical subsentential expressions, and the expressions whose use they governed will be singular terms and predicates respectively.

But these are not the logical locutions appealed to in the argument against asymmetrically significant substituted-for expressions. On the contrary, the only logical locutions required for that argument are those whose roles are

definable solely in terms of the behavior of *sentences*, before any sort of subsentential substitutional analysis has been undertaken. The argument does not depend on any particular features of the sentential contents that are available to begin with, determining the proprieties of material inference that provide the targets for substitutional codification in (implicit or explicit) SMSICs. All that matters is the availability of the expressive power of logical sentential connectives.

There is of course no absolute necessity that such vocabulary be available in a language. It would be a mistake to conclude from the true premise that something can be thought of as *propositionally contentful* only in virtue of its relation to proprieties of *inferential* practice, to the conclusion that such practice must be *logically* articulated. Such a move depends on the formalist error of assimilating all correctnesses of inference to logical correctness of inference, thereby denying the possibility of material, content-conferring inferential proprieties. Material proprieties of inference are antecedent to formal proprieties in the order of explanation, because to say that an inference is valid or good in virtue of its K form (for instance logical form) is just to say that it is a (materially) good inference, and it cannot be turned into one that is not good by replacing non-K (for instance nonlogical) vocabulary by (syntactically cocategorial) non-K vocabulary. There is nothing incoherent about a language or stage in the development of a language in which the only vocabulary in play is nonlogical.<sup>41</sup> But insofar as the material contents associated with substituted-for expressions are introduced and modified in units corresponding to asymmetrically significant SMSICs involving those expressions, the language containing them is not just in fact bereft of the expressive power of logical vocabulary, it is actually precluded from acquiring it (until and unless the offending subsentential semantic structure is reorganized). That is a reason not to have substituted-for expressions behave this way semantically, even in languages in which logical locutions have not yet been introduced.

Having to do without logical expressions would impoverish linguistic practice in fundamental ways. The use of any contentful sentence involves implicit commitment to the (material) correctness of the inference from the circumstances of appropriate application associated with that sentence to the consequences of such application. Introducing conditionals into a language permits these implicit, content-conferring, material-inferential commitments to be made explicit in the form of assertional commitments. This is important at the basic, purely sentential level of analysis for the same reason it becomes important later at the subsentential level, when identity and quantificational locutions can be introduced to make explicit the SMSICs that confer distinguishable material-inferential content on subsentential expressions. In each case, once made explicit in the form of claims, those content-conferring commitments are brought into the game of giving and asking for reasons. They become subject to explicit objection, for instance by

confrontation with materially incompatible assertions, and equally to explicit justification, for instance by citation of materially sufficient inferential grounds. The task of forming and nurturing the concepts we talk and think with is brought out of the dim twilight of what remains implicit in unquestioned practice into the daylight of what becomes explicit as controversial principle. Material contents, once made explicit, can be shaped collectively, as interlocutors in different situations, physically and doxastically, but in concert with their fellows, provide objections and evidence, claims and counterclaims, and explore possible consequences and ways of becoming entitled to assert them. Logic is the linguistic organ of semantic self-consciousness and self-control. The expressive resources provided by logical vocabulary make it possible to criticize, control, and improve our concepts. To give this up is to give up a lot.<sup>42</sup> Yet, it has been argued, it is a direct (if unobvious) consequence of leaving open the possibility of introducing such inferentially explicating vocabulary that the subsentential expressions that are substituted for will be singular terms, and their corresponding sentence frames will be predicates, as judged by the symmetric and asymmetric forms of their respective substitution-inferential significances.<sup>43</sup>

## VI. OBJECTIONS AND REPLIES

### 1. *Are Singular Terms Symmetrically Substitution-Inferentially Significant Substituted-For?*

Before pursuing further the significance of this result, it will be well to look a little more closely at the argument that has been offered for it. There are several sorts of objection it might elicit, which call for different sorts of clarification. The argument went like this. Singular terms and predicates were distinguished, as essentially subsentential expressions, by the coincidence of a particular syntactic substitutional role with a particular semantic substitutional role. Singular terms are substituted *for*, and the significance of an occurrence is determined by the content associated with it in the form of a set of *symmetric* simple material substitution inferential commitments linking that term to others. Predicates are substitutional sentence *frames*, and the significance of an occurrence is determined by the content associated with it, in the form of a set of *asymmetric* simple material substitution inferential commitments linking that predicate to others. So to ask why there are singular terms is to ask why these two sorts of substitutional roles coincide as they do.

The possibility that both what is substituted for and the resulting sentence frames might have symmetric substitutional significance is dispatched by the observation that the sentences displayed by this analysis as substitutional variants of one another stand in asymmetric substitution-inferential relations to one another. So the question, Why are there singular terms?

reduces to the question of why syntactically substituted-for expressions have symmetric rather than asymmetric substitution-inferential significances. For a substituted-for expression to have such a significance is for the contribution it makes to the substitution-inferential potential of sentences in which it occurs to be articulated as a set of SMSICs relating it to other expressions. An asymmetric SMSIC relating  $t$  to  $t'$  dictates that for any sentence  $Pt$  in which  $t$  has primary substitution-semantic occurrence, the inference from  $Pt$  to  $Pt'$  is a good one, though not conversely, where  $Pt'$  is a substitutional variant of  $Pt$  resulting from it by substituting  $t'$  for  $t$ . It was then pointed out that any language expressively rich enough to contain conditional or negating locutions—in short any language equipped with the conceptual resources of elementary sentential logic—precludes the existence of any such asymmetrically significant SMSICs governing substituted-for expressions. For those resources suffice to ensure that whenever, for some sentence frame  $P\alpha$ , the inference from  $Pt$  to  $Pt'$  is a good one but not conversely, there is another sentence frame  $P'\alpha$  such that the inference from  $P't'$  to  $P't$  is a good one but not conversely. It follows that in such languages discerning asymmetrically significant substituted-for expressions would codify no inferences at all, permit no assignment of material content, in the form of a set of SMSICs, to any substituted-for expression. This way of combining syntactic and semantic substitutional roles is accordingly ruled out, and only that actually instantiated by singular terms and predicates remains. Thus the availability of the expressive capabilities of sentential logic dictates that subsentential substitutional analysis discern singular terms and predicates, rather than some other fundamental subsentential categories. It will be helpful to consider some different kinds of objection that might arise to this argument.

One potentially vulnerable premise is the definition of singular terms (and so of predicates) relied upon. The key final move in the argument, invoking the expressive power of logical locutions to form new sentences (and so new sentence frames) purports to show only why expressions that play the syntactic role of substituted-for must play the semantic role of being symmetrically substitution-inferentially significant. It might be denied that this result (for the moment treating the argument as successful) shows anything of importance about singular terms. For it might be denied that the syntactic and semantic substitutional roles that have been identified as characterizing this subsentential category of expression pick out its fundamental properties.

This complaint can be subdivided. It might be denied that the dual syntactic-semantic substitutional role attributed to singular terms is one they actually play—that is, that playing that dual role provides at least a necessary condition for expressions to be functioning as singular terms. This seems a difficult line to pursue. Singular terms are substituted for in sentences, generating derivative substitutional sentence frames. These two roles are clearly discriminable by the fact that the resulting frames have fixed possibilities of combination with substituted-for expressions (fixed adicities), always being

exhibited by sentences with the same number of substituted-fors,<sup>44</sup> while those latter expressions are subject to no analogous constraint. Of course, singular terms can also play the role of the expression that is substituted in. Thus there can also be expressions of the derivative category of singular term frames, such as "the father of  $\alpha$ ," which are substitutional patterns according to which singular terms can be assimilated.

This much is not special to substituted-for expressions. There is no reason " $\alpha$  wrote about  $\beta$ " cannot be understood as a one-place sentence-frame frame, for instance that exhibited by sentence frames such as " $\alpha$  wrote about Rousseau" and " $\alpha$  wrote about Kant," as well as its playing the role of two-place sentence frame exhibited by any sentences exhibiting any of those one-place frames.<sup>45</sup> Sentences, in virtue of the direct pragmatic (paradigmatically assertional) significance of their use, are the original substituted-ins, the ones that get the whole substitution-analytic enterprise off the ground. Subsentential expressions are pragmatically significant, and so semantically contentful only in a substitutionally indirect sense (for instance, it is only via their substitutional relations with other expressions that they can be said to stand in inferential or incompatibility relations). Thus the fact that they can be substituted in, in no way qualifies the characterization of singular terms as substituted for. By contrast to the case of sentences, it is only because terms can play the latter syntactic substitutional role that they can play the former one.

Alternatively, one might argue that playing the semantic substitutional role of being governed by symmetrical SMSICs does not provide a necessary condition for a substituted-for expression to be a singular term. Here there are two possible cases. The claim might be that some terms are governed by asymmetrical SMSICs, or it might be rather that the use of some genuinely occurring singular terms is not governed by SMSICs at all.

For a first try at constructing singular terms that stand in one-way, or asymmetrical, inferential relations, one might consider definite descriptions of the following form: {the  $x$ }[ $Dx$ ] and {the  $x$ }[ $Dx \& Fx$ ]. This, however, will not provide a counterexample to the symmetry claim. To generate such a counterexample, it would have to be the case that in general either

- (1)  $P(\{\text{the } x\}[\text{D}x])$  entails  $P(\{\text{the } x\}[\text{D}x \& \text{F}x])$ ,  
and not always vice versa

or

- (2)  $P(\{\text{the } x\}[\text{D}x \& \text{F}x])$  entails  $P(\{\text{the } x\}[\text{D}x])$ ,  
and not always vice versa.

But (1) does not hold because an object may qualify as {the  $x$ }[ $Dx$ ] without thereby qualifying as {the  $x$ }[ $Dx \& Fx$ ], namely in those cases in which it is not also  $F$ . Description (2) does not hold because an object may qualify as {the  $x$ }[ $Dx \& Fx$ ] without thereby qualifying as {the  $x$ }[ $Dx$ ], namely in those cases

where, though only one thing is both *D* and *F*, more than one thing is *D*. These correspond to failures respectively of the descriptive and definiteness conditions on the applicability of definite descriptions.

## 2. Asymmetrically Substitution-Inferentially Related Singular Terms?

Seeing where this sort of prospective counterexample falls down suggests a much more serious prospect.<sup>46</sup> Consider the relation between pairs of terms such as 'Benjamin Franklin' and 'Benjamin Franklin, who was a printer'. Anyone who is committed to the claim:

Benjamin Franklin, who was a printer, invented lightning rods  
is thereby committed to the corresponding claim:

Benjamin Franklin invented lightning rods,

but not everyone who is committed to the second claim is thereby committed to the first. In particular, anyone who has no opinion about whether Benjamin Franklin was a printer, or who denies that he was, can endorse the second without the first.

Such examples point out that there is a kind of term-forming operator that, given a term *t*, produces a compound term of the form '*t*, who (or which)  $\Phi$ s', which systematically induces asymmetric substitution inferences relative to the first. The existence of such a recipe for producing terms from terms requires a genuine restriction in the scope of the claims made about terms so far. The example shows that there are some pairs of terms and some sentence frames such that substitution inferences relating those terms in those sentence frames are asymmetric. Acknowledging the existence of such cases, however, does not require relinquishing any claims on which the preceding argument relies.

As a preliminary, it may be noticed that the asymmetry pointed to here involves only relations to compound terms of a special form, and not terms generally—by contrast to predicates, *all* of which stand in one-way inferential involvements to other predicates that are *not* formed from them at all, never mind in a particular way. Again, the formation of terms of the form '*t*, who  $\Phi$ s' presupposes the existence of terms not of that form, which can appear in the ordinary predication of the form  $\Phi t$ , from which these special compound terms are formed (of course, *t* could itself be of the form '*t*, who  $\Psi$ s', but it cannot in principle be terms of this form all the way down). Thus the existence of terms that behave this way is parasitic on the existence of terms that do not.

But the point of central importance is that, as could be deduced already from the argument concerning recipes for generating logically compound predicates that reverse inferential polarity, it is *not* the case even for these

special pairs of terms that their substitution-inferential significance is governed by asymmetric SMSICs. To be so governed involves suitable substitutional variants standing in one-way inferential involvements for *all* predicates. And someone who is committed to the claim:

It is *not* the case that Benjamin Franklin, who was a printer, invented lightning rods

is *not* thereby committed to the claim:

It is *not* the case that Benjamin Franklin invented lightning rods.

Anyone who denies that Benjamin Franklin was a printer but who believes he did invent lightning rods can endorse the first claim but not the second. A similar point will hold for sentence frames formed from conditionals in which the terms in question appear in the antecedent. Thus pairs of terms of the form  $\langle t, \text{ and } t, \text{ who } \Phi s \rangle$  fail to be *systematically* asymmetrically substitution-inferentially significant in the way predicates are. So the symmetric and asymmetric *patterns* of substitution-inferential significance can still be appealed to in distinguishing the category of singular terms from that of predicates.

The restriction of asymmetric significance to some sentential contexts points up that these cases belong in a box with asymmetric inferences such as that from:

The skinniest person in the room can't fit through the narrowest door

to

The fattest person in the room can't fit through the narrowest door,

which are also not reversible.<sup>47</sup> These examples clearly turn on interactions between the predicates used to form definite descriptions and those involved in the sentence frame in which the description is embedded. Just by their nature, such asymmetries do not generalize across sentence frames generally and so have no systematic significance of the sort appealed to in the substitutional account of the difference between singular terms and predicates.<sup>48</sup>

### 3. *Symmetrically Substitution-Inferentially Related Expressions That Are Not Singular Terms?*

The second possible form of complaint about the necessity of the semantic condition of symmetrical substitutional significance stems from the observation of expressions that intuitively seem to be singular terms but that do not play a symmetric substitutional role. Consider occurrences of

expressions that function syntactically as singular terms, that is, that are intersubstitutable, saving sentential well-formedness, with substituted-fors that also have primary substitution-semantic occurrences.<sup>49</sup> The semantic significance of those occurrences could be of any of four different kinds. First, they might be primary substitution-semantic occurrences. In that case, which of the substitution inferences the expression is materially involved in are taken to be good or correct is governed by the SMSICs undertaken by the primary assessor of those inferences—typically, us (who are discussing them). These are the occurrences here treated as characteristic and fundamental.

Second, at the other end of the spectrum, there can be occurrences of what are syntactically singular terms that are semantically inert, for instance appearances in direct quotation, or that are unprojectable, such as the ‘it’ in “It is raining.” The right thing to say about these seems to be that they are singular terms merely in a syntactic sense—that is, syntactically substitutable for full-blooded singular terms—but not functioning semantically as singular terms. The proposed substitutional definition of the dual functional role of singular terms explicitly leaves room for this sort of occurrence.

Third, in between primary substitution-semantic occurrences and semantically accidental occurrences, are what may be called secondary semantic occurrences. These have a systematic semantic significance for the propriety of inferences, and it is determined by a set of (symmetric) SMSICs. But they fail to qualify as primary substitution-semantic occurrences because the SMSICs relevant to the assessment of the propriety of inferences in which the expressions are materially involved are not the SMSICs associated with the one assessing those inferences. The central examples are occurrences in opaque positions in propositional attitude ascriptions. Thus my assessment of the propriety of the inference from “Carlyle believed (or acknowledged commitment to the claim) that Kant ascribed to each of us a duty to make ourselves perfect and others happy” to “Carlyle believed (or acknowledged commitment to the claim) that the author of *Dreams of a Spirit-Seer* ascribed to each of us a duty to make ourselves perfect and others happy” does not depend on whether, according to me, Kant is the author of *Dreams of a Spirit-Seer*; that is, it does not depend on the SMSICs that govern my primary uses of those expressions. But it is governed by some (symmetric) SMSICs—namely those I attribute to Carlyle, rather than those I undertake myself (that is, take to be true).

The definition of singular term in the full-blooded sense could be sharpened so as to acknowledge explicitly the possibility of secondary semantic occurrences. Since these are defined by government by (symmetric) SMSICs, such a clarification, qualification, or emendation is entirely in the spirit of the original proposal. The only kind of occurrence of expressions that syntactically qualify as singular terms that would be troublesome to the definition offered is if there should be a fourth sort of occurrence, intermediate between secondary semantic occurrences and semantically accidental ones.

These would be occurrences that were systematically significant for the goodness of inferences involving sentences they occur in, but not according to the pattern of government by SMSICs (symmetric or not). It is hard to know what to say about this possibility, in the absence of a specific candidate. The response to such a candidate ought to be to try to show that that role, like that of secondary semantic occurrences of singular terms, presupposes and is built on their primary role as governed by symmetric SMSICs. The possibility of a counterexample from this direction cannot simply be dismissed, though.

#### *4. Does the Definition of Singular Terms Offer Sufficient Conditions?*

A more promising line of attack might seem to be to focus on the sufficiency rather than the necessity of the characterization—to accept that singular terms do play the dual syntactic and semantic substitutional role attributed to them, but to deny that specifying these roles by itself suffices to pick out singular terms. Here the idea might be that there are other, genuinely central or essential roles of singular terms that are not derivable from those substitutional roles already mentioned, in that there could be expressions that satisfied the substitutional criteria but did not qualify as genuine singular terms through failure to play the genuinely critical roles. For instance, it might be noticed that the referential, rather than the inferential, role of singular terms has been ignored. This objection may be responded to in two ways—concessively and aggressively.

The concessive response would be to acknowledge that the characterization of singular terms omits features that may turn out to be essential, so that that characterization does not suffice to pick out singular terms. It may immediately be pointed out, however, that such an admission would not be very damaging to the argument presented. For that argument still goes through for an important class of expressions, call them ‘ralugnis mrets’, which contains singular terms as its most prominent proper subset. Even if this class should turn out to be substantially wider than that of singular terms, the question, Why are there any ralugnis mrets? would still be a fundamental one and would have to be addressed as the first part of an explanation of why there are singular terms. For that question would then assume the form “Why are there ralugnis mrets that also have the additional discriminating feature  $F$  (the hypothetical specific difference required to distinguish singular terms within the broader class of ralugnis mrets)?”

The aggressive response, though, is more interesting and potentially more satisfying. It is to attempt to derive the remaining features of singular-term usage from the dual substitutional characterization of them as symmetrically and substitution-inferentially significant substituted-for expressions. (Recall the suggestion earlier that having its substitution-inferential significance

determined exclusively by symmetric SMSICs can help provide an analysis of purporting to refer to or pick out one and only one object.) The explanatory strategy to which the present work (in particular, Part 2) is conceived as a contribution is to try to explain the referential and representational relations linguistic expressions stand in by appeal to the inferential, substitutional, and anaphoric relations they stand in. (These are three successive levels of analysis of material-sentential contents, each one of which presupposes the prior levels.) This is an ambitious undertaking, to be sure, and one this chapter does not make it possible fully to assess, since the substitutional relations discussed here provide only some of the necessary raw materials. Any difficulties of principle that arise in pursuing it will be forces pushing toward the concessive, rather than the aggressive, response to objections to the sufficiency of the characterization of singular terms.

But it would seem to be good Popperian methodology to adopt the strongest, most easily falsifiable, apparently sustainable commitment, to see where it falls down. In this case, concrete reason to be pessimistic about deriving the remaining representational properties would be concrete information about the missing feature *F*, which must be added to the substitutional characterization of singular terms to make it sufficient to pick them out within the wider class of ralugnis mrets. Detailed pursuit of the aggressive explanatory enterprise requires two further discussions: of anaphoric commitments that link unrepeatable tokenings and explain how anaphoric inheritance according to such links determines which SMSICs govern the use of those tokenings, and of how those links are made explicit in the form of *de re* ascriptions of propositional attitude. (*De re* ascriptions, such as "Carlyle believed of or about the author of *Dreams of a Spirit-Seer* that he ascribed to each of us a duty to make ourselves perfect and others happy" are the primary locutions with which we discuss representational relations. They permit us to talk about what people are talking about, and so provide a suitable explanatory target for an account of representational properties of linguistic expressions.)

### 5. *Can the Same Argument Be Used to Show That Sentences or Predicates Must Have Symmetric Inferential Significances?*

Another line of attack would be to object that the argument cannot be correct, for if it were, it would prove too much. For, it might be asked, why cannot one apply exactly the same considerations used to show that substituted-for *terms* must have symmetric inferential significances in order to show that *sentences* must have symmetric inferential significances? Sentences can be substituted for, as well as substituted in; they can appear as antecedents of conditionals or as negated; but their occurrences have an asymmetric inferential significance. Why does this not show that there must

be some defect in the argument purporting to show that there cannot be asymmetrically significant substituted-for expressions in a language containing logically expressive locutions? It does not because, unlike singular terms and predicates, sentences are not essentially subsentential expressions. Since sentences can occur freestanding, possessing pragmatic significance all on their own, their semantic content is not to be understood exclusively in terms of the contribution their occurrence makes to the inferential behavior of compound sentences in which they occur as significant components.

Sentences have a directly inferential content, since they can play the role of antecedent and consequent in inferences. Essentially subsentential expressions such as singular terms and predicates, by contrast, have an inferential content at all only in a substitutionally indirect sense. The inferentially articulated pragmatic significance of asserting a sentence is asymmetric already in the direct sense, since swapping premises for conclusions does not in general preserve the correctness of inferences. Considered just as expressions that are substituted for in other expressions—that is, just as embedded expressions—sentences need only be sorted into equivalence classes of various sorts, accordingly as they are symmetrically intersubstitutable preserving pragmatically important properties of the compound sentences in which they are embedded. This is exactly the procedure that was employed in Sections I and II, which investigated just such embedded occurrences of sentences and the sorts of subsentential sentential content they involve. Because the same sentences that stand in one-way inferential involvements as freestanding also are sorted into equivalence classes as substituted for, those equivalence classes can accordingly then be sorted into asymmetric families.

But this is entirely a consequence of their roles as premises and conclusions, and there can be no analog of this role for essentially subsentential expressions. As the discussion in the opening sections shows, the way to move from the sort of content associated with freestanding uses of sentences to that associated with embedded uses is broadly substitutional in nature. But it does not take the form of government by simple material, substitution-inferential commitments, which is all that is available for essentially subsententially occurring expressions. In Section II two sorts of sentential context of embedding for sentences were distinguished as inferentially ‘extensional’ in different senses, accordingly as intersubstitutabilities that preserve pragmatic statuses are determined by the freestanding assertional and inferential behavior of the component sentences, respectively. Both of these notions of extensional context depend essentially on how the expressions that occur in those contexts behave freestanding. Thus these notions do not apply to essentially subsentential expressions. Yet it is only by conflating them with the related but distinct notion of primary substitution-semantic occurrence (exclusive government by SMSICs), which applies to essentially subsentential expressions, that it could seem that the argument used to

forbid asymmetric significance of essentially subsentential substituted-for expressions would forbid it also to substituted-for sentences.

The objection was that the argument that essentially subsentential substituted-for expressions cannot have asymmetric substitution-inferential significance proves too much, for if good it would also show that sentences, which can also be substituted for, cannot have asymmetric substitution-inferential significance. That objection can be seen to fail because it ignores the further semantic resources available for expressions that are inferentially significant in a direct sense, as well as the substitutionally indirect one that the argument addresses. This immediately suggests a further objection that would not be vulnerable to this sort of response: to modify the previous objection by appealing to predicates, rather than sentences, to try to show that the argument for the symmetric significance of substituted-for expressions would prove too much. For predicates are essentially subsentential expressions, and the argument acknowledges that they do have asymmetric substitution-inferential significances, even in languages that have conditionals and negation in them. Although they are not expressions that are substituted for—indeed, as substitution frames, they are defined explicitly by their contrasting and complementary syntactic substitutional role—it must nonetheless be conceded by the argument that it makes sense to talk of predicates being replaced. For the semantic substitutional characterization of predicates as having the indirect inferential significance of their occurrence determined by families of asymmetric SMSICs presupposes such a notion. Why does not the argument from the expressive necessity of inferentially inverting sentential contexts to government of essentially subsentential substituted-for expressions by symmetric SMSICs go through equally well for essentially subsentential, merely replaceable, expressions—in particular, predicates? This is really another way of asking what role is played in the argument by the fact that the essentially subsentential expression involved is syntactically categorized as substituted for, and not as a substitution frame.

The argument was that for the simple material, substitution-inferential commitments associated with singular terms to govern not only occurrences in logically atomic sentences but also occurrences in embedded, inferentially inverting, logically articulated sentences, the significances associated with expressions by those commitments must be symmetric. Since asymmetric sentential substitution inferences are to be projected on the basis of the substitutional decomposition of sentences into compounds of essentially subsentential expressions, it follows that the other, derived, substitutional syntactic category—sentence frames—must be governed by SMSICs having asymmetric significance. It is now objected that there must be something amiss with this argument, for not only terms but predicates are recognizable when embedded in the antecedent of a conditional, or inside a negated sentence. So how is it possible for their occurrence to have asymmetric significance?

Here it is important to keep clearly in mind the distinction between substitutionally defined sentence frames and the predicate-letters or other expressions used as marks for them. Predicates, as substitution frames, are defined to begin with as equivalence classes of sentences, those that can be turned into one another by substitution for expressions of the substitutionally basic category.<sup>50</sup> ' $P\alpha$ ', as a one-place sentence frame has been represented here, stands for such an equivalence class. Suppose it is the class of which "Carlyle wrote *Sartor Resartus*" and "Ruskin wrote *Sartor Resartus*" are members, because substitutional variants of the 'Carlyle'  $\leftrightarrow$  'Ruskin' sort. But "If Carlyle wrote *Sartor Resartus*, then Carlyle had an ambiguous relationship with Hegel" is not obtainable as the result of substituting terms for terms in these sentences. It is not a member of the equivalence class denoted by ' $P\alpha$ '. In short, while the term 'Carlyle' has primary semantic occurrence in the conditional, as well as in its antecedent, the sentence frame " $\alpha$  wrote *Sartor Resartus*" has primary syntactic, as well as semantic, occurrence only in the antecedent (when freestanding), and not in the conditional as well. This is reflected in the fundamental difference between *substituting* for an expression and *replacing* it, as those words have been used here. What is replaced is always a frame associated with the whole sentence.<sup>51</sup> Thus although there is a sense in which a sentence can exhibit more than one sentence frame—for instance, "Carlyle wrote *Sartor Resartus*" exhibits not only " $\alpha$  wrote *Sartor Resartus*" but also " $\alpha$  wrote  $\beta$ "—it is not in general possible to replace one of them, resulting in a sentence that still exhibits others of them. By contrast, a sentence may contain many occurrences of different substituted-for expressions, and substitution for one of them by and large results in sentences in which the rest still occur.<sup>52</sup> Thus frames have adicities, and substituted-fors do not.

It is a joint consequence of the requirement that primary semantic occurrences of either substituted-for expressions or of substitutional frames must have asymmetric significances, and the requirement that anything that has primary semantic occurrence both in freestanding sentences and embedded in inferentially inverting contexts such as negations and the antecedents of conditionals must be governed by symmetric SMSICs, that substituted-for and substitutional frames cannot both have primary semantic occurrence (the kind whose significance is governed by the SMSICs associated with the expression) in those embedded contexts. It is a consequence of the substitutional definition of sentence frames that they do not have primary occurrence in those contexts. This is not to say, of course, that no connection can be recognized between "X wrote *Sartor Resartus*" and "If Carlyle wrote *Sartor Resartus*, then Carlyle had an ambiguous relationship with Hegel" (if that were so, it would not be possible to state modus ponens without equivocation). On the contrary, the former is a frame exhibited by freestanding occurrences of the sentence that appears as antecedent in the latter conditional. Logical vocabulary produces inferentially inverting contexts because of its expressive role in making explicit inferential and incompatibility relations

among sentences. That same role guarantees that the contribution made to the use of a logical compound by the occurrence in it of component sentences is determined by the freestanding inferential role of those component sentences, which is what is being made explicit.<sup>53</sup> Thus the only contribution the predicate occurring in the antecedent need make to the projection of a semantic role for the conditional is to help project the semantic role of the antecedent as a freestanding, unembedded sentence. This much, however, is settled by discerning primary semantic occurrences of the predicate (replaceable ones whose significance is governed by SMSICs) only in the antecedent as a freestanding sentence. Thus in fact no semantic projectability is lost by refusing to discern in conditionals occurrences of the sentence frames exhibited by their antecedents.

#### *6. Why Must It Be Possible to Substitute for Singular Terms in Logically Compound Sentences?*

Offering this response to the objection naturally elicits a question as rejoinder: Why not treat occurrences of singular terms this way? That is, why not treat them as having their significance determined by the two-step process of examining first their contribution to the freestanding use of sentences and then the contribution of that freestanding use to the use of logically compound sentences in which they occur? The basic argument shows that both syntactic substitutional kinds of essentially subsentential expression cannot be taken to have primary semantic occurrence in conditionals; and it has been shown that, as defined, sentence frames do not. But what is the warrant for the asymmetry in treatment of the two syntactic substitutional sorts? Why see even terms as having primary semantic occurrence in conditionals, as well as in their antecedents? Asking this question makes it possible to highlight what is in some sense the heart of the difference between substituting for substitutionally basic subsentential expressions and replacing substitutionally derived sentence frames. There is an expressive reason for insisting on discerning primary semantic occurrences of singular terms in logically compound sentences, of which conditionals are paradigmatic, which has no analog for predicates.

To say that 'Carlyle' can play the role of substituted-for expression in "If Carlyle wrote *Sartor Resartus*, then Carlyle had an ambiguous relationship with Hegel" as well as in "Carlyle wrote *Sartor Resartus*" is to say that 'Ruskin', for instance, can be substituted for it in that context, to yield *conditional* substitutional variants, such as "If Ruskin wrote *Sartor Resartus*, then Ruskin had an ambiguous relationship with Hegel." These conditional substitutional variants define a conditional substitution frame: "If  $\alpha$  wrote *Sartor Resartus*, then  $\alpha$  had an ambiguous relationship with Hegel." The question accordingly becomes, What is the special virtue of discerning such frames? To yield the result in question, it must be a virtue that does not correspondingly attach to discerning the second-order conditional 'frames'

that would result from assimilating conditionals that were variants under replacement of first-order frames occurring in their antecedents—what might be symbolized in the example by “If  $\Phi\alpha$ , then  $\alpha$  has an ambiguous relationship with Hegel.”

The virtue in question is an expressive one, namely that conditional sentence frames must be discerned if the conditional locution that is used to make explicit material-inferential relations among sentences is to be able also to make explicit material substitution-inferential relations among sentence frames.<sup>54</sup> The material content associated with sentence frames, in virtue of which discerning expressions playing that essentially subsentential substitutional role contributes to the projectability of material-inferential contents for novel sentences, can be factored into simple material substitution-inferential commitments relating frames to other frames. When these content-conferring commitments remain implicit—that is, do not take the form of assertional commitments—they determine the significance of replacing one frame by another. When they are made explicit as the contents of logically articulated sentences, it is as quantified conditionals. It is a necessary condition of introducing quantificational locutions that one be able to discern the conditional sentence frames that are the quantificational substitution instances. In a language that has the logical expressive resources supplied by quantifiers, a typical predicate SMSIC might be made explicit as follows:  $\{x\}[\text{if } x \text{ is a whale, then } x \text{ is a mammal}]$ . Even in a relatively expressively impoverished language, one that still lacks the subsentential quantificational logical locution, if it has the sentential logical expressive resources supplied by the conditional, then the ability to project the semantic contents of novel material conditionals requires the ability to discriminate the substitutional variants that make up the conditional sentence frame “If  $\alpha$  is a whale, then  $\alpha$  is a mammal.” That is, in a language containing a conditional, mastery of the material inferential content associated with the logically atomic frame “ $\alpha$  is a whale” requires being able to assimilate conditional sentences into frame-equivalence classes accordingly as they are mutually accessible by substitution for expressions of the basic substitutional syntactic category. It is this mastery that may then be made explicit in the form of quantified conditionals.

The picture that emerges, then, is that substituted-fors and substitution for them (substitution mappings indexed by them) must be discerned in order to define frames in the first place, and again in defining replacement for them. Frames need to have material contents associated with them in order for singular terms to do so because projection requires the cooperation of both. Furthermore, logical frames are needed in order to make explicit the material contents associated with singular terms because identity locutions are such frames. *Conditional* (and negated)—that is, logically compound—frames are needed to make explicit the material substitution inferential and incompatibility contents of frames. Conditional sentence-frames—formed by assimilat-

ing conditionals according to accessibility relations defined by substituting for various basic essentially subsentential expressions—are required in order to codify explicitly the SMSICs that govern nonlogical first-order sentence frames.

There is no comparable necessity to be able to distinguish logically compound second-order frames by assimilating conditionals according to replacement of substitutionally derivative first-order frames in antecedents. Doing so is not required in order to codify and express explicitly the material contents either of singular terms or of first-order frames.<sup>55</sup> So essentially subsentential substituted-for expressions must be taken to have primary semantic occurrence in conditionals, as well as in the sentences embedded in those conditionals. For the corresponding conditional sentence frames articulate the substitutionally indirect, material-inferential contents governing the significance of primary semantic occurrences of nonlogical sentence frames. This reason does not analogously require that sentence frames be taken to have primary semantic occurrence in logically compound sentences.

Since it has been shown that discerning primary semantic occurrences of an expression both in the nonlogical sentences whose content is explicitated and in the inferentially inverting logical sentences that explicitate them requires that those occurrences be governed symmetrically by SMSICs, and that either frames or substituted-for expressions must be governed asymmetrically, it follows that frames and substituted-for expressions cannot both be taken to have primary semantic occurrence in logically compound sentences. Taken together, these arguments show why it is substituted-for expressions, and not the resulting substitutionally derivative sentence frames, that have the pattern of primary semantic occurrence across logically atomic and logically compound sentences that requires government by symmetric simple material substitution-inferential commitments. Since singular terms are essentially subsentential expressions that play the dual syntactic substitutional role of being substituted for, and the semantic substitutional role of having a symmetric substitution-inferential significance, this explains why there are singular terms.

### *7. Can the Substitutional Significance of the Occurrence of a Subsentential Expression Be Determined in Different Ways for Different Contexts?*

One final objection should be considered. This stems from the thought that the inferential patterns associated with SMSICs are too rigid. Why should it be required that *all* of the primary occurrences of subsentential expressions have their significance determined by a SMSIC in the *same* way? If this requirement is relaxed, it seems that there is a way to evade the argument against the possibility of asymmetrically substitutional commitments governing expressions that are substituted for. Suppose, then, that the

terms  $a$  and  $b$  are linked, not by a symmetric relation of intersubstitutability that could be made explicit by an assertible substitution license in the form of an *identity*  $a = b$  (defining an equivalence class), but by an asymmetric relation of *domination* that could be made explicit by an assertible substitution license in the form of an *inequality*  $a \triangleright b$ . When this possibility was considered above, it was insisted that the significance of such a relation be that for *all* predicates or sentence frames  $P\alpha$ , if  $Pa$  then  $Pb$ , but not necessarily vice versa. The suggestion being considered now is that this requirement be relaxed.

Suppose that the logically atomic predicates are sorted into two classes, according to their *inferential polarity*. If  $P\alpha$  has *positive* inferential polarity, then if  $Pa$ , then  $Pb$ , but not necessarily vice versa, just as before. By contrast, if  $P\alpha$  has *negative* inferential polarity, then if  $Pb$ , then  $Pa$ , but not necessarily vice versa, under the same assumption that  $a \triangleright b$ —that is, that  $a$  dominates  $b$ . It may turn out (but on the assumption being considered it need not) that all the logically atomic sentence frames have positive polarity. In any case, as the crucial step in the overall argument shows, there will be some logically compound sentence frames that have negative polarity. If  $P\alpha$  has positive polarity, then  $\sim P\alpha$  and  $P\alpha \rightarrow r$  have negative polarity. To keep track, one might express all the logical compounds in disjunctive normal form and count the number of negations the term placeholder  $\alpha$  is within the scope of. If it is odd, the polarity of its proximal logically atomic frame is reversed by the whole context; if even, that original polarity is retained.<sup>56</sup> Now it seems that it is possible to project substitution-inferential properties for logical compounds on the basis of an asymmetric relation of domination between expressions that are substituted for, by projecting the polarities of those compounds and applying the generalization that is appropriate for the polarity of the compound in each case. Although there is no *single* generalization specifying the significance of a certain asymmetric SMSIC of the sort demanded, there will be a *pair* of them, of just the sort demanded, one for each of the two polarities.<sup>57</sup>

If the suggestion that asymmetrically significant substitutional commitments governing the inferential significance of the occurrence of expressions that are substituted for can be accommodated in this way could be made to work, it would be devastating for the overall argument that has been offered here. It will not work, however, and something further can be learned from seeing why not. The general response is straightforward. Such a procedure will work as offered only if all predicates are one-place, including logically compound ones. If not, the polarity of a predicate can be different in different argument places. Thus  $Pa \rightarrow Pb$  will have opposite polarities for the two argument places—one in the antecedent and the other in the consequent. The two-place predicate  $P\alpha \rightarrow P\beta$  will not be sorted, then, into either polarity class by the procedure outlined above, and its inferential properties will accordingly not be determined by either pattern corresponding to the domination relation between  $a$  and  $b$ .

One might try to overcome this difficulty by assigning polarities not to *predicates* or *sentence frames* but to *argument places*. Then, assuming that the underlying frame  $P\alpha$  has positive polarity, in  $P\alpha \rightarrow P\beta$ , the  $\alpha$  position will have negative polarity, and the  $\beta$  position will have positive polarity. But this proposal will still not determine what should be said about the inferential relations between  $Pa \rightarrow Pa$  and  $Pb \rightarrow Pb$  in the case where  $a \triangleright b$ : the first argument place has negative polarity and is being weakened by the substitution of  $b$  for  $a$ , so on that basis it should be the case that the overall claim is being strengthened inferentially by the substitution, while the second argument place has positive polarity and is being weakened by the substitution of  $b$  for  $a$ , so on that basis it should be the case that the overall claim is being weakened inferentially by the substitution.

Even if it is possible to fix up the proposal so as to deal with this difficulty, there is another that is decisive. One can take a sentence with two terms occurring in it at argument places of different polarities and form from it a *one-place* predicate or sentence frame:  $P\alpha \rightarrow Q\alpha$ , which can be represented as  $R\alpha$ . In this sentence frame,  $\alpha$  has both positive and negative polarity. This is fatal to the scheme suggested for keeping track of polarities to permit projection of substitution-inferential proprieties in the face of asymmetrically significant substituted-for expressions.

Why not, then, exclude from the range of projection predicates that would require being assigned to each or neither of the polarities? Because these predicates are expressively essential. They are the ones that are required to codify the inferences involving predicates, for instance in the way that will eventually be made explicit by the use of quantifiers to bind conditional predicates ("Whatever walks, moves"). Thus the proposal cannot be carried through and poses no threat to the overall argument presented here.

## VII. CONCLUSION

The title of this chapter asks a double question: What are singular terms, and why are there any? The strategy of the answer offered to the first query is to focus on substitution. The fundamental unit of language is the sentence, since it is by uttering freestanding sentences that speech acts are performed. Thus sentences are fundamental in the sense that it is coherent to interpret a community as using (its practices conferring content on) sentences but not subsentential expressions, while it is not coherent to interpret any community as using subsentential expressions but not sentences. But in fact there are good reasons why any community that uses sentences should also be expected to use subsentential expressions, indeed subsentential expressions of particular kinds.

The notion of substitution provides a route from the discrimination of the fundamental sentential expressions to the discrimination of essentially subsentential expressions. To carve up sentences substitutionally is to assimilate

them accordingly as occurrences of the same subsentential expressions are discerned in them. Such a decomposition is accomplished by a set of substitution transformations. The functional significance of discerning in a sentence an occurrence of one out of a set of expressions that can be substituted for is to treat the sentence as subject to a certain subclass of substitution transformations relating it to other, variant sentences. So the expressions that are substituting and substituted for can be used to index the transformations.<sup>58</sup> Two sentences are taken to exhibit the same substitutional sentence frame in case they are substitutional variants of one another—that is, are accessible one from the other by substitution transformations. These substitutional assimilations define two basic substitution-structural roles that essentially subsentential expression kinds could play. The first half of the answer to the first question, “What are singular terms?” is, then, that *syntactically*, singular terms play the substitution-structural role of being substituted for, while predicates play the substitution-structural role of sentence frames.

The second half of the answer to that question is that *semantically*, singular terms are distinguished by their *symmetric* substitution-inferential significance. Thus if a particular substitution transformation that corresponds to substituting one singular term for another preserves some semantically relevant sentential status (commitment, entitlement, truth, or whatever) when only primary occurrences are involved, no matter what the sentence frame, then the inverse transformation also preserves that status, regardless of frame. By contrast, every sentence frame is involved in weakening inferences where there is some other frame such that replacing primary occurrences of the first by the second always preserves the relevant sentential status, no matter what structure of substituted-for expressions is exhibited, while the converse replacement is not always status-preserving. Because the simple material substitution-inferential commitments that articulate the semantic content associated with singular terms are symmetric, their transitive closure partitions the set of singular terms into equivalence classes of intersubstitutable substituted-for expressions. It is in virtue of this defining character of their use that singular terms can be said to “purport to refer to just one object.”

The full answer to the question, What are singular terms? is then that singular terms are substitutionally discriminated, essentially subsentential expressions that play a dual role. Syntactically they play the substitution-structural role of being substituted for. Semantically their primary occurrences have a *symmetric* substitution-inferential significance. Predicates, in contrast, are syntactically substitution-structural *frames*, and semantically their primary occurrences have an *asymmetric* substitution-inferential significance. This precise substitutional answer to the first question supplies a definite sense to the second one.

To ask why there are singular terms is to ask why expressions that are

substituted for (and so of the basic substitution-structural kind) should have their significance governed by symmetric commitments, while sentence frames (expressions of the derivative substitution-structural kind) should have their significance governed in addition by asymmetric commitments. The strategy pursued in answer to this question is to focus on the use of logical vocabulary to permit the explicit *expression*, as the content of sentences, of relations among sentences that are partly constitutive of their being contentful. To say that subsentential expressions are used by a community as substituted-fors and substitution-structural frames is to say that the contents conferred by the practices of the community on the sentences in which those expressions have primary occurrence are related systematically to one another in such a way that they can be exhibited as the products of contents associated with the subsentential expressions, according to a standard substitutional structure. The problem of why there are singular terms arises because that structure need not, for all that has just been said, assume the specific form that defines singular terms and predicates.

But suppose the condition is added that the sentences whose proper use must be codifiable in terms of the proper use of their subsentential components is to include (or be capable of being extended so as to include) not only logically atomic sentences but also sentences formed using the fundamental sentential logical vocabulary, paradigmatically conditionals and negation. This condition turns out to interact in intricate ways with the possibility of substitutional codification of sentential contents by subsentential ones—ways that when followed out can be seen to require just the combination of syntactic and semantic substitutional roles characteristic of singular terms and predicates. So the answer offered is that the existence of singular terms (and so of their complementary predicates) is the result of a dual expressive necessity. On the one hand, the material-inferential and material-incompatibility commitments regarding sentences must be implicitly substitutionally codifiable in terms of material-inferential and material-incompatibility commitments regarding the subsentential expressions that can be discerned within them or into which they can be analyzed, if the contents of novel sentences are to be projectable. On the other hand, those same commitments regarding sentences must be explicitly logically codifiable as the contents of assertional commitments, if the contents of nonlogical (as well as logical) sentences are to be available for public inspection, debate, and attempts at improvement. It is these two expressive demands, each intelligible entirely in terms of considerations arising already at the sentential level, that jointly give rise to the structure of symmetrically significant substituted-fors and asymmetrically significant substitution-structural sentence frames that defines the functional roles of singular terms and predicates.

The argument presented here may be called an *expressive deduction* of the necessity of basic subsentential structure taking the form of terms and predicates. A language must be taken to have expressions functioning as

singular terms if essentially subsentential structure is (substitutionally) discerned in it at all, and the language is expressively rich enough to contain fundamental sentential logical locutions, paradigmatically conditionals (which permit the assertorily explicit expression of material-inferential relations among nonlogical sentences) and negations (which permit the assertorily explicit expression of material-incompatibility relations among nonlogical sentences). The only way to combine the presence of logical vocabulary with a subsentential substitutional structure that does not take the term/predicate form is to preclude the formation of semantically significant logically compound sentence frames, by denying substituted-for expressions primary occurrence in logically compound sentences. The expressive cost of this restriction, however, is also substantial.

Unless sentence frames formed by substitutional assimilation of logically compound sentences are already available, it is not possible to introduce logical vocabulary (in the case of singular terms and predicates, identity and quantificational locutions) that will do for the commitments articulating the contents of subsentential expressions what the conditional and negation do for the commitments articulating the contents of sentences—namely express them explicitly in the form of assertional commitments. To put the point otherwise, the expressive power of sentential logical vocabulary derives in part from the interaction between a direct and a substitutionally indirect mode of making explicit the commitments that articulate sentential contents. The direct mode permits the formulation as the content of assertional commitments—of inferential commitments, for example. Without the expressive capacities provided here by conditionals, reasons could be demanded and debated for premises and conclusions, but not for the material inferences whose correctnesses are part and parcel of the content of those premises and conclusions. The substitutionally indirect way of making explicit the commitments that articulate sentential contents is by making explicit the commitments that articulate the contents associated with the subsentential expressions into which they can be analyzed, which commitments regarding subsentential expressions implicitly codify the same commitments regarding sentences that can also be made explicit directly. If a language has sentential logical vocabulary suitable to play both sorts of explicating role, then its subsentential structure is obliged to take the specific form of singular terms and predicates.

Logical vocabulary has the expressive role of making *explicit*—in the form of logically compound, assertible sentential contents—the *implicit* material commitments in virtue of which logically atomic sentences have the contents that they do. Logic transforms semantic practices into principles. By providing the expressive tools permitting us to endorse in what we say what before we could endorse only in what we did, logic makes it possible for the development of the concepts by which we conceive our world and our plans (and so ourselves) to rise in part above the indistinct realm of mere tradition,

of evolution according to the results of the thoughtless jostling of the habitual and the fortuitous, and enter the comparatively well-lit discursive marketplace, where reasons are sought and proffered, and every endorsement is liable to being put on the scales and found wanting. The expressive deduction argues that subsentential structure takes the specific form of singular terms and predicates because only in that way can the full expressive benefits of substitutional subsentential analysis—codifying material correctnesses implicit in the use of sentences in material correctnesses implicit in the use of subsentential expressions—be combined with those afforded by the presence of full-blooded logical vocabulary of various sorts, performing its task of making explicit in claims what is implicit in the practical application of concepts.

In other words, languages have singular terms rather than some other kind of expression so that logic can help us talk and think in those languages about what we are doing, and why, when we talk and think in those languages. The full play of expressive power of even purely sentential logical vocabulary turns out to be incompatible with every sort of substitutional subsentential analysis except that in which essentially subsentential expressions playing the substitution-structural role of being substituted for have symmetric, substitution-inferential significances, and those playing the substitution-structural role of sentence frames have asymmetric, substitution-inferential significances. For to play its inference-explicitating role, the conditional, for instance, must form compound sentences whose antecedent substitution-position is inferentially inverting. Only symmetrically significant expressions can be substituted for, and so form sentence frames, in such a context. That is why in languages with conditionals, subsentential structure takes the form of singular terms and predicates.

In the opening paragraph of Section III it was pointed out that the principle that singular terms are used to talk about particular objects can be exploited according to two complementary directions of explanation. One might try to give an account of what particulars are, without using the concept *singular term*, and then proceed to define what it is to use an expression as a singular term by appeal to their relations to particulars. Or one might try to give an account of what singular terms are, without using the concept *particular*, and then proceed to define what it is for something to be a particular by appeal to their relations to expressions used as terms. (It should of course be admitted that in either case the talking about relation will require substantial explanation, though that explanation may have to look quite different depending on which explanatory strategy it is conceived as abetting.) The answer presented here to the question, What are singular terms? does not appeal to the concept of objects. So it provides just the sort of account required by the first stage of the second (Kant-Frege) strategy for explaining the concept of objects.

It is not the business of this chapter to pursue the later stages of that

direction of explanation, nor, therefore, to argue for its ultimate viability. But it is worth pointing out here that in the context of this order of explanation, to explain why there are singular terms is in an important sense to explain why there are objects—not why there is something (to talk about) rather than nothing (at all), but rather why what we talk about comes structured as propertied and related objects. “The limits of language (of that language which alone I understand) means the limits of my world.”<sup>59</sup> To ask the question, Why are there singular terms? is one way of asking the question, Why are there objects? How odd that the answer to both should turn out to be: because it is so important to have something that means what *conditionals* mean!

## Appendix I: From Substitutional Derivation of Categories to Functional Derivation of Categories

In functional-categorial grammars of the sort Lewis discusses in “General Semantics,” one starts with some basic categories and defines derived categories by functions whose arguments and values are drawn from the basic categories, and from the derived categories already defined. Thus where singular terms ( $T$ ) and sentences ( $S$ ) are the basic categories, predicates, ( $T \rightarrow S$ ), are defined syntactically as functions taking (ordered sets of) terms as arguments and yielding sentences as values. Semantically, they are interpreted by functions taking (ordered sets of) whatever sort of semantic interpretant is assigned to terms as arguments and yielding as values whatever sort of semantic interpretant is assigned to sentences. Items of the basic categories play the roles of arguments and values of the functions associated with items of the derived categories. Items of derived categories of course play the roles of functions, but they also can serve as arguments and values of other functions.

These three roles correspond to substitution-structural roles. To play the role of value of a function is to be an expression that is substituted in. To play the role of argument of a function is to be an expression that is substituted for. To play the functional-categorially derivative role of a function is to be a substitutional frame—that is, a substitutionally derivative role. The triadic division of substitution-structural roles is accordingly orthogonal to that of functional-categorially basic and derived categories, in that derived categories, for instance, play all three substitution-structural roles. The two sorts of structure are nevertheless intimately related. The substitutional syntactic structure is a way of thinking about and constructing the functional-categorial syntactic structure (and thereby the corresponding semantic one).

The substitutional construction corresponding to a functional-categorial

hierarchy standardly generated by using sentences and terms as basic categories begins with sentences and substitutional transformations relating them. The sentences act as expressions to be substituted in, corresponding to the values of predicate functions. As to the expressions that are substituted for, they correspond to the indices of the substitution transformations. If they are, like the sentences substituted in, antecedently distinguishable, then the facts concerning which transformations remove occurrences of which terms, and which produce occurrences of other terms, can be used to index the substitution transformations, assigning to each a set of pairs of substituted and substituting terms. If only the transformations and not the term occurrences are given, then the sentences substituted in can be indexed by the substitution transformations that apply nontrivially to them, in order to determine what terms occur in them. Each transformation is then assigned a pair of sets of sentences—those it applies nontrivially to, and those it nontrivially results in. (A nontrivial substitution transformation is one that results in some sentence different from that to which it is applied. Intuitively, a transformation will apply nontrivially to a sentence only if one or more of the expressions it substitutes for occurs in the sentence.) In either case, predicates as sentence frames are defined as equivalence classes of substitutionally variant sentences.

Depending upon how the substitution transformations are conceived, it may take some special effort to see to it that a proper equivalence relation is defined from these substitutional accessibility relations. Thus if substitution for  $t$  needs to replace all occurrences of  $t$ , then it need not be the case that wherever substituting  $t'$  for  $t$  in  $s$  yields  $s'$ , that substituting  $t$  for  $t'$  in  $s'$  yields  $s$ . This failure of immediate symmetry is evident if one substitutes 'Hegel' for 'Kant' in "Hegel wrote about Kant"—the converse substitution will not recover the original sentence from "Hegel wrote about Hegel." Similar phenomena afflict transitivity. These may be resolved either by defining substitutional accessibility in terms of the symmetric and transitive closures of these basic substitutional relations or by permitting partial substitutions in the base relation, at the cost of making each transformation one-many instead of one-one. For present purposes, it does not matter which route is adopted.

Defining sentence frames as equivalence classes of substituted-in expressions in this way suffices to determine their role as functions that apply to sets of substituted-in expressions. Applying the function to such an argument is just selecting some of the substitutionally variant sentences contained in the equivalence class, depending upon which substitution transformations apply nontrivially to it. This is another way of saying that selecting the right substitutionally variant instance depends on what substituted-for expressions occur in it. It may be noticed that at this stage, nothing corresponding to the order of arguments for a predicate function has been distinguished; no way has been supplied to tell the difference between "Brutus killed Caesar"

and "Caesar killed Brutus." These are recognizable as distinct members of the equivalence class that may be denominated " $\alpha$  killed  $\beta$ ." But each of them is a result of applying that frame to the unordered set {'Brutus', 'Caesar'}. This is patently insufficiently discriminating for the purposes of codifying inferences; this point holds even before quantifiers, for it concerns the antecedently important implicit proprieties of inference that will be made explicit as the contents of assertions with the aid of quantifiers. So (to stick to the simplest sort of substitution inference) "Brutus killed Caesar" follows from "Brutus killed Caesar" but not from "Caesar killed Brutus." Finer discrimination is thus required.

This requirement should come as no surprise, if for the moment one thinks about substitution transformations in functional-categorial terms, rather than the other way around. Basic substitution transformations are  $(T \rightarrow T)$ s, functions substituting one term for another, which induce  $(S \rightarrow S)$ s. But it is seen in the body of the chapter that codification of simple inferences by substitution requires consideration not only of the inferential significance of substitution for terms but also of the inferential significance of replacement of predicates. This operation corresponds categorially to a function from predicates to predicates (hence inducing one from sentences to sentences), namely to a  $((T \rightarrow S) \rightarrow (T \rightarrow S))$ .

Understanding such an operation requires understanding predicates,  $(T \rightarrow S)$ s, not in their role as functions or frames but in their role as arguments and values of higher-order functions. (It is here that the connection Dummett rightly perceives between complex predicates and quantification emerges, for quantifiers are of course  $((T \rightarrow S) \rightarrow S)$ s.) For, in addition to their role as functions, the full-fledged derivative categories  $(X \rightarrow Y)$  of an unrestricted functional-categorial grammar can also serve as arguments for further, higher-order derived categories, such as  $((X \rightarrow Y) \rightarrow Z)$ , and as values of such categories as  $(Z \rightarrow (X \rightarrow Y))$ . Talk of playing the role of argument and value is, in substitutional terms, talk of playing the substitutional roles of being substituted for and substituted in. The analog to being substituted for, for substitutionally and functionally derivative sorts of expressions, has been called here 'replacement'. Supposing that replacement can be defined, the role of sentence frames as values of functions—that is, as expressions that can themselves be substituted in (and therefore be understood as the result of broadly substitutional relations)—will follow straightforwardly. No new considerations are introduced by this further role, however, so it is not further considered here.

Defining replacement of one sentence frame by another is a more complex affair. This is the analog for substitutional frames of substituting one expression for another, which underlies the inference for instance from "Brutus killed Caesar" to "Brutus injured Caesar." (This is a propriety of practice that, in idioms expressively rich enough to count as logically articulated, can be made explicit in the principle  $(x)(y)[(x \text{ killed } y) \rightarrow (x \text{ injured } y)]$ .) It is with

respect to this operation that sentence frames must be individuated as finely as complex predicates, and not just as simple ones. Replacing " $\alpha$  killed  $\beta$ " by " $\alpha$  injured  $\beta$ " requires keeping the argument places straight. At this point structure is required that has no analog whatsoever at the level of simple substitution for basic expressions.

Understanding substitution for basic expressions requires that sentences be assimilated into equivalence classes corresponding to frames. Replacement of one substitutionally derivative frame by another requires not only those equivalence classes but a mapping from one to another that has special properties. In particular, there must be a bijection mapping the two equivalence classes onto each other so as to preserve the substitutional relations within each class. With respect to such a mapping, replacement of one predicate by another in a sentence exhibiting it then results in the element of the replacing equivalence class that is the image under that mapping of the first sentence. An example will make clear what is intended.

The set of sentences corresponding to " $\alpha$  killed  $\beta$ "—call it  $S$ —has the form {"Brutus killed Caesar," "Brutus killed Brutus," "Caesar killed Brutus," "Caesar killed Caesar," "The noblest Roman killed the conqueror of Gaul," . . .}. The set of sentences corresponding to " $\alpha$  injured  $\beta$ "—call it  $S'$ —has the form {"Brutus injured Caesar," "Brutus injured Brutus," "Caesar injured Brutus," "Caesar injured Caesar," "The noblest Roman injured the conqueror of Gaul," . . .}. The trouble is that these are unordered sets. Since at lower levels the occurrences of terms have been distinguished, it is already possible to specify that the result of replacing " $\alpha$  killed  $\beta$ " by " $\alpha$  injured  $\beta$ " in "Brutus killed Caesar" must be an element of  $S'$  in which the terms 'Brutus' and 'Caesar' both occur. So "Caesar injured Caesar" and "The noblest Roman injured the conqueror of Gaul" are ruled out. But nothing said so far makes it possible to choose between "Brutus injured Caesar" and "Caesar injured Brutus."

What is required is that the set  $S$  of sentences corresponding to " $\alpha$  killed  $\beta$ " be put in one-to-one correspondence with the set  $S'$  of sentences corresponding to " $\alpha$  injured  $\beta$ ," so that  $h(\text{"Brutus killed Caesar"}) = \text{"Brutus injured Caesar,"}$   $h(\text{"Caesar killed Brutus"}) = \text{"Caesar injured Brutus,"}$  and so on. Then to replace " $\alpha$  killed  $\beta$ " by " $\alpha$  injured  $\beta$ " in "Brutus killed Caesar," one simply applies the function  $h$ . The formal criterion of adequacy for a function  $h$  to be able to play this role is that:

*If  $h(s) = s_1$  and if  $\text{Sub}\{s, s_2, t, t'\}$   
(that is,  $s_2$  results from  $s$  by substituting  $t'$  for  $t$ ),  
then there must exist an  $s_3$  such that  $\text{Sub}\{s_1, s_3, t, t'\}$  and  
 $h(s_2) = s_3$ .*

In the example, since  $\text{Sub}\{ \text{"Brutus killed Caesar,"} \text{ "The noblest Roman killed Caesar,"} \text{ "Brutus,"} \text{ "The noblest Roman"} \}$ , this means that if  $h(\text{"Brutus killed Caesar"})$  is "Brutus injured Caesar," then there must be some sentence,

namely "The noblest Roman injured Caesar," such that it is the case both that  $h$ ("The noblest Roman killed Caesar") is "The noblest Roman injured Caesar," and that  $\text{Sub}$  ("Brutus injured Caesar," "The noblest Roman injured Caesar," "Brutus," "The noblest Roman"). The notion of frame replacement makes sense only where such a mapping  $h$  has been defined from substitutional variants that are elements of one substitution-frame equivalence class to those that are elements of another.

Of course it is clear from this example that if there is one such mapping, there may well be others. For instance,  $h'$  could satisfy the condition if  $h'("Brutus killed Caesar") = "Caesar injured Brutus"$  instead of "Brutus injured Caesar." Selecting a substitution-structure preserving isomorphism  $h$  suffices to define the operation of predicate replacement that is employed in the semantic discussion of substitution inferences in the broad sense, which involves not only substituting for basic expressions but replacing substitutionally derived ones. This is all that is appealed to in the argument of this work.

To define the full functional-categorial hierarchy of derived categories, however, not only must frames be replaceable, but sentence-frame frames must be definable from them. This is part of playing the role of argument for higher-level functions. If the notion of predicate replacement is to be extended so as to be fully analogous to substitution for basic expressions (as the argument does not require), further structure still is needed. In particular, for this syntactic operation, one must be able to assimilate substituted-in expressions (sentences) accordingly as the same sentence-frame frame is exhibited—what "Kant admired Rousseau, and Kant wrote about Rousseau" has in common with "Kant lived longer than Rousseau, and Kant had a shorter name than Rousseau" and "Kant wrote more than Rousseau, and Kant wrote more carefully than Rousseau." Defining equivalence classes of sentences accessible from one another by replacing predicates with predicates requires more than the pairwise isomorphisms required to define replacement of predicates in the first place. It requires a set of such isomorphisms that link all the interreplaceable predicates into an equivalence class. This can be formulated as a requirement on a set of pairwise replacement-defining substitution-preserving isomorphisms. A structure  $\langle R, H \rangle$  is a replaceability equivalence structure, in case:

1.  $R = \{P_i / \text{each } P_i \text{ is a predicate of the same adicity } n\}$  and
2.  $H = \{h[P_i, P_j] / P_i, P_j \text{ are elements of } R, \text{ and } h[P_i, P_j] \text{ is a substitution-preserving isomorphism between } P_i \text{ and } P_j\}$ , and
3.  $H$  is reflexive, symmetric, and transitive over  $R$ , in that:
  - (a)  $h[P_i, P_i]$ , an identity relation, is an element of  $H$
  - (b)  $h[P_i, P_j]$  is the inverse of  $h[P_j, P_i]$
  - (c)  $h[P_i, P_k]$  is the composition of  $h[P_i, P_j]$  and  $h[P_j, P_k]$ .

Conditions (a), (b), and (c) need to be specially stipulated because of the potential multiplicity of isomorphisms mapping  $P_i$  onto  $P_j$ , and so qualifying

to be included in  $H$  as  $h(P_i, P_j)$ . This means that specifying such a structure amounts to picking one complex predicate from the set of those associated with each given simple predicate. Each structure  $\langle R, H \rangle$  permits the definition of predicates as objects in the full-blooded sense of substituted-for expressions. Thus invariants of substituted-ins, across replacement within these classes  $R$ , as defined by the associated set of mappings  $H$ , permit the definition of genuine derived categories of higher-order sentence frames resulting from replacement of predicates. These same constructions of frames, by assimilation of substituted-ins, and of substituted-fors out of derived frames of lower levels, will be repeated at each level to generate the full hierarchy of functional categories.

## Appendix II: Sentence Use Conferring the Status of Singular Terms on Subsentential Expressions—An Application

In Section IV of this chapter, an account is offered of what it is to use expressions as singular terms and predicates. That account is couched in terms of substitution-inferential relations among sentences. One consequence of the argument is accordingly that a theorist who analyzes some target system of linguistic practices by discerning the use of expressions as singular terms and predicates is obliged to show how that analysis is supported by appropriate features of the use of the sentences that contain them. The substitution-inferential structure described here puts substantial constraints on sentential practices, which must be satisfied if they are to be claimed to be sufficient to confer on subsentential components the semantic significance of singular terms and predicates. Where these constraints are not observed, erroneous conclusions will be drawn.

A prominent instance is what is often made of Quine's famous 'gavagai' example, from Chapter 2 of *Word and Object*. The example is forwarded as an argument for the thesis of the indeterminacy of translation. The significance of the example is typically understood to lie in its promise of a general recipe for generating alternate translation schemes by reindividuation. Specifically, wherever there is a 'straight' translation scheme, rendering a target-language sortal 'gavagai' by home-language sortal  $K$ , for instance 'rabbit', it is possible to produce a distinct and competing scheme that renders it instead by something that individuates more finely (or less, but it will suffice here to concentrate on the finer discriminations), for instance 'undetached rabbit part' or (temporal) 'rabbit stage'. The point is to be that since sentences are the smallest linguistic units that can be used to make a move in the language game, the evidence of linguistic practice directly constrains only the interpretation of sentences. This leaves considerable slack in how responsibility for the use of those sentences is indirectly apportioned between the subsentential linguistic units the theorist chooses to discern. The

considerations advanced in the body of this chapter do not provide reason to quarrel with the general conclusion but do give reason to quarrel with this example.

The idea is that what get construed in the straight translation scheme as predication addressed to singular terms governed by the sortal 'rabbit' are construed by the derived translation scheme as predication addressed to singular terms governed by the sortal 'undetached rabbit part'. Thus "There is a large rabbit" becomes something like "There is an undetached rabbit part of a large contiguous collection of such parts." As Quine indicates, "This is the same rabbit as that one" becomes "This undetached rabbit part belongs to the same contiguous collection of such parts as that one." The strategy is to take what appear to be sentences about rabbits, which predicate ordinary properties of them, as instead sentences about rabbit parts, which predicate of them gerrymandered properties involving the contiguous wholes they belong to.

From the point of view of the present analysis, the difficulty with such a derived scheme is that if the sentences as it construes them are to count as genuinely using some expressions as singular terms invoking parts, there must be some predication of them that do not address them solely through the wholes in which they appear. Not all the predicates that appear in the derived translated language can be of the sort that result from the recipe for retranslating what appear as predicates on the straight translation. That is, the use of sentences as translated must be governed by some symmetric simple material substitution-inferential commitments—which license substitutions of one part term for another—while insisting on a finer discrimination than that of their belonging to the same contiguous whole. These will be the symmetric SMSICs that could be (though they need not be) made explicit in the derivative translation by the use of genuine identity locutions such as "This is the same undetached rabbit part as that one." These will govern substitution inferences involving genuine predicates of undetached rabbit parts. Thus if the predicate  $P$  meant "... is a broken foot," a symmetric SMSIC governing terms  $a$  and  $b$  will license indifferently the inference from  $Pa$  to  $Pb$  and vice versa. It will be a commitment to the identity of the undetached rabbit parts  $a$  and  $b$ . It is one of the fundamental commitments of the present analysis that unless their use is such as properly to be governed by such symmetric SMSICs,  $a$  and  $b$  are not genuine singular terms. The point then is that derived translations of what are construed in the straight translation as predication applying to rabbits will not serve as contexts permitting genuine identity commitments regarding undetached parts.

This point can be seen intuitively, without appeal to the technical notion of simple material substitution inferential commitments. If 'gavagai' is to be construed as a genuine sortal, the language containing it must contain the apparatus needed to individuate the items it sorts. It must have a use for some notion that appears in the language as translated by the derived scheme

as 'same gavagai'. But the reindividuative strategy of construing apparent references to wholes as references to parts offers no assurance that the language being translated can be taken to have the apparatus needed to distinguish parts. Consider the suggestion that 'gavagai' means 'undetached organic molecule contained in a rabbit'. The natives presumably cannot identify and individuate molecules, and no amount of gerrymandering of their actual linguistic practice could construe it as already containing sufficient apparatus to do so.

The example of a reindividuative derivation of an alternative to the straight translation seems to work only because the theorist, working in a metalanguage rich enough to contain the full individuative and referential apparatus needed to make some expressions mean 'undetached rabbit part', or even 'undetached organic molecule contained in a rabbit', *stipulates* that a native expression is to be understood as used in the way such expressions are used. What the present considerations show is that this possibility does not ensure that the uses of the sentences attributed to the natives are themselves sufficient to *confer* that significance on the subsentential expressions they employ.<sup>60</sup> The result is a substantial asymmetry between the languages attributed to the natives on the straight construal and the derived construal. The straight construal attributes an autonomous language, in the sense that the use of the sentences attributed to the natives suffices by itself to make the subsentential expressions mean what they are taken by the theorist to mean. By contrast, the derived construal attributes a language that is not autonomous, in the sense that using the sentences in the way the natives are taken to is not enough to make the subsentential expressions mean what they are taken by the theorist to mean. Since no natural language could be like this—only an artificial language whose use is stipulated in some richer metalanguage could be—the straight construal is clearly theoretically preferable.

Thus the considerations advanced here concerning what it is for sentences to be related by substitution inferences in such a way that they count as containing occurrences of singular terms and corresponding predicates puts constraints on the theorist's discrimination of subsentential structure generating the use of sentences. These constraints are not satisfied by the proposed retranslations by reindividuation that would render what can be understood as 'rabbit' by 'undetached rabbit part' or 'rabbit-stage' (and dual considerations will apply to schemes that would move up to the less finely individuated 'rabbit-hood', rather than down to more finely individuated sortals).

To say this is to take issue with one (prominent) argumentative strategy, not with the indeterminacy thesis as such. For one thing, the present account begins with proprieties, including inferential proprieties, of the use of sentences, not with Quine's spare foundation of patterns of irritation of sensory surfaces (of theorist and native). Again, there are many other ways into indeterminacy not addressed by these conferral considerations, most notably

those Davidson develops involving the possibility of trading off attributions of beliefs and desires attributed to individuals, and the meanings attributed to their utterances. However, the fundamental point of this chapter has been to disagree with Quine's claim (offered at *Word and Object*, p. 53, as a lesson of the 'gavagai' example) that "terms and reference are local to our conceptual scheme," that "the very notion of term" is "provincial to our culture."

## Anaphora: The Structure of Token Repeatables

---

Untwisting all the chains that tie  
The hidden soul of harmony.

MILTON, L'Allegro

As fast as thou shalt wane, so fast thou grow'st  
In one of thine from that which thou departest,  
And that fresh blood which youngly thou bestow'st  
Thou mayst call thine when thou from youth convertest  
....

Let those whom nature hath not made for store,  
Harsh, featureless and rude, barrenly perish

....  
She carved thee for her seal, and meant thereby  
Thou shouldst print more, nor let that copy die.

SHAKESPEARE, Sonnet 11

---

### I. FREGE'S GRUNDLAGEN ACCOUNT OF PICKING OUT OBJECTS

#### 1. *Introduction*

The first step in understanding why and in what sense claims represent or are about objects is to see what sort of conceptual content can be associated with the use of singular terms—the expressions that purport to refer to or represent objects—and correlatively with the predicates that denote the properties of objects and the relations they stand in. Only what has propositional (assertible or believable) content can serve as premise and conclusion—can both be given as a reason and have reasons given for it—and hence play a directly inferential role of the primary sort. But all sorts of conceptual content are essentially inferentially articulated. So the conceptual contents of singular terms and predicates must be understood in terms of their indirectly inferential role—the contribution their occurrence makes to the inferential potential of sentences containing them. As Frege puts it: “We ought always to keep before our eyes a complete proposition. Only in a proposition have the words really a meaning . . . It is enough if the proposition taken as a whole has a sense; it is this that confers [*erhalten*] on its parts also their content.”<sup>1</sup>

One of the principal technical conceptual debts the inferentialist semantic

tradition owes to Frege is the idea of using *substitution* to understand how the directly inferential articulation of what is expressed by sentences induces an indirectly inferential articulation of what is expressed by their semantically significant parts. The previous chapter investigated the substitutional species of inferential relations and showed how to understand singular terms and predicates in terms of the roles that expressions of those categories play in substitution inferences. It showed further the sort of independently characterizable expressive impoverishment to which a discursive structure is doomed unless its subsentential substitution-inferential structure takes the specific form of singular terms and predicates. Subsentential structure may be eschewed entirely, though the cost is substantial. For one then forgoes the expressive empowerment provided by the combinatory generation of novel interpretable sentences from familiar sentence-parts, which looms so large in our actual discursive practice. If semantically significant, essentially subsentential structure is discerned substitutionally, however, it can take a form other than that of singular terms and predicates only by relinquishing the full semantic explicating expressive resources otherwise provided by sentential logical locutions, paradigmatically the conditional. This, it was claimed, is why there are singular terms (and so predicates, since the two categories come as a package). This same argument provides the ultimate reason why sententially atomic propositionally contentful claims are, or purport to be, about objects, and to represent those objects as having properties and standing in relations. The connection between singular-term usage and purported representation of objects can be filled in a bit by looking at how Frege, in the *Grundlagen*, understands the representation of objects in purely substitutional terms.<sup>2</sup>

## 2. Objects Are Given to Us by the Use of Singular Terms

One of Frege's concerns in the *Grundlagen* is to explain "how numbers are given to us."<sup>3</sup> In order to do that, he must consider the wider question of how particular objects are "given to us" cognitively. On the face of it, explaining what it is for our thought and talk to pick out or be directed at objects seems particularly difficult for the case of numbers, since, as he puts it, "we cannot have any ideas or intuitions of them."<sup>4</sup> Translated from the neo-Kantian idiom he is employing here, this means that the aboutness of numerical thought can be understood neither as derived from the supposedly more primitive aboutness of subjective mental pictures nor as a feature of the way in which thought about numbers is causally influenced by the numbers it is about. In fact in this context the abstractness of number is a philosophical boon because it requires Frege to address in its most general terms the question of what it is to pick out objects with our concepts—undistracted by such ultimately misleading features of some prominent special cases as the presence of mental pictures of or causal commerce with what is

thought and talked about. Instead, the role played by causal contact in the ability to pick out perceivable objects in thought and talk must be understood in terms of some more general conception of object-directedness.

Frege calls the grammatical category of expressions used to talk and think about objects "proper names": "The name of a thing is a proper name [*Eigenname*]."<sup>5</sup> This usage elides the distinction—of the first importance for Frege's project in the *Grundlagen*—between lexically simple singular terms, such as 'Frege', and definite descriptions formed from predicates and sortals, such as 'the author of the *Grundlagen*'. Frege's discussion focuses on the latter for two reasons: numerical expressions are formed in this way, and the definite article makes *explicit* the singular referential purport that is implicit in the use of other singular terms.

We speak of "the number 1," where the definite article serves to class it as an object.<sup>6</sup>

The definite article purports to refer to a definite object.<sup>7</sup>

The general question Frege is addressing is how expressions must be used for them to *succeed* as singular terms by referring to, picking out, or giving us a cognitive grip on definite objects—as "ways in which objects are given to us," ways of "arriving at determinate" objects, or "symbols signifying objects."<sup>8</sup> The issue of what it is to use an expression as a name of an object is ultimately a normative one; it is to be responded to by specifying proprieties of practice. Since the use of the definite article makes singular referential purport explicit, those proprieties can be brought out into the open by asking (in deontic scorekeeping terms) what sort of *commitment* is expressed by the use of the definite article, and what is required for *entitlement* to that commitment.

Frege insists that the issue of entitlement to singular referential purport is an important one. The use of the definite article stands in need of *justification* [*Rechtfertigung*.]<sup>9</sup> The definite article is used in forming definite descriptions from predicates—what he calls "the definition of an object in terms of a concept under which it falls."<sup>10</sup> Frege is explicit about what is required for a justification of such a use of a definite article:

If, however, we wished to use [a] concept for defining an object falling under it, it would, of course, be necessary first to show two distinct things:

1. that some object falls under this concept;
2. that only one object falls under it.<sup>11</sup>

These are the paired conditions, of existence and uniqueness, on which Russell later erected his theory of descriptions.

It may seem, that however it is with definite descriptions, explaining the object-directedness of thought (the way it puts us in touch with particular

objects our judgments are about) need involve attending only to the first of these. Showing what it means for atomic judgments to be analyzable in terms of the application of predicates would seem to suffice, for objects should then emerge as what the predicates are judged to apply *to* or be true *of*. In fact, however, understanding this ‘*to*’ and this ‘*of*’ requires mastery of the sort of practical issues of identity and individuation that are appealed to in the second condition. For in the absence of such considerations, one grasps only the use or application of whole sentences (what it is to take them to be true)—not yet what it is to apply them *to* something or take them to be true *of* something. That a judgment is directed toward an *object* is intelligible only in the context of practices of identifying objects as the same again, and individuating them as distinct.

### *3. Judgments Expressing Our Recognition of an Object as the Same Again Are Substitution Licenses*

This is to say that the use of expressions as singular terms essentially involves, not only norms that could be made explicit as criteria of *application*, but also norms that could be made explicit as criteria of *identity*. Frege formulates this categorial point as the demand that “if we are to use a symbol *a* to signify [*bezeichnen*] an object, we must have a criterion for deciding in all cases whether *b* is the same as *a*, even if it is not always in our power to apply this criterion.”<sup>12</sup> For what an expression makes cognitively available for us to “have a definite character” as an object our judgments are about, it is necessary that “it can be recognized again beyond doubt as the same, and can be distinguished from every other.”<sup>13</sup> As indicated by the qualification “even if it is not always in our power to apply this criterion” in the previous passage, in spite of the epistemic flavor of “recognition” and “beyond doubt,” the requirement is not that we in fact be able to apply the implicit criterion of identity or be infallible in our recognitions. It is just that a notion of *correctness* of identifications and discriminations must have been settled somehow. The normative status must have been instituted, even though any particular attitudes, attributions, and assessments may get it wrong.

The demand for an implicit criterion of identity associated with the use of a singular term is presupposed by the uniqueness condition on the application of definite descriptions, but it is not restricted to those singular terms in which the singular referential purport is marked overtly by the use of a definite article. In its absence, no sense could be made of the notion that terms (including those that are not definite descriptions) implicitly involve a specifically *singular* referential purport. An implicit criterion of identity provides the “authority to pick out [particulars] as self-subsistent objects that can be recognized as the same again [*selbständige, wiedererkennbare Gegenstände zu unterscheiden*].”<sup>14</sup> What does it mean for such authority or enti-

tlement (which could be made explicit in the form of a criterion of identity) to be in place? The key fact is that “objects too can change their properties without that preventing us from recognizing them as the same [*sie als dieselben anzuerkennen*]”.<sup>15</sup> Recognizing an object as the same again is making a certain kind of judgment, what Frege calls a “recognition judgment.” Thus “For every object there is one type of proposition which must have a sense, namely the recognition-statement.”<sup>16</sup>

Indeed, the use Frege makes of the concept of a recognition judgment shows that he is committed to a much stronger claim. Not only is fixing the sense of recognition judgments *necessary* for entitlement to use an expression as singular term, it is *sufficient*. And once an expression has qualified as entitled to its singular referential purport, it is a way in which a determinate object can be picked out or given to us.<sup>17</sup>

How, then, are numbers to be given to us, if we cannot have any ideas or intuitions of them? Since it is only in the context of a proposition that words have any meaning, our problem becomes this: To define the sense of a proposition in which a number word appears. That, obviously, leaves us still a very wide choice. But we have already settled that number words are to be understood as standing for self-subsistent objects. And that is enough to give us a class of propositions which must have a sense, namely those which express our recognition of a number as the same again . . .

In doing this, we shall be giving a general criterion for the identity of numbers. When we have thus acquired a means of arriving at a determinate number and of recognizing it again as the same, we can assign it a number word as its proper name.<sup>18</sup>

That an expression is used as a singular term, and so has singular referential purport—that it is a way in which determinate objects can be made available to judgment (“arrived at,” “given to us”—is a significance that performances can be accorded in the context of practices of keeping deontic score on special sorts of commitment and entitlement. It emerges from the passages quoted above that the central technical concept Frege employs to explain the commitments and entitlements that define singular term usage is that of *fixing the sense of a recognition claim*. The rest of this section is devoted to exploring how Frege uses this concept to elaborate his understanding of what it is to talk or think about particular objects.

Securing singular reference is for Frege “a matter of fixing the content of a recognition-judgment [*Wiedererkennungsurtheils*]”.<sup>19</sup> Recognition judgments have the form of *identity* claims. Identity claims express recognition of an object as “the same again” when given or referred to in two different ways. To establish reference to a particular object by a given expression, one must settle what would make true or false various identities in which that expression occurs (even if one is not in a position to tell of each such identity

whether it is in fact true or false). Since the singular referential purport of terms amounts to claiming that recognition judgments involving those terms have a definite sense,<sup>20</sup> taking it that the significance of asserting an identity involving a term has been settled is treating the term as referring to or picking out an object. That is why “to use the symbol ‘=’ is likewise to designate [something] an object.”<sup>21</sup>

Frege’s problem was set by the fact that the absence of causal contact with and mental images of numbers made the possibility of picking them out as objects of thought and knowledge seem particularly mysterious. Reconceiving the problem of securing singular reference in terms of recognition judgments yields the result that “to obtain the concept of Number, one must fix the sense of a numerical identity.”<sup>22</sup> The general account of what it is to talk and think about particular objects accordingly shows how our cognitive and conceptual grasp on numbers can be made intelligible in terms of our capacity to take or treat sentences involving numerical terms as expressing identity claims.

Our aim is to construct the content of a judgment [*den Inhalt eines Urtheils zu bilden*] which can be taken as [*auffassen lässt*] an identity such that each side of it is a number.<sup>23</sup>

In the same way with the definitions of fractions, complex numbers and the rest, everything will in the end come down to the search for a judgment-content [*beurtheilbaren Inhalt*] which can be transformed [*verwandelt*] into an identity whose sides precisely are the new numbers. In other words, what we must do is fix the sense of a recognition-judgment for the case of these numbers.<sup>24</sup>

So Frege’s explanatory strategy begins with the idea that particular objects are to be distinguished as what can be recognized as the same again—in the sense that the norms governing the use of terms referring to them would be made explicit by associating with them not only criteria of application but also criteria of identity. The recognition judgments that express the applicability of such norms are thus to be construed as identity claims. To carry this strategy through to completion, Frege must address two further issues, one quite general and the other specific to the case of numbers (as abstract objects). The general question is what it is to “fix the sense of an identity”: How must a sentence be used, what sort of significance must it be accorded, in order to confer the content of an identity claim? The question specific to numbers is then what is required to confer such content on claims involving numerical expressions.

Frege’s answer to the first question is straightforward, and just as it should be from the point of view of the discussion of using expressions as singular terms in Chapter 6. Identity claims make explicit *substitution* licenses. “In universal substitutability [*allgemeinen Ersetzbarkeit*] all the laws of identity

are contained.”<sup>25</sup> Since identity claims are the form of recognition judgments, recognizing an object as the same again is itself to be understood in terms of substitutional commitments. “When are we entitled to regard a content as that of a recognition-judgment? For this a certain condition has to be satisfied, namely that it must be possible in every judgment to substitute without loss of truth the right-hand side of our putative identity for its left-hand side.”<sup>26</sup> The consequences of application distinctive of identity claims consist in the undertaking of substitution-inferential commitments. What is made assertorionally explicit as a claim of the form  $a = b$  is commitment to a pattern of inferences requiring doxastic (assertional) commitment to the claim expressed by  $Pa$  whenever one undertakes doxastic commitment to the claim expressed by  $Pb$ , and vice versa. Frege understands particular objects as what we get cognitive and conceptual access to by using expressions as singular terms, and he offers a substitutional construal of what it is to use expressions as singular terms. The proprieties governing the circumstances and consequences of their application are those codified explicitly in identity claims, which have the significance of symmetrical substitution licenses.

#### *4. The Maximal Substitutional Requirement on Using an Expression as a Singular Term*

There is more to introducing a new term by “fixing the sense of an identity” involving it, however, than just understanding what one is committing oneself to in asserting such an identity. Ordinary cases of term introduction are special in a way that tends to obscure what more (beyond a general understanding of identity) is required to fix the sense of identities in which the new expression occurs. The sort of example that best highlights what else Frege takes to be needed is that of introducing not only new terms but new objects. The lesson appears most clearly from consideration of the role played by *abstraction* in his account of how we can become entitled to use numerals as names of definite objects.

The key point is that to be entitled to introduce a new term as the name of an object, one must settle when it would be *correct* to recognize the object picked out as the same again; in this way one distinguishes it from all other objects.<sup>27</sup> Frege officially insists that to do this one must see to it that the truth or falsity of *all* identities involving it has been settled. Doing so is settling when it would be correct to recognize the object picked out as the same again, and thus distinguishing it from all other objects. When what settles the truth-values of these identity claims involving a term is made explicit, it takes the form of a criterion of identity.<sup>28</sup>

In run-of-the-mill cases of term introduction, this requirement is quite easy to satisfy. For in the central cases a new singular term is being introduced to refer to an object that can already be referred to by using other terms

already available in the language. When a proper name is introduced for a person, place, or perceivable thing, there are typically already-individuating sortals in place appropriate to it, and it can be picked out by definite descriptions using those sortals, combined with specifications distinguishing it from others of its kind (for instance, spatiotemporal ones): ‘the person who just came out of the front door of Jay’s Bookstall’, ‘the Northwest corner of the intersection of Forbes and Meyran avenues’, ‘the black telephone in that corner’, and so on. Assuming that the use of these antecedently available terms is already in order (as far as Frege’s official requirement is concerned), all that is required to introduce a new term  $a$  is commitment to a reference-fixing identity. For under these circumstances, if the term ‘(the  $x$ )[Dx]’ is already in use in the language, then by hypothesis the truth-values of all identities of the form (the  $x$ )[Dx] =  $t$  (where  $t$  is another term already in use in the language) have been settled. The introducing stipulation that  $a$  = (the  $x$ )[Dx] then automatically settles the truth-values of all the identities involving  $a$  and antecedent vocabulary:  $a$  =  $t$  just in case  $t$  = (the  $x$ )[Dx], and not otherwise. In these cases, then, committing oneself to the truth of a single-identity claim linking the novel term to a familiar one serves to fix the sense of *all* the identities involving the novel term, for it settles all their truth-values.

Clearly this technique is not available for introducing new terms for *new* objects—ones that cannot already be referred to in the antecedent vocabulary.<sup>29</sup> The problem of introducing numerical expressions referring to numbers, Frege says, is the problem of fixing the sense of numerical identities. He does this by the method of *abstraction*: a particular way of explaining the use of novel terms (referring to novel objects) by means of the use of familiar terms (referring to familiar objects). The idea is this: Where  $\lceil a \rceil$  and  $\lceil b \rceil$  are terms whose use is already established, new terms of the form  $\lceil fa \rceil$  and  $\lceil fb \rceil$  can be introduced wherever there is an equivalence relation  $R$  available defined on the old vocabulary. For one can then define the sense of identities involving  $f$  terms by stipulating that

$$\lceil fa \rceil = \lceil fb \rceil \text{ iff } Rab.$$

In this way, if  $\lceil a \rceil$  and  $\lceil b \rceil$  are terms designating lines, one can introduce new terms of the form  $\lceil \text{direction of } a \rceil$  and  $\lceil \text{direction of } b \rceil$  (and hence the new sortal or object-kind *directions*) by appeal to the equivalence relation . . . is parallel to \_\_ defined on lines:

$$\text{the direction of } a = \text{the direction of } b \text{ iff } a \text{ is parallel to } b.^{30}$$

In just the same way, if  $\lceil a \rceil$  and  $\lceil b \rceil$  are terms designating collections of already-available objects, one can introduce new terms of the form  $\lceil \text{number of } a \rceil$  and  $\lceil \text{number of } b \rceil$  (and hence the new sortal or object-kind *numbers*)

by appeal to the equivalence relation . . . can be put in one-to-one correspondence with   defined on collections of objects:

the number of *a* = the number of *b*  
iff *a* can be put in one-to-one correspondence with *b*.<sup>31</sup>

The claim that the relevant equivalence relation obtains between the familiar objects accordingly serves as the content of a judgment that can be taken as or transformed into an identity relating numerical (or direction) expressions, as Frege requires in the two passages quoted above.<sup>32</sup> That the judgment *Rab* can be reconstrued as an assertion of identity involving terms referring to novel abstract objects—rather than just as asserting a relation between familiar concrete (relative to this construction) ones—depends just on *R* being an equivalence relation; to be entitled to the reconstrual of such claims as putting us in cognitive and conceptual touch with abstract objects is just to be entitled to characterize *R* as reflexive, symmetric, and transitive. For since “in universal substitutability all the laws of identity are contained,” it follows that

in order to justify our proposed definition of the direction of a line, we should have to show that it is possible, if line *a* is parallel to line *b*, to substitute

‘the direction of *b*’

everywhere for

‘the direction of *a*’.

This task is made simpler by the fact that we are being taken to know of nothing that can be asserted about the direction of a line except the one thing, that it coincides with the direction of some other line. We should thus have to show only that substitution was possible in an identity of this type, or in judgment-contents containing such identities as constituent elements. The meaning of any other type of assertion about directions would have first of all to be defined, and in defining it we can make it a rule always to see that it must remain possible to substitute for the direction of any line the direction of any line parallel to it.<sup>33</sup>

Showing that the relation *R* on which the abstraction is based is an equivalence relation entitles one to regard *Rab* as an identity relating the new expressions *fa* and *fb* (circumstances of application). Regarding it that way is undertaking a substitutional commitment to the propriety of the inference from *P(fa)* to *P(fb)*, and vice versa, for any sentential context in which one discerns a primary occurrence of the new terms (consequences of application). The doctrine of abstraction Frege puts forward here is the claim that the significance of attributing this constellation of entitlement and commit-

ment is as taking the subject of those deontic statuses to be in a position to make judgments (to think and talk) about a new range of abstract objects—which may be thought of as equivalence classes of the old ones. This is how objects, paradigmatically mathematical ones, which we do not have causal commerce with (and need not be able to form mental images of) can be “given to us.”<sup>34</sup>

As indicated above, Frege’s official view is that to introduce a new term one must settle the truth-values of *all* identities relating it to other terms. This requirement leads to disastrous results in the later *Grundgesetze*, and Frege never does find an acceptable way to satisfy it for the introduction of terms referring to abstract objects.<sup>35</sup> In any case, the sort of abstractive definition just considered “fixes the sense of numerical identities” only in the sense of settling the truth-values of identities, both sides of which are numerical expressions—in the general case identities of the form  $fa = fb$ , but not of the form  $fa = c$ , where ‘ $c$ ’ is a bit of antecedent vocabulary, a term referring to an object that is concrete relative to the abstractive method of term-and-object introduction. The significance of the failure of abstractive definitions to meet the very strong condition Frege puts on term introduction—what one must do or show in order to be entitled to use an expression as a singular term—depends not only on whether it is possible to satisfy that condition in some other way but also on the reasons there are for endorsing that condition.

Frege’s basic insight is that the essential singular referential purport involved in singular-term usage consists in the role such terms play in identity claims. Since he further analyzes what is expressed by identity claims in terms of the significance of such claims as intersubstitution licenses, this amounts to taking singular referential purport to consist in a structure of symmetric substitutional commitments. It is in terms of the undertaking and attributing of such substitutional commitments that the scorekeeping significance of using a singular term to express a claim is to be understood. Frege takes it that the strong condition he imposes on successful term introduction is a consequence of this substitutional analysis of what it is to use an expression as a singular term. For he takes it that unless the truth-values of *all* identities involving the candidate term have been settled, it has not been settled what one would be committing oneself to by employing it to make claims (for the identities merely make substitutional commitments explicit, that is, assertible). Abstractive definitions settle whether in using one of the new terms to make a claim of the form  $P(fa)$  one is thereby committing oneself also to  $P(fb)$ , but they do not settle for arbitrary  $c$  whether one is committing oneself to  $P(c)$ .

Appealing to symmetric substitutional commitments (a species of inferential commitment) to explain what it is to use an expression as a singular term—the fundamental Fregean insight that is developed in detail in the previous chapter—does not necessitate the maximalist reading of what is

required for successful introduction of terms that Frege thinks follows from it. Frege thinks that there is something wrong with using an expression where it has been settled (whether or not anyone in particular is in a position to tell—a matter of status rather than of attitudes) that in endorsing a sentence in which it appears, one is thereby committed to the claims expressed by *some* substitutional variants of that sentence (and precluded from entitlement to commitments to the claims expressed by various other substitutional variants of that sentence), if there are some *other* substitutional variants on which one is *not* thereby counted as taking up a stance.<sup>36</sup> But what is wrong with its being settled that when I claim that the largest number that is not the sum of the squares of distinct primes is odd, I am thereby in some sense committing myself (whether I know it or not) to the claim that 17,163 is odd, am making a claim incompatible with the claim that 17,163 is even, and am not taking a stand on the question of whether England or the direction of the earth's axis is odd? Why would not such a situation count as one in which it had been settled exactly what I am and am not committing myself to (and similarly for entitlements), and so one in which a perfectly *definite* sense is associated with the numerical expressions involved, even though that sense is not *complete* in the way that Frege wants to require?

For many purposes it may be appropriate to insist on Frege's strong condition that the truth-values of all identities be settled; these may even include the purposes that motivate the development of the formal language Frege uses in the *Grundgesetze*. The issue is not even one of whether, relative to these purposes, a language in which this condition is imposed is *better* than one in which it is not. The question of interest at this point is rather whether there is some way of relaxing Frege's condition while maintaining the features of singular term use that make it appropriate to think of them as purportedly (and in favored cases successfully) picking out particular objects. Furthermore, it would be of interest to know just how weak the condition on the symmetric substitutional commitments associated with an expression could be made without endangering its singular referential purport. What is the *minimal* substitutional requirement (or necessary condition) on using an expression to pick out an object in thought, corresponding to the *maximal* substitutional requirement (or sufficient condition) that Frege imposes?

##### 5. *The Minimal Substitutional Requirement on Using an Expression as a Singular Term*

A good place to begin in addressing this question is to notice that even according to Frege, to fix the sense of a novel term (for instance a numerical expression) it is *not* in fact sufficient merely to settle the truth-values of all the identities it can occur in. For it would not suffice for term

introduction to settle the truth-value of all the nontrivial identities—all those that relate the term to some *other* term—as *false*. No criterion of identity is implicitly associated with the expression  $'a'$  by stipulating that  $a = a$ , but that if  $'b'$  is any expression distinct from  $'a'$ , then  $a = b$  is false. Settling the truth-values of all the identities involving the new expression in this way does not even implicitly involve associating with it an object that can be recognized as the same again. “All identities would then amount simply to this, that whatever is given to us in the same way is to be reckoned as the same. This, however, is a principle so obvious and sterile [*unfruchtbar*] as not to be worth stating. We could not, in fact, draw from it any conclusion which was not the same as one of our premises. Why is it after all that we are able to make use of identities with such significant results in such diverse fields? Surely it is rather because we are able to recognize something as the same again even though it is given in a different way.”<sup>37</sup>

Objects are essentially things that can be recognized as the same again, even though given in different ways. That is why they are things for which the issue of identity arises—why using the identity sign with an expression is treating it as referring to an object.<sup>38</sup> To be an object is to be something that can be referred to in different ways; to associate an object with an expression as its referent requires settling what would count as *another* way of picking out that same object. Frege’s maximalist claim is that introducing a term as picking out a definite object requires settling *every* other way of picking out that same object. The corresponding minimalist claim is that it requires settling at least *some* other way of picking out that same object.

The thought can be put more clearly by shifting from material mode to formal mode: from talk of objects to talk of the substitutional significance of singular terms by means of which talk of objects is officially to be understood. The basic idea is that unless the occurrence of a candidate term in the expression of a claim has *some* substitution-inferential significance (unless it commits one to *some* further claim that is expressed by a sentence resulting from the first by substitution of another term for the candidate), then the candidate is not functioning as a singular term at all. Its occurrence is not semantically significant in the way terms are; it is substitutionally idle, thus inferentially idle, and therefore semantically idle to discern its occurrence at all.<sup>39</sup> The minimal condition on using an expression as a singular term that emerges from understanding the characteristic substitutional role terms play is just that it must have been settled that the occurrence of the putative term have *some* (symmetric) substitution-inferential significance. As elaborated in the previous chapter, for the occurrence of an expression to have a significance of the kind characteristic of singular terms, its use must be governed by some simple material substitution-inferential commitments (SMSICs)—commitments of the sort that can be expressed explicitly as non-trivial identity claims or recognition statements. Where Frege demands a *complete* set of substitutional commitments associated with each term, the

minimal demand compatible with a substitutional understanding of singular terms (motivated by the observation that Frege would not permit all the nontrivial identities to be settled as false) is that a *nonempty* set of substitutional commitments be associated with each term; at least one nontrivial identity must be settled as true.<sup>40</sup>

Talk of objects as what can in principle be recognized as the same again—what can be given to us or referred to in different ways—reflects the structure of substitutional significance that the occurrence of bits of subsentential vocabulary must have for them properly to be understood as having the indirectly inferential content characteristic of the use of singular terms. The singular referential purport of such vocabulary consists in the fact that the deontic scorekeeping significance of its use is to be determined by symmetric substitutional commitments that link it to other vocabulary. These are the commitments that are made explicit by the nontrivial identity claims that Frege calls “recognition judgments”—which he takes to express the recognition of an object as the same again, though given in two different ways. This much of Frege’s thought in the *Grundlagen* can be taken over without a consequent commitment to the requirement that the truth-values of *all* nontrivial identities must be settled in order for a singular term to have been properly introduced.

Even the minimal claim that settling the truth of some nontrivial identities involving a candidate singular term is a necessary condition for using it as a name of an object, however, has consequences that can seem mysterious unless the substitutional gloss on that claim is kept firmly in mind. For it follows that the idea of an object that can be picked out or referred to only in one way is not an idea of an *object* at all. (Recall the discussion above in 6.4.) A language cannot refer to an object in one way unless it can refer to it in two different ways. This constraint will seem paradoxical if referring to an object by using a singular term is thoughtlessly assimilated to such activities as using a car to reach the airport or using an arrow to shoot a deer: even if only one car or one arrow is available and impossible to reuse, what one is doing can still genuinely be driving to the airport or shooting the deer. Why should referring be different, something that cannot be done one way unless it can be done two ways? Understanding an expression’s purporting to refer to an object in terms of its use being governed by proprieties articulating its significance according to substitution-inferential commitments dispels the puzzlement that can otherwise attend this phenomenon. An object that can be referred to in only one way is the sound of one hand clapping.

So for an expression to be used as a singular term, there must be *some* substantive substitutional commitment undertaken by the one who uses it. It is not necessary that either the one who undertakes that commitment or the one who attributes it—by attributing a doxastic commitment that would be avowed by the assertion of a sentence containing the singular term—be able to specify just what the content of that commitment is. But it is only

where the interpreter takes it that there *is* some such substitutional commitment included in the significance of the underlying doxastic commitment that the one who undertakes that commitment is interpreted as using a singular term to make a claim about an object. This is just the conclusion that was drawn in Chapter 6: the category of singular terms should be understood as comprising expressions whose proper use is governed by simple material substitution-inferential commitments (SMSICs) linking them to other such expressions. Taking an expression to be a singular term—taking it to purport to pick out a particular object—just is taking its use to be governed by some such SMSICs. When such a simple material substitution-inferential commitment linking two expressions is made propositionally explicit (as an assertible), it takes the form of a nontrivial identity claim. That is why to introduce an expression as a singular term, one must somehow settle the truth-value of at least one such identity (of what can be so expressed in an idiom with suitable explicating resources—that is, logical vocabulary). Purported reference to objects must be understood in terms of substitutional commitments linking diverse expressions.

### 6. Substitutional Triangulation

This *substitutional holism*—according to which mastery of the use of one expression as a singular term involves mastery of the use of many—is the reflection at the subsentential level of the *inferential holism* according to which mastery of the use of one expression as a sentence (even one that can be used to make noninferential reports) involves mastery of the use of many.<sup>41</sup> Carving up sentences according to their substitutional relations to one another is just a method for extending the notion of content-conferring, inferentially articulated deontic significance to the subsentential level—to expressions that cannot themselves play the directly inferential roles of premises or conclusion of inferences. The conceptual content expressed by a sentence depends on its place in a network of inferences relating it to other sentences; the conceptual content expressed by a singular term depends on its place in a network of substitutions relating it to other terms. The substitutional roles that determine the pragmatic significance of the occurrence of singular terms are a kind of indirectly inferential role because substitutional commitments are a kind of inferential commitment.

Another topic this minimal substitutional requirement for using an expression as a singular term illuminates concerns picking out objects by conceptual *triangulation*. Triangulation strategies arise from consideration of a fundamental problem concerning the discrimination of a particular *stimulus* to which some sort of response is reliably keyed. In his discussion, Davidson introduces the familiar point this way: "Why say the stimulus is the ringing of the bell? Why not the motion of the air close to the ears of the dog—or even the stimulation of its nerve endings? Certainly if the air were

made to vibrate in just the way the bell makes it vibrate it would make no difference to the behavior of the dog. And if the right nerve endings were activated in the right way, there still would be no difference."<sup>42</sup> Typically, there is a whole causal chain of covarying events culminating in a response. In the standard case, the occurrence of one is accompanied by the occurrence of all the rest. Under these circumstances, the response being keyed to one of the event kinds is its being keyed to all the rest. How is one element of the chain to be singled out as the stimulus? What is the nature and source of the privilege that distinguishes one element from another?

One strategy for assigning such privilege, and therefore picking out as the stimulus one element from the whole chain of covarying event types that culminates in a response of the specified type, is (as Davidson goes on to suggest) to look to *proximity* to the eventual response. The justification for seizing on causal proximity of stimulating event to the response as what matters is maximizing the relative *reliability* of the connection between the occurrence of events of the distinguished stimulus type and the occurrence of events of the distinguished response type. The proximal element of the chain is the one that most reliably brings about the response. For prior occurrences in the chain elicit the response only in the cases where they succeed in bringing about an event of the proximal type, while events of that type can elicit the response regardless of whether they have themselves been brought about in the standard way. The trouble is that such a proximal theory of stimuli will always yield the result that the stimuli being responded to are at the sensory surfaces or within the nervous system of the responding organism.

In the context of the project of using reliable differential responsive dispositions as a model to understand which objects basic empirical concepts are being applied to, the adoption of such a policy for the discrimination of stimuli is disastrous. For what is classified by the protoconcepts that repeatable responses are going proxy for is not bells and tables and rabbits but only states of the responding organism. Nothing that looks like one of our ordinary empirical concepts, applying to ordinary observable objects, is within reach of such an approach. A distal strategy is required in order to get the protoconcepts represented by reliably differentially elicited noninferential response types to count as classifying and so applying to ordinary observable objects and properties. Understanding them this way involves respecting the language-learning situation in which these reliable differential responsive dispositions are established.

The most popular approach to identifying distal stimuli as what is classified by the exercise of reliable differential responsive dispositions is to appeal to *triangulation*. This is a strategy for picking out or privileging one bit of the causal chain of covarying event types that reliably culminates in a response of a distinguished type, by looking at the *intersection* of two such chains. The insight it develops is that the best way to pick a single *point* (the

stimulus) out of a *line* (the causal chain of covarying event-types that reliably elicit a response of the relevant type) is to *intersect* it with another line—another causal chain corresponding to another reliable differential responsive disposition.

One writer who employs such a triangulation strategy to address the problem of picking out distal stimuli as what a response is *about* is Dretske.<sup>43</sup> In order to pick out the distal stimulus he looks to the upstream intersection of two distinct “flows of information” (or causal chains of reliably covarying event-types) that reliably culminate in responses of the same type. A simple example of the sort of system he has in mind would be a thermostat that keeps the temperature of a room within a certain range by turning a furnace on and off. If the thermostat has only one way of measuring temperature—for instance by the bending of a bimetallic strip until it touches either the left electrical contact (too cold) or the right one (too warm)—there is no way, Dretske acknowledges, to say that what the system is responding to is the temperature of the room, rather than the temperature of the bimetallic strip or the curvature of the bimetallic strip or the closing of the circuit between the bimetallic strip and one or the other of the contacts. Notice that a pragmatist appeal to practical *consequences* of the response in question is of no help here; turning the furnace on affects not only the temperature of the room but also that of the bimetallic strip, its curvature, and so its relation to the electrical contacts.

The idea is that one can be entitled to such a description if the thermostat is slightly more complicated and has another causal route to the same response (turning the furnace on or off). If the thermostat has a second sensor—for instance a column of mercury supporting a float with an electrical contact that completes one circuit to turn the furnace on whenever the float is below one point (too cold) and turns it off whenever the float is above another point (too warm)—then the system has two ways of responding to the change in temperature in the room. Although for this second route by itself (just as for the first by itself) there is no feature of the system that entitles one to say it is responding to changes in the temperature of the room rather than to the temperature of the mercury or the length of the mercury column or the closing of the switches, when the two routes are considered together, they intersect in just two places—upstream at the change of temperature in a room (which is included in the “flow” or causal chain corresponding to each route) and downstream in the response of turning the furnace on or off.<sup>44</sup> Dretske shows how the general strategy of looking to the *intersection* of two reliable differential responsive dispositions might be funded from the resources of the responding system itself.

One might worry that Dretske has not in fact succeeded in responding to the general worry about how to justify describing the system as responding to a distal stimulus rather than a proximal one. For there is an objection available to his strategy that seems to reinstate the original worry. Why, it

might be asked, ought we not to conclude that even in the two subsystem case, what is responded to is a proximal stimulus, but a *disjunctive* one? The system turns on the furnace just in case *either* the temperature of the bimetallic strip is low enough *or* that of the mercury column is low enough, or alternatively, in case the curvature of the bimetallic strip is far enough to the left *or* the mercury column is short enough. (Again, pragmatic appeal to the practical consequences of entering this state will not solve the problem.)

This worry is connected to the complaint voiced already in Chapter 2, to the effect that mere differential responsiveness is not sufficient for identifying the responses in question as applications of *concepts*. The rationalist supplementation suggested there—that what is distinctive of the conceptual is the *inferential* role played by the responses that stimuli differentially elicit—is also what is required to exploit the triangulation strategy in connection with genuine concepts in a way that responds to the worry about disjunctive proximal stimuli.

Consider a man who reliably responds (as one wants to say) to the visible presence of rabbits by saying “Gavagai.” Suppose further that he is reliably differentially responding not just to rabbits, but to the presence of the distinctive (according to him) rabbit flies that are for him decisive evidence of the presence of rabbits, or that the visual cue he is using, as determined by a physiologist of perception, is a glimpse of the fluff around the tail of the rabbit. What is it about the situation in virtue of which he can be said nonetheless to be reporting not the presence of the rabbit flies or of the fluffy tail but the presence of a rabbit? The inferentialist response is that the difference is not to be found in the reliable differential responsive dispositions, not in the causal chain of covarying events that reliably culminates in the response ‘gavagai’, to which not only the rabbit but the flies or the fluffy tail belong. It lies rather in the inferential role of the response ‘gavagai’. For instance, does the commitment undertaken by that response include a commitment to the claim that what is reported can fly? Or is the claim expressed by ‘gavagai’ incompatible with the further characterization of the item reported as flying? If it is incompatible, then it is not the flies that are being reported. What determines which element of the causal chain of covarying events that reliably elicit the report is being reported is the *inferential* role of the report, what it *entails*, what is *evidence* for it, what it is *incompatible* with.

Assuming that the observable predicate corresponding to ‘flying’ has already picked out the things that fly, noticing that the report ‘gavagai’ could mean rabbit flies in case its applicability entails the applicability of ‘flying’ and could not mean rabbit flies in case its applicability is incompatible with the applicability of ‘flying’ is just what is wanted to pick out the distal stimulus the concept expressed by ‘gavagai’ is being applied to or is classifying. But the appeal to inference and incompatibility may seem just to put off the issue. How does ‘flying’ get to apply properly to flying things, and not

to whatever cues we in fact use in discriminating flying things—in short to one element of the causal chain of covarying event types that reliably culminate in its application? The answer must be that what the appeal to inferential role does is establish a sort of *triangulation*, or intersection of flows of information or reliable differential responsive dispositions. If ‘gavagai’ is used so as to entail ‘flying’, then whatever is properly responded to by the former expression must be properly responded to by the latter, so what is classified as gavagai must also be classified as flying, so ‘gavagai’ must apply to rabbit flies, and not to the rabbits that are their invariable (we are supposing) concomitants. In short, the appeal to inferential role, in addition to reliable differential responsive dispositions, involves triangulation of the sort that Dretske invokes, where two (or more) different reliable responsive dispositions of the system are invoked, so that their intersection can pick out a unique element of the causal chain of covarying events as the stimulus being classified by a response of a certain type. Because ‘flying’ will *not* be taken to apply to lots of things that merely hop, we can be sure that it does not mean *flying or hopping*, and so that ‘gavagai’ does not mean something disjunctive like *rabbit or rabbit-fly*.<sup>45</sup>

In sum, to make the triangulation approach to distinguishing distal stimuli work, one needs to look further ‘downstream’ from the response, as well as ‘upstream’—just as orthodox functionalism would lead one to expect. What picks one kind of thing out as what is being *reported*, from among all those that are being differentially *responded* to, is a matter of the inferential commitments that response is involved in. These inferential consequences of going into a state make it clear that what is being classified is something outside the system. They are what determine that a physicist is reporting the presence of a mu-meson in a bubble chamber, and not simply a large hook-shaped pattern. For the consequences of classifying something as a microscopic mu-meson are quite different from those of classifying something as a macroscopic hook-shaped trace. It is the lack of such consequences that makes Dretske’s dual thermometer liable to a disjunctive proximal interpretation. The conclusion is that *causal triangulation* by intersecting causal chains associated with reliable differential responsive dispositions must be supplemented by *inferential triangulation* associated with different concepts.

The minimal condition on singular reference that has been extracted from Frege in this section amounts to the demand that objects be picked out by *substitutional triangulation*. Taking it that an expression is being used to pick out an object is taking it that that *same* object could be picked out in some *other* way—that some commitment-preserving substitutions involving that expression are in order. Substitutional commitments are compound inferential commitments, corresponding to *patterns* of simple inferential commitments. Substitutional articulation is a kind of inferential articulation, and substitutional triangulation is a kind of inferential triangulation. The notion of substitutional commitments is what is needed to explain what

it is to take two distinct claims (whether responsively elicited or not) to be applications of concepts to the *same* object. The significance of causal triangulation is to be understood in terms of the supporting role it can play in this sort of substitution-inferential triangulation. It cannot by itself provide an analysis of picking out objects. And as Frege's discussion of picking out abstract objects shows clearly, however important a role it plays in the way perceptible objects are given to us, causal triangulation is not even a necessary component of the substitution-inferential triangulation that is what our cognitive grip on objects in general consists in.

### 7. Conclusion and Prospectus

This is by no means to say that the discussion of substitutional triangulation here and in the previous chapter suffices to understand what our talking and thinking about objects consists in. That discussion addresses primarily the issue of what it is for it to be *objects* (and their properties and relations) that our talk (and so our thought) purports to be about. To understand fully what it is for our thought and talk to purport to be *about* them requires an account of the crucial *social* dimension of the substitutional triangulation that structures the contents expressed by the use of singular terms (and predicates) and of the inferential triangulation that structures the contents expressed by the use of sentences. The way in which the social structure of the broadly inferential articulation of discursive practice bears on the nature of the conceptual contents that practice confers on the intentional states it institutes (and on the performances that express them) is already implicit in the discussion of discursive practice in terms of deontic scorekeeping, in Chapter 3. It is the task of Chapter 8 to make it explicit, and thereby to show how the representational dimension of conceptual content arises out of, and essentially depends on, differences in social perspective among the various discursive practitioners.

A further shortcoming in the account of picking out objects in terms of substitutional triangulation as adumbrated so far is that it is primarily addressed to the phenomenon of *purported* singular reference. Although general reasons have been offered motivating a direction of explanation that begins with the notion of representational purport, it remains to say something also about the *success* of such purport. To this end, the next section discusses what we are doing when we take it that a singular term succeeds in referring, in that the object the term purports to refer to actually *exists*. An account is offered of *existential* commitment as a kind of *substitutional* commitment. This story in turn permits an analysis of the commitments characteristic of the use of expressions as definite descriptions, and so shows how to extend the deontic scorekeeping model from languages with predicates and lexically simple singular terms to ones that contain definite descriptions as well.

The rest of the chapter then addresses the structures that make substitu-

tional triangulation (and so recognizing an object as the same again) possible when *unrepeatable* expression tokenings are involved. This means above all the *deictic* or *demonstrative* use of terms that is so important for understanding the role causal triangulation can play in substitutional and inferential triangulation, and so ultimately for understanding what is distinctive about *empirical* knowledge. Moving to the level of *unrepeatable* tokenings requires discerning a finer structure of *token recurrence* below that of *substitution*, just as the finer substitutional structure had to be discerned below that of *inference*. The key concept in this account is that of *anaphora*. Explaining the anaphoric linkage of tokenings in terms of the inheritance of the determination of substitutional commitments provides an official account in deontic scorekeeping terms of the phenomenon by means of which the traditional semantic vocabulary, ‘true’ and ‘refers’, was explained in Chapter 5, redeeming the promissory note issued there. So by the end of this chapter the full three-leveled structure of fundamental concepts in terms of which conceptual content is to be understood in the semantic portion of the present account will have been made available: *inference*, *substitution*, and *anaphora*. At that point the semantic raw materials will be on hand to be combined with the underlying pragmatics to yield in Chapter 8 an account of representation by conceptual contents, on the semantic side, and objectivity of conceptual norms, on the pragmatic side.

## II. DEFINITE DESCRIPTIONS AND EXISTENTIAL COMMITMENTS

### 1. *Forming Singular Terms from Predicates*

To talk about the singular referential purport of singular-term usage is to talk about what kind of substitutional commitments one must attribute (and, as will emerge in the next chapter, acknowledge) in order for what one is doing—the practical deontic scorekeeping attitude one is adopting—to count as taking someone to be using an expression as a singular term.<sup>46</sup> To be doing that, one must treat the use of the expression as governed by proprieties determined by symmetric simple material substitutional commitments—commitments that in languages with sufficient logical expressive power are made explicit in the form of the nontrivial identity claims that Frege calls “recognition judgments.” The substitutional commitments involving a singular term that a scorekeeper attributes and undertakes determine the pragmatic significance, for that scorekeeper, of each use of that term. That the significance for deontic scorekeeping of its occurrences is to be determined in this way is what treating it as a singular term (as purporting to pick out an object) consists in.

The referential purport that is in this way acknowledged or attributed concerns the committive antecedents and consequences of application of singular terms as such. Chapter 6, some of the points of which were reca-

pitulated in Frege's terminology in the previous section, showed how those committive circumstances and consequences of application can be understood substitutionally. But referential *purport* is one thing, referential *success* is another. What is the difference between taking it that an expression has been introduced as purporting to refer to a definite object and taking it that it in fact picks out or gives us a cognitive or semantic grasp on such an object? This is a question about a certain kind of *entitlement* to the substitutional commitments in which singular referential purport consists, and so in a broader sense about the appropriate circumstances of application of singular terms as such.

The deontic attitudes that constitute taking the singular referential purport characteristic of singular terms to be successful emerge most clearly from consideration of what is involved in taking someone to be entitled to use a definite description formed from a predicate. This is what Frege calls "the definition of an object in terms of a concept under which it falls."<sup>47</sup> As an example of a definite description that is defective—whose referential purport is not successful because it involves substitutional commitments the user cannot in the relevant sense be entitled to—he considers the expression 'the largest proper fraction'. The predicate that description is formed from is one that can be used to express commitments with appropriate entitlements.

The expression "the largest proper fraction" has no content, since the definite article purports to refer to a definite object [*der bestimmte Artikel den Anspruch erhebt, auf einen bestimmten Gegenstand hinzuweisen*]. On the other hand, the concept "fraction smaller than 1 and such that no fraction smaller than 1 exceeds it in magnitude" is quite unexceptionable: in order, indeed, to prove that there exists no such fraction, we must make use of just this concept, despite its containing a contradiction. If, however, we wished to use this concept for defining an object falling under it, it would, of course, be necessary first to show two distinct things:

1. that some object falls under this concept;
2. that only one object falls under it.

Now since the first of these propositions, not to mention the second, is false, it follows that the expression "the largest proper fraction" is senseless.<sup>48</sup>

Suppose that a predicate *P* has been introduced and is in use. (This is a supposition that has been given definite content in substitutional terms by the discussion of Chapter 6.) The problem Frege is addressing is to make explicit what else is required for it to be proper to take someone to be entitled to use a definite description formed from it—a singular term of the form 'the *P*', or as it may be expressed more generally,  $\lambda x(Px)$ . Treating  $\lambda x(Px)$  as a singular term is taking it that its use is governed by symmetric simple

material substitution-inferential commitments, that is, that there is some true nontrivial identity involving it. Thus, there must be some true recognition statement of the form  $\exists x(Px) = a$ .

As already indicated, the two conditions Frege imposes are that there be something that is  $P$ , and that it be unique:

1.  $Pa$  and
2. For any  $y$ , if  $Py$ , then  $y = a$ .

The second condition—uniqueness—can be straightforwardly parsed in the substitutional idiom already available. For it just amounts to saying that for any terms  $t, t'$ , if  $Pt$  and  $Pt'$ , then  $t = t'$ ; the substitutional commitments associated with claims of these forms have already been explained. The first condition requires more discussion, however. In the *Grundlagen* Frege frames his dispute with the formalists in terms of the necessity of proving the existence of an object falling under the concept, or as he puts it, “producing something that falls under it.”<sup>49</sup> What must one do to satisfy this requirement?

## 2. Substitutional Commitments Expressed by Quantifiers

It too can be understood in substitutional terms. Existential commitments are a kind of substitutional commitment, related to, but not identical with, the substitutional commitments involved in the use of quantifiers. As the discussion of the formation of complex predicates in the previous chapter indicates, universal and particular quantifiers are logical locutions that have the expressive function of making propositionally explicit conjunctive and disjunctive substitutional commitments. Attributing commitment to a claim of the form  $(\exists x)Px$  is attributing commitment to *all* claims of the form  $Pa$ . Such a substitutional rendering of quantification has been criticized as inadequate in cases where, for cardinality reasons, there are not enough singular terms to pick out all the objects one is quantifying over. It is very important that we be able to make claims about all real numbers—for instance that every one can be represented by converging sequences of rational numbers—even though we are in principle limited to the use of at most a countable number of singular terms referring to them.<sup>50</sup> In fact, however, that the stock of available substituends is in this way *limited* threatens a substitutional construal of quantifiers only if that stock is in addition conceived of as being *fixed*.

It is of the essence of singular-term usage that new terms can always be introduced—both new terms for familiar objects and terms that introduce unfamiliar objects, paradigmatically by description. We cannot indeed extend our language so as to have separate terms for *all* real numbers at once, but *each* real number can be picked out. For there is no real number that we cannot specify—for instance by a definite description in terms of converging

sequences of rational numbers. The substitution instances  $Pa$  one becomes jointly and severally committed to by committing oneself to a claim of the form  $\{x\}Px$  include, not only those formed from terms  $a$  that are currently in the language, but all those that *could* be introduced (not necessarily simultaneously). The substitutional construal of quantificational commitments requires that the expressive powers of a set of discursive practices be conceived in the wider sense that takes account of the possibility of introducing novel expressions, rather than in the narrower sense that restricts attention to locutions already actually in use. This latter view amounts to freezing an idiom: taking a snapshot of it and evaluating its expressive capacities in abstraction from the process by which it develops. It is encouraged by thinking of languages as formal objects (perhaps set-theoretic structures) that have fixed vocabularies. If languages are instead conceived as living practices, then the ways in which new vocabulary is introduced take their place as fundamental aspects of those practices—as central as the ways in which new claims are made. Frege is the father of the formal approach to languages, but his project in the *Grundlagen* leads him to be vitally concerned with the process of introducing novel expressions functioning as singular terms, not only for unfamiliar objects of familiar kinds (by description), but even for unfamiliar objects of unfamiliar kinds (by abstraction).

Similarly, the use of a particular quantifier in connection with a complex predicate makes explicit a disjunctive substitutional commitment to the effect that for some term  $a$ ,  $Pa$ . To be entitled to such a claim one may, but need not, be able to produce the relevant substitution instance. The vindicating substituend  $a$  need not even already be in the language; one is committed only to the possibility of introducing such a term. The point of the existence requirement Frege imposes on entitlement to introduce definite descriptions is that a certain kind of bare stipulation is not in general enough to entitle one to such term introduction. One is not permitted without further ado to introduce the expression  $\exists x(Px)$  and then, relying on the fact that  $P(\exists x Px)$  (whenever use of the definite description is appropriate), to use that description as the substituend that vindicates the claim made by use of a particular quantifier. The large question of interest in this section is precisely what that existential condition on the employment of definite descriptions comes to.

Though it is common to do so, it is not necessary, however, to extend the existential condition Frege imposes on the use of definite descriptions to whatever counts as a vindicating substituend for a particular quantification. It is for this reason that the general formal notion of particular (that is, disjunctive) quantification should be distinguished from the substantive notion of specifically existential quantification. Free logics distinguish particular quantificational commitments from existential commitments so as to allow an idiom in which ‘Pegasus is a winged horse’ can count as true, even though Pegasus does not physically exist, and so in which ‘Pegasus’ can serve

as a substituend that vindicates a particular quantificational commitment to there being some winged horses. Frege, of course, does not want to talk this way, taking it that because 'Pegasus' has sense but no referent, 'Pegasus is a winged horse' cannot be true. For many purposes (certainly for Frege's), this policy is no doubt the best. Nonetheless there is nothing incoherent about scorekeeping practices that permit particular quantificational commitments to be vindicated by term substituends with respect to which the scorekeeper does not undertake existential commitments, and considering such ontologically relaxed idioms highlights certain important features of genuinely existential commitments.

The substitutional significance of particular quantification is entirely determined by features of discursive scorekeeping practices that have already been discussed if it is stipulated that a particular quantificational commitment with respect to the predicate  $P\alpha$  is to be equivalent to the commitment expressed by  $\sim\{x\}\sim Px$ . The negation of a claim  $p$  was defined in Chapter 2 as its minimum incompatible: the inferentially strongest claim that is commitment-entailed by every claim incompatible with  $p$ . Thus the claim that for some  $x$ ,  $Px$  is incompatible with any claim that for every term  $a$  entails some claim  $Qa$  that is incompatible with  $Pa$ . So the particular quantificational claim that for some  $x$ ,  $Px$  is both commitment- and entitlement-entailed by any claim of the form  $Pa$  (but not necessarily just by these).

### *3. Sortally Restricted Substitutional Commitments*

In his technical systems (both that of the early *Begriffsschrift* and that of the late *Grundgesetze*) Frege offers a substitutional account of the formation of complex predicates and of the formation of sentences from them by the application of quantifiers. In each case the scope of the quantifiers is unrestricted: *every* well-formed singular term can serve as the substituend that vindicates a particular quantification and can serve as a potential counterinstance to a universal quantification. One consequence of running these systems with their quantifiers wide open is that in order to give quantificational claims the force Frege wants them to have—above all for the assertion of claims formed by particular quantification to involve the undertaking of specifically *existential* commitment—Frege must ensure that it can be proven that every well-formed singular term has a referent. As Russell notoriously showed, another consequence—in the context of the expressive power provided by unrestricted formation of complex predicates or sentential functions by substitution—is that the resulting systems are inconsistent. That unpalatable result has prompted the investigation of how the various theoretical commitments that conspire to produce it might be relaxed so as to avoid it. One popular candidate is Frege's insistence on unrestricted quantification; the strategy of placing restrictions on the substitution instances relevant to the semantic evaluation of claims formed by the application of

particular and universal quantifiers is what lies behind the theory of types Russell develops in *Principia Mathematica* to evade Frege's difficulty.

Restricted quantification, however, is not of merely technical interest. Indeed, Frege's unrestricted version appears as an artificial extrapolation once it is realized that in natural languages ordinary quantificational tropes are sortally restricted.<sup>51</sup> The central uses of quantifiers are to make claims such as:

Every integer is the sum of nineteen or fewer fourth powers.<sup>52</sup>

Some nineteenth-century German philosophers did not care about ontological issues.

All bank employees must wear neckties.

A deer made those tracks less than an hour ago.

Each of us has intentional states.

The central quantificational construction is *every K* or *some K*, where *K* is a sortal expression such as 'dog' or 'book'. 'Everyone' and 'someone' have the sense of 'every person' and 'some person', and even the apparently wide open 'everything' usually carries some restriction, either explicitly, as in

Everything the author says about propositional content is confused,

or implicitly, as in

Everything is a disappointment in the end.

The sortal restriction puts conditions on allowable substituends, so that even though 'the author of *The Stones of Venice*' is a perfectly good singular term, substitution instances formed from it are not relevant to the semantic evaluation of "Every integer is the sum of nineteen or fewer fourth powers."

As Frege indicates in the *Grundlagen*, sortals are like predicates, except that they have not only criteria and consequences of *application* but (like singular terms) also criteria (and so consequences) of *identity*. For many purposes, '... is a dog' functions predicatively, just as '... is large' does. But if *a* is a dog and *b* is a dog, it makes sense to ask whether *a* is the *same* dog as *b*. Sortals have associated with them practices of identifying and individuating the things they apply to, as nonsortal predicates do not. So in order to introduce the sortal 'number', Frege insists on "a general criterion for the identity of numbers [*Kennzeichen für die Gleichheit von Zahlen*]."<sup>53</sup> When made explicit in the form of a claim, such a criterion has the form:

If *x* is a *K* and *y* is a *K* and *Rxy*, then *x* is the same *K* as *y*.

Introducing a sortal, like introducing a predicate, requires fixing the sense of claims formed by substitution into sentence frames of the form "*α* is a *K*"

(“ $\alpha$  is  $P$ ”), but it requires in addition establishing a criterion of identity. Satisfaction of this additional constraint ensures that  $K$ 's can be *counted*.

In fact, establishing a criterion of identity (and so a sortal) is not only sufficient for countability, it is necessary as well. Unsortedized ‘things’ or ‘objects’ cannot be counted. There is no answer to the question how many things there are in this room; there is one number of books, another of molecules, another of atoms, another of subatomic particles. As Frege says: “If I place a pile of playing cards in [someone's] hands with the words: Find the Number of these, this does not tell him whether I wish to know the number of cards, or of complete packs of cards, or even say of honour cards at skat. To have given him the pile in his hands is not yet to have given him completely the object he is to investigate; I must add some further word—cards, or packs, or honours.”<sup>54</sup> Counting is intelligible only with respect to a sortal concept.

‘Thing’ and ‘object’ are pseudosortals. They can occupy the syntactic positions occupied by sortals, but they do not individuate as sortals must. They are mere placeholders for sortals, used when for some reason—often the clumsy disjunctiveness of the sortal that would be required—one does not want to specify the relevant sortal explicitly. When we say something like “Put everything that is on top of the desk into the drawers,” we usually mean all the middle-sized bits of dry goods: books, papers, pens, paper clips, and so on. We do not mean ‘things’ such as designs in the dust, cool spots, drops of water, and so on. One of the central uses of ‘one’ in English is as an *anaphoric prosortal*—an anaphoric dependent standing in for a sortal that is its antecedent—as in “John quoted an English philosopher, and I quoted a German one,” or “Eric wants an ice cream cone, and Russell wants one too.” In these examples ‘one’ is used in the ‘lazy’ anaphoric way, where it is replaceable by its antecedent sortal. Like the pronoun ‘it’, however, it is promiscuously available to stand in for a wide variety of antecedents. ‘Thing’ and ‘object’ are what one gets if one misunderstands this grammar and instead construes ‘one’ as expressing a genuine sortal.

Frege in fact makes exactly this objection to the attempt to press the term ‘unit’ (or ‘one’) into generalized duty in place of substantive sortals in theorizing about counting. His own view is that the invocation of substantive sortal concepts cannot in this way be avoided; he takes it rather that “a concept [is] the unit relative to the Number which belongs to it.”<sup>55</sup> Not all concepts will do; only those expressed by sortals (rather than predicates without individuating criteria of identity): “The concept ‘syllables in the word “three”’ picks out the word as a whole, and as indivisible in the sense that no part of it falls any longer under the same concept. Not all concepts possess this quality. We can, for example, divide up something falling under the concept ‘red’ into parts in a variety of ways, without the parts thereby ceasing to fall under the same concept ‘red’. To a concept of this kind no finite number will belong. The proposition asserting that units are isolated

and indivisible can, accordingly, be formulated as follows: Only a concept which isolates what falls under it in a definite manner, and which does not permit any arbitrary division of it into parts, can be a unit relative to a finite Number.<sup>56</sup> This insight ought to have led Frege to see quantifiers as coming with sortal restrictions on the admissible term substituends. For quantifiers quantify; they specify, at least in general terms, *how many*, and how many there are depends (as Frege's remarks about playing cards indicate) on *what* one is counting—on the sortal used to identify and individuate them. As Frege saw clearly, the use of quantifiers depends on the use of the singular expressions that provide their substitution instances. It is best therefore to think not only of quantifiers but of singular terms as properly introduced only in connection with some at least implicit sortal.

Definite descriptions should be explicitly sortally restricted: 'the *man* in the brown suit', 'the *book* that Carlyle had to rewrite because of Mill's maid', and so on. Individual proper names and demonstratives and other indexical expressions cannot properly be understood except in terms of their associated sortals. Thus one cannot simply point in the direction of a statue of a man on a horse and christen "it" 'Lumpl'. It matters whether one is naming the statue or the lump of clay it is made of. If the former, reshaping it into a statue of a mother with a child destroys Lumpl; if the latter, not—for the transformed figure is the same lump, but a different statue.<sup>57</sup> If I hold up my copy of Kant's first Critique and ask "Has Eric read this?" my remark is susceptible to two different sorts of readings, depending on whether the demonstrative is associated with the sortal that individuates books according to the content of the text or rather (as might arise if I have just discovered a large jelly stain defacing the page that sets out the Table of Judgments) the sortal that individuates them according to particular physical copies. 'This' or 'that' used by itself should on these grounds always be understood as elliptical for 'this *K*' or 'that *K*'. Again, it is important that 'I' implicitly invokes the sortal that individuates persons—it is a *personal* pronoun. For I, who am flying to London, am the same person who last month flew to Philadelphia, while I am *not* the same passenger who did so.<sup>58</sup> The discussion in this vicinity about "relative identity" is prompted by this sort of observation. But it often takes the form of a mysterious thesis about *things*, rather than a clear one about the conditions that ought to be met to count as having introduced (or understood) a singular term (even a tokening of a demonstrative) as having a definite reference.<sup>59</sup> Such a confusion is the result of thinking of sortally unrestricted quantification and singular-term usage as conceptually fundamental, and seeing sortal restrictions as optional additions—rather than seeing the restricted case as fundamental, and unrestricted quantification as a dangerous and often unwarranted extrapolation based on a misunderstanding of the way pseudo- and prosortals such as 'thing', 'object', 'one', and 'item' function. Frege's requirement that to introduce a referring expression one must fix the sense of identities involving it—settle how it is

to be distinguished from other objects—turns out to require that it be associated, implicitly or explicitly, with an individuating sortal concept.

#### 4. Existential Commitments

Existential commitment is a species of substitutional commitment. It can be thought of as a particular quantificational commitment that involves a special sort of restriction on the vindicating substituends that determine the content of that commitment. Generically, the restriction on substituends characteristic of existential commitments is akin to the sortal restriction involved in quantificational commitments—and more generally in the use of any expressions conveying singular referential purport. The structure of the restriction is different, however. What is distinctive of specifically existential commitments is the special role played by a distinguished class of admissible substituends, here called *canonical designators*. The difference between the substitutional function of canonical designators and that of sortally qualified substituends in ordinary quantification is what stands behind the principle that existence is not a property—that existent objects are not a *kind* of object.

The best way to appreciate the role an expression must play to be functioning as a canonical designator with respect to a species of existential commitment is by examples. Three different sorts of existential commitment are considered here, corresponding to *numerical* existence, *physical* existence, and *fictional* existence: the sense in which there is a number such that every number greater than it is the sum of distinct primes of the form  $4n + 1$ , the sense in which there is a pen on my desk, and the sense in which there is someone who keeps house for Sherlock Holmes, respectively. By looking at these different kinds of existential commitment, it is possible to see what they all have in common, in virtue of which they deserve to be understood as species of a genus—so that ‘exists’ can be understood as univocal, in spite of the important differences between commitment to the existence of particular numbers, of physical objects, and of fictional characters. The claim is that what these different sorts of existential commitments share—what makes them all varieties of *existential* commitments—is the way in which their pragmatic significance is determined by a set of expressions playing the role of canonical designators. The manifest differences between them are consequences of the very different sorts of expressions that serve as canonical designators in each case.

It is clear how the sense of the expression “the smallest natural number such that every larger one is the sum of distinct primes of the form  $4n + 1$ ” is to be determined. The predicate this definite description is formed from has clear circumstances and consequences of application, and its inclusion of the specification ‘the smallest’ ensures that if that predicate applies to any natural number, it is to a unique one. (In this way it can be contrasted with

the predicates 'largest possible fraction' or 'most rapidly converging sequence', which Frege considers.) What does it mean to say in addition that in fact there is a number to which the predicate applies—that the smallest natural number such that every larger one is the sum of distinct primes of the form  $4n + 1$  exists? What is it for the definite description not only to purport to refer to or pick out a particular number but actually to succeed in doing so?

In this case its success consists in the truth of this identity:

The smallest natural number such that every larger one is the sum of distinct primes of the form  $4n + 1 = 121$ .

Entitlement to the existential commitment implicit in the use of the definite description, and so entitlement to use that description, can be secured by entitlement to *any* identity of this form. Of course the existential claim is not equivalent to the substitutional commitment that is made explicit by this particular identity—or indeed, any other of this form. The existential commitment is rather equivalent to the disjunctive claim that some identity of this form is true; the significance of the existential commitment is determined by its being incompatible with any claim that is incompatible with all claims of the form of this identity. The significance of such existential commitments is accordingly to be understood, and their propriety assessed, in terms of the class of vindicating substituends supplied by identities of the form of the one above.

What is the relevant form? Not just any identity will do. For instance

the smallest natural number such that every larger one is the sum of distinct primes of the form $4n = 1$	=	40 less than the smallest natural number such that every larger one is the sum of distinct primes of the form $6n - 1$
---	---	---

is an identity (substitution license) that does not, like the one above, settle it that the smallest natural number such that every larger one is the sum of distinct primes of the form  $4n + 1$  exists. It just links that question to the question of whether the smallest natural number such that every larger one is the sum of distinct primes of the form  $6n - 1$  exists. If the latter number exists, then so does the former. The claim is, however, that in identifying the first number with 121 (or the second with 161), one is doing more than merely settling this sort of conditional existential question. One is in that case settling the categorical existential question of whether the existential commitments implicit in the use of these definite descriptions are in order, whether those descriptions are successful singular referring expressions, whether the numbers they purport to specify exist.

To say this is to say that the issue of the success of their singular referential purport does not arise for expressions such as '121' and '161' in the same way that it does for expressions such as 'the smallest natural number such

that every larger one is the sum of distinct primes of the form  $4n + 1'$ . It is to take a frankly inequalitarian approach to referential purport and its success.<sup>60</sup> Numerals are semantically privileged ways of picking out numbers. By contrast to definite descriptions of numbers, the well-formedness of numerals suffices for their referential success, guaranteeing that they pick out a corresponding object. Furthermore, distinct numerals are guaranteed to correspond to distinct objects. According this privileged status to a class of singular terms is treating them as *canonical designators* of a kind of object.

In the paradigmatic case of natural numbers, numerals such as '121' can serve as canonical designators because they are systematic abbreviations of successor numerals: elements of the sequence

$$0, 0', 0'', 0''', 0'''' \dots$$

Peano's axiomatization ensures that numerals of this form cannot fail to refer to numbers, and further that their lexical distinctness (marked by the number of successor marks they bear) suffices for the distinctness of the numbers they refer to. Claiming that some noncanonical numerical expression succeeds in referring to a number is just claiming that it is intersubstitutable (saving discursive commitments) with some element in the canonical sequence of successor numerals. Existential commitment with respect to this sort of object, natural numbers, is a disjunctive substitutional commitment linking the candidate numerical expression to some canonical substituend. Saying which number a numerical expression refers to is producing the canonical designator that is intersubstitutable with it. (In this sense one has not yet said how many seconds there are in a century when one asserts the identity

$$\text{the number of seconds in a century} = 100 \cdot 365^{1/4} \cdot 24 \cdot 60 \cdot 60.$$

One has only given a recipe that would make it possible, with some work, to say which number the definite description picks out—a recipe that guarantees at least that it does pick out some definite number.) Once the use of some expressions as canonical designators has been established, Frege's requirement that entitlement to use an expression as a singular term depends on its having been settled which object it refers to in a sense that includes distinguishing that object from others—the requirement that becomes explicit in the introduction conditions he imposes on definite descriptions—is satisfied for numerical expressions by settling it that there is some canonical designator linked to the expression in question by a true recognition statement: a nontrivial identity claim making explicit a simple material substitution-inferential commitment. Ensuring that novel singular terms are suitably substitutionally linked to canonical designators establishes both the existence and the uniqueness of the objects they pick out, and so secures the

success of the singular referential purport that distinguishes them as singular terms.

So the notion of canonical designators makes it possible to think of existential commitment as a kind of substitutional commitment. It can be thought of as a particular quantificational commitment in which the vindicating commitments that determine its content are restricted to canonical designators. It is important to notice, however, how differently such a restriction functions from the sortal restrictions associated with quantificational commitments generally. The sortal restriction is a restriction to kinds of *objects*, while the restriction to canonical designators is a restriction to kinds of *expression*. In substitutional terms, this means that if a sortal qualifies one of an equivalence class of intersubstitutable (that is, coreferential) singular terms, it qualifies all of them. If Kant is a person and Kant is the author of "Was ist Aufklärung?" then the author of "Was ist Aufklärung?" is a person. So the sortal restriction does not discriminate among different ways of referring to the same thing. Not so for the restriction to canonical designators. '9' is a canonical designator of a natural number,<sup>61</sup> and 9 is the number of solar planets, but 'the number of solar planets' is not a canonical designator of a natural number. It is of course a designator of a natural number: 'natural number' specifies a sortal, picks out a kind of thing. But 'canonical designator' picks out a kind of expression, not a kind of thing. So the structure of the restriction on admissible substituends involved in existential commitment is quite different from that involved in ordinary sortally restricted quantification. Existence is not a predicate or property, and existing things are not a *kind* of thing.

### 5. The Role of Canonical Designators

For a class of singular terms (for instance successor numerals) to have the status of canonical designators with respect to a kind of objects (for instance natural numbers) is a matter of the significance their use has according to the relevant discursive scorekeeping practices. The institution of that significance presupposes, rather than establishes, entitlement to use those expressions as singular terms, however. As Frege would be the first to insist, one cannot simply stipulate that the use of successor numerals as singular terms is in order. Like all expressions, their use must be governed by some nontrivial identities for it to count as the use of expressions as singular terms at all. The point of Frege's disagreement with the formalists is that merely laying down the Peano axioms is not enough to satisfy this requirement. Identities of the form

$0''' = \text{the successor of } 0''$ ,

$0''' = \text{the successor of the successor of } 0'$ ,

and so on will not do because such identities are in the relevant sense trivial;

they do not link two different ways of picking out an object because the expressions flanking the identity sign are merely notational variants of each other. Thus they cannot serve as genuine recognition statements; they do not express substantive material substitution-inferential commitments. Frege's response to this situation in the *Grundlagen* is to link the use of the successor numerals to that of other expressions already in use, by means of the method of abstraction. In this way the use of successor numerals is determined by its relation to the process of counting the previously discriminated objects that fall under some sortal. His definition then permits versions of the Peano postulates to be *proven* to hold for expressions introduced in this way. Thus entitlement to the use of successor numerals as singular terms is secured, and they become available for duty as canonical designators. One must be entitled to use expressions as designators first, and only then can they serve as *canonical* designators, which can be appealed to in explaining what existential commitment consists in.

As Frege and others have shown, once one understands existence claims regarding expressions that purport to pick out natural numbers, one can systematically extend that understanding to existence claims regarding rational, real, and complex numbers, and so on. The canonical designators that give sense to the notion of existence for rational numbers can be pairs of successor numerals (corresponding to ratios of natural numbers), for real numbers they can be converging sequences of canonical specifications of rational numbers, for complex numbers pairs of such canonical specifications of reals, and so on. (Notice that in order to be entitled to use a sequence of canonical rational number designators as a canonical designator of a real, one is obliged to *prove* the convergence of the sequence of numbers those designators pick out.) The idea behind the use of successor numerals as canonical designators in explicating existential commitments as a species of substitutional commitment is that to say that some numerical expression succeeds in referring—to say that a number corresponding to it *exists*—is to say that it has some address in the structured space mapped out by the successor numerals. This idea can be carried over, with some differences, to the case of *physical* existence.

To say that some physical object expression succeeds in referring, that the object it designates exists, is to say that it exists *somewhere* in space and time, that it occupies some spatiotemporal region. This is to say that it has some address in the structured space of spatiotemporal coordinates centered on the speaker.<sup>62</sup> The speaker who takes it that Pegasus does *not* (and never did) exist, while P. T. Barnum's elephant Jumbo *does* (or did) is claiming that a continuous spatiotemporal trajectory cannot be traced out connecting the region of space-time occupied by the speaker<sup>63</sup> to one occupied by Pegasus, while such a trajectory can be traced out connecting the speaker's region with that occupied by Jumbo. It is not that Pegasus must be conceived as not taking up any space or surviving for any time; it follows immediately from

his being a horse that he does both. It is that the region he occupies is inaccessible from here and now—"You can't get there from here." He does not exist in *our* space and time, the one that defines physical existence for us. The analog in the case of physical existence of the structured address space defined by the successor numerals is the structured address space defined by egocentric spatiotemporal coordinate descriptions. Thus the term 'Pegasus' is not properly intersubstitutable with any expression of the form 'the {winged} horse located at  $\langle x, y, z, t \rangle$  from here', while the term 'Jumbo' is intersubstitutable with an expression of the form 'the elephant located at  $\langle x, y, z, t \rangle$  from here'.<sup>64</sup> Thus, like numerical existential commitments, physical existential commitments can be understood as substitutional commitments involving a class of canonical designators (again a kind of *expression*, not a kind of *thing*).

Of course there are also disanalogies between the way the spatiotemporal designators that are canonical for physical existence work and the way the successor numerals that are canonical for numerical existence work. Herenow centered coordinate specifications of accessible spatiotemporal regions are, like successor numerals, guaranteed to succeed in their referential purport. But the canonical designators of physical *objects*, as opposed to the regions they occupy, must include sortal information as well: the statue and the lump of clay *may* occupy just the same spatiotemporal region over the whole course of their existence. The sortals relevant in this case are those where identity (or difference) of the spatiotemporal regions occupied guarantees identity (or difference) of the objects within the sortal (as opposed to across sortals, as in the lump/statue case). Thus if  $\text{horse}_1$  occupies region  $r$  and  $\text{horse}_2$  occupies region  $r$ , then  $\text{horse}_1 = \text{horse}_2$  (and if not, not).<sup>65</sup> The individuation of horses is parasitic on spatiotemporal individuation, in that if one has used the horse-specific criteria of application of 'horse' to stick labels only on horses, one then uses spatiotemporal coincidence or divergence to decide how many horses have been labeled and how many labels each horse has.

As in the case of numerical existence, these existential commitments can be understood as substitutional commitments involving physically canonical designators only where one is entitled to apply those canonical designators—which in this case are formed from sortals plus specifications of accessible space-time regions. In the case of successor numerals, this could be done wholesale—'producing' objects for them to refer to by abstraction, so that the Peano postulates could be proven. As in the case of using converging sequences of canonically designated rational numbers as canonical designators of real numbers, however, not only must one pick out a privileged general form as canonical, one must also settle which of the designators of that form are suitable for endorsement as canonical. Just as one must prove the convergence of each sequence of rational numbers that is put forward as a substituend that could vindicate a commitment regarding the existence of a

real number, so one must show in the physical case for each sortal-plus-region pair that the sortal properly applies to the region—that the region specified is occupied by an elephant. The variety of spatiotemporally individuating sortals means that there is nothing useful and general to say about how one becomes entitled to claims applying a sortal to a region. The appropriate circumstances of application for applying the sortal-derived predicate ‘... is an elephant’, or ‘... is occupied by an elephant’ to a particular space-time region are quite different from those of ‘... is (or is occupied by) an electromagnetic force field’. But these details concern the use of these particular sortals and predicates, not the notion of existence in general. The surplus significance of a commitment to physical existence lies in the *accessibility* to the one undertaking the commitment (via a continuous trajectory from here-now) of a spatiotemporal region to which the sortal (or its derived predicate) is properly applicable. For that reason appealing to the notion of a predicate or sortal being applicable to a region does not make this way of thinking about physical existence circular.

As a final example, *fictional* existence, existence in or according to a story, can be understood as having the same shape as that common to physical existence and the various sorts of numerical existence. To say that in or according to the Sherlock Holmes stories Holmes’s housekeeper exists (or that the expression ‘Holmes’s housekeeper’ succeeds in referring to an individual) is to say that that expression is intersubstitutable with some singular term that actually appears in the story (in this case a tokening of ‘Mrs. Hudson’). The singular terms that appear in the text that defines the fictional context can be considered as the canonical designators. Thus the claim that according to those stories Holmes’s archenemy exists but his fairy godmother does not involves undertaking a substitutional commitment regarding a canonical designator in the first instance, and a commitment incompatible with it in the second. Even if the phrase ‘Holmes’s archenemy’ does not ever appear in the text, the fact that ‘Professor Moriarty’ does occur there and that it can be deduced from what is said about him that this term is intersubstitutable with ‘Holmes’s archenemy’ ensures that ‘Professor Moriarty’ is a canonical designator that can vindicate the substitutional existential commitment. That no such term plays this role with respect to ‘Holmes’s fairy godmother’ is the thrust of the denial of even fictional existence to that creature.

For some purposes it is useful to consider as canonical designators not only expressions that actually occur in the stories but also those, like ‘Sherlock and Mycroft Holmes’s maternal grandmother’, whose applicability is entailed by what is said there, though they are never actually used. These boundaries are hazy, for it is not clear what auxiliary hypotheses one is entitled to appeal to in extracting the consequences of what we are told in the story. In most settings regularities of nature, even if not explicitly mentioned, seem safely carried over, but beyond that the matter seems one for decision rather than

discovery. To say this is just to say that the notion of fictional existence is itself hazy outside those objects actually mentioned in the text in question. Another regard in which fictional existence is ontologically indeterminate as far as singular reference goes is that a difference in canonical designators need not here entail a difference of objects. The story may simply be silent (even by implication) on the subject of whether the person who had last hired the hansom cab Holmes is riding in at a certain point was Mrs. Hudson (or Sherlock and Mycroft Holmes's maternal grandmother) or not. One can constrain such individuative issues by importing the physical spatiotemporal individuating apparatus into the fictional context; Pegasus was spatiotemporally accessible to Perseus, according to the story. Like the invocation of regularities of nature, these constraints only go so far, and certainly in typical cases far underdetermine the identity and individuation of the fictional objects referred to by canonical designators.

The point of this discussion does not reside in the particular choices that have been offered here as candidates for sets of canonical designators corresponding to different sorts of existence. It is that existential commitments can be understood as a special kind of substitutional commitment (akin to but distinct from sortally restricted particular quantificational commitments) by using the notion of a privileged set of substituends playing the special substitutional role of canonical designators. What one is doing in claiming that the largest number that is not the sum of the squares of distinct primes exists (its canonical designation is '17,163') is different in specific ways from what one is doing in claiming that the tallest edifice in Washington, D.C., exists; numerical existence is different from physical existence. But these different sorts of existence, or even senses of exist, have a structure in common that qualifies them both as notions of existence. It is that common structure that the notion of existential commitments as substitutional commitments restricted to canonical designators seeks to capture.

If something like this account of existential commitments is right, then kinds of existence are to be individuated by the associated sets of canonical designators. Picking different sets of canonical designators gives different senses of existence. So one who treats specifications of real numbers by *pairs* of sequences of specifications of rationals such that sufficiently late elements of one sequence are arbitrarily close to those of the other would, strictly speaking, mean something different by saying that a certain real number exists. Of course, as long as each canonical designator in the one set is intersubstitutable (coreferential) with a designator in the other, this difference would not make a difference. It would be significant only when, perhaps against all expectation, apparently equivalent specifications turn out to diverge in hitherto unconsidered cases.<sup>66</sup>

A set of canonical designators has a *structure*—paradigmatically that of the successor numerals or coordinatized specifications of spatiotemporal regions—which systematically provides *addresses* for all the existing objects

of the class in question. As was indicated above, this need not be a totality that is present in the language at any given moment; the structure may provide only recipes for producing suitable canonical designators, or even just criteria for recognizing them, as is the case for some definitions of real numbers in terms of converging sequences of rational numbers. It remains to say what it is to treat such a set of designators as privileged in the particular way necessary for them to be functioning as canonical designators defining a sort of existential commitment. This is a matter of the consequences of undertaking and attributing existential commitments. One of the consequences, of course, is that being entitled to an existential commitment regarding a definite description is a necessary condition of being entitled to use that description at all. Only scorekeepers who attribute an entitled commitment to the existence of some  $x$  such that  $Fx$  to an interlocutor take the use of a definite description of the form  $/x(Fx)$  by that interlocutor to be appropriate.

More important, anyone who does *not* undertake a commitment to the physical existence of the object referred to by a term  $t$  cannot endorse any claims of the form  $Pt$ , where  $P$  is a physical predicate. Under these circumstances one can be entitled only to endorsements of the fictional, 'according to the story' truth of a claim, which differ from endorsements that are not in the fictional mode—for instance in the unavailability as premises for practical reasoning of the claims that appear inside the scope of the explicitating 'according to the story' operator. Similarly, one does not endorse numerical claims whose expression essentially involves terms with respect to which one is not prepared to undertake numerical existential commitments, and one does not endorse fictional claims whose expression essentially involves terms with respect to which one is not prepared to undertake fictional existential claims. It is this intimate connection between existential and doxastic commitments that led Frege to forbid nonreferring terms in his ideal languages.

Picking out the set of descriptions of accessible spatiotemporal regions as playing the role of canonical designators with respect to claims of physical existence includes offering a gloss on what someone is saying who denies that Pegasus (or any winged horse) existed, or ever flew over Greece. Negative existential judgments—claims to the effect that something or other does not exist—have been the source of considerable philosophical confusion over the years. (A paradigmatic example is the doctrine that nonexistent entities must at least have some sort of "subsistence" for us to be able to refer to them in denying that they exist.) But understanding existential claims as expressing substitutional commitments with respect to a class of canonical designators yields a straightforward reading of the sense of such negative existentials. Denying numerical existence to the largest integer, or physical existence to Bellerophon's flying horse, is committing oneself to something incompatible with *all* of the identities, one side of which is a canonical designator of the

relevant sort and the other side of which is the expression ‘the largest integer’ or ‘Bellerophon’s flying horse’—or expressions anaphorically dependent on them.

On this relaxed account, there is no reason to boggle at claims that numbers or other abstract objects exist. One must insist only that a determinate sense have been given to such claims, by specifying the relevant class of canonical designators. Once it has been settled that a class of expressions functions as singular terms, if some of the sentences in which they have primary occurrences are true, then it is a criterion of adequacy on the specification of a class of terms as canonical designators relative to such claims that there be canonical substituents for them. Where those conditions are satisfied, corresponding existential claims have been given definite sense and are themselves true. (Of course, how *interesting* they are is another matter.)<sup>67</sup>

### III. SUBSTITUTION, TOKEN RECURRENCE, AND ANAPHORA

#### 1. *Inference, Substitution, and Anaphora*

Discursive practice has at its center the game of giving and asking for reasons; social deontic scorekeeping practices confer specifically *propositional* contents because they are *inferentially* articulated. The previous chapter examined the *substitutional* fine structure discernible within that inferential articulation—the substructure in virtue of which subsentential expressions can play an indirectly inferential role by making a systematic contribution to the propriety of inferences in which the sentences they occur in serve as premises or conclusions. That claims are articulated according to substitution inferences in turn presupposes a further level of structure. For substitution is not definable for individual unrepeatable expression tokenings. It requires some notion of token *repeatability*.

To begin with, in order to be available as reasons, sentences used to express claims must be at least in principle repeatable—both within and across interlocutors. Furthermore, for an expression to be used as a singular term (or predicate), it must be possible for it to occur in different sentences, combined with different predicates (or terms).<sup>68</sup> The definition of substitution inferences requires that occurrences of the same expression, whether term or predicate, be identifiable in both premise and conclusion. The first section of this chapter showed how Frege’s triangulation principle (according to which purporting to pick out a definite object depends on its having been settled what would count as recognizing that object as the same again) can be understood in terms of the substitutional construal of what it is to use an expression as a singular term. That discussion simply presupposed the availability of repeatable terms.

When analytic focus is sharpened from repeatables such as ‘Benjamin Franklin’ and ‘the inventor of bifocals’ to particular tokenings of such types

as 'this man' and 'it', however, more must be said about how such unrepeatables are grouped together into term repeatables. The structure that emerges as crucial to generalizing substitutional considerations so as to encompass expressions of this sort is just that appealed to in Chapter 5 in explaining the use of the traditional semantic vocabulary 'true' and 'refers'—namely *anaphora*. The rest of this chapter investigates this phenomenon, adding a final level of semantic analysis to the two considered already. The result is a tripartite theoretical semantic structure whose key concepts are *inference*, *substitution*, and *anaphora*.

## 2. Substitution and Repeatability

What sort of relations do the repeatable terms, predicates, and sentences that have been under discussion thus far stand in to the unrepeatable tokenings that instantiate them in the actual performances that are accorded pragmatic significance by discursive scorekeepers? What is it for two tokenings to be occurrences of the *same* term or sentence? Up to this point the examples considered have done nothing to discourage the supposition that those repeatables are just lexical *types*: equivalence classes of lexically cotypical tokenings. But cotypicality is neither necessary nor sufficient for tokenings to be occurrences of the same term in the semantically relevant sense of 'same term'. It is not sufficient because cotypicality cannot guarantee the correctness of substitution inferences of the form:

This organism is a mammal,  
therefore this organism is a vertebrate.

For the two tokenings of the type ⟨this organism⟩ may involve different demonstrations, and so be governed by different material substitution-inferential commitments. In that case they are not guaranteed to be coreferential (= intersubstitutable), as different (primary) occurrences of a single term are.

Cotypicality is not necessary for tokenings to be occurrences of the same term in the semantically relevant sense of 'same term' because the relation between the predicates *does* guarantee the correctness of substitution inferences of the form:

This organism is a mammal,  
therefore it is a vertebrate.

For the tokening of the pronoun is guaranteed to be governed by the very same material substitution-inferential commitments as the tokening that is its anaphoric antecedent. In the sense required by the discussion of the previous chapter, the latter of these is a proper substitution inference, and the former is not. The latter involves what are in the semantically relevant sense two occurrences of the same term, while the former does not.

These examples emphasize that the idea of substitution inferences (and

hence the assimilation of sentences as substitutional variants of one another) presupposes that repeatable expressions can *recur*—occurring now in one context, now in another. This makes no sense applied directly to unrepeatable tokenings. One might, of course, literally cut a token out of one written context and paste it into another, but this is just the sort of case that points to the need to focus on *tokenings* or particular unrepeatable *uses* of tokens, rather than on the tokens themselves. For the same token may be used to perform various different speech acts at different times. One sign saying “Dig here” may be moved from place to place on the campus during a treasure hunt; the different tokenings of ‘here’ have different referents (for instance, they are intersubstitutable with different definite descriptions of locations) at different times, even though just one token is involved. Token recurrence may be determined by cotypicality, but that is not the only structure of token recurrence. Indeed, as will emerge, if what it is for a term to have a cotypicality recurrence structure is for *all* tokenings of a given type to be tokenings of that same term, and hence be guaranteed to be coreferential, then *no* sort of expression has such a recurrence structure—not even proper names or definite descriptions.

Practical acknowledgment or attribution of expression recurrence (treating some tokenings as tokenings of the same term or sentence) is an attitude that is implicit in the adoption of substitutional, and hence of inferential commitments. A particularly vivid reminder of the implicit presupposition of the reidentifiability of *terms* that stands behind the substitutional reidentifiability (and hence identifiability) of *objects* is provided by consideration of the way identity claims work in making substitutional commitments explicit. Identity locutions make explicit the claim that two terms pick out the same object. Their defining use is accordingly in explicit substitution inferences of the form:

$$\begin{aligned} \Phi a, \\ a = b \\ \text{therefore } \Phi b. \end{aligned}$$

The correctness of an inference of this form depends on the tokening of the type *(a)* occurring in the first premise and the tokening of that type occurring in the second premise being tokenings of the same term, and similarly for the tokening of type *(b)* occurring in the second premise and in the conclusion. Clearly identity locutions cannot be used to make *this* sort of implicit presupposition explicit in turn, on pain of an infinite regress. For the use of identity locutions to license substitutions presupposes the possibility of tokening recurrence between elements of identity claims and elements of the sentences substitutionally governed by them. Of course, like the substitutional commitments made explicit by identity locutions, commitments to treating some tokenings as recurrences of others can be made explicit if suitable further vocabulary is introduced. If the two premises and

the conclusion of the inference above are denominated (i), (ii), and (iii), respectively, then the metalinguistic apparatus for talking about tokenings that was introduced in Section IV of Chapter 5 can be used to do just that.

The first tokening of type ⟨a⟩ can then be picked out as /a/\_i, the second as /a/\_ii, and the claim that they are tokenings of the same term (not just of the same type) expressed by saying “⟨a⟩\_ii is a recurrence of ⟨a⟩\_i” abbreviated as Recur⟨a/\_i, /a/\_ii⟩. The reason the previous substitution inference is a good one is then that Recur⟨/This organism/\_premise, /it/\_conclusion⟩, and the reason the prior inference need not be a good one is that it can happen that ~Recur⟨/This organism/\_premise, /This organism/\_conclusion⟩. So it is possible to express the sort of inference that is licensed by explicit identity claims by saying that the inference from

$$\begin{array}{ll} \Phi_a, & \text{and} \\ a = b, & \text{to} \\ \Phi_b & \end{array}$$

is a good one provided that Recur⟨a/\_i, a/\_ii⟩ and Recur⟨b/\_ii, b/\_iii⟩. Of course, such a further explication is of use only insofar as one can implicitly recognize the different occurrences of ‘a’, ‘⟨a⟩\_i’, and so on as recurrences of the same expressions. It is not in principle possible to use explicit stipulations to eliminate the need for reliance on implicit capacities to recognize recurrences.<sup>69</sup> For this reason, an implicit token recurrence relation is appealed to in what follows (much as an implicit substitutional variation relation was appealed to in the previous chapter).

### 3. Token Recurrence

There are two varieties of substitutional equivalence. These are *intraterm* and *interterm*, or *de jure* and *de facto* equivalences of tokenings. The former are (taken to be) binding on all interlocutors; the latter vary from doxastic repertoire to doxastic repertoire, according to the particular substitutional commitments undertaken by or attributed to an individual. Each attributor takes recurrence to bind all, in keeping track of significances of identificatory commitments and invocations of them by term use. But the identificatory or substitutional commitments themselves vary from individual to individual.

Substitutional structure requires both sorts. They cannot be defined separately, apart from their role in such a structure; one cannot have the one sort of equivalence without the other. What the intraterm equivalences are *for* is to be the vehicles of interterm substitutional commitments. These latter in turn presuppose them, in that they could not otherwise have content. The model of invariance of something under substitution involves changing something and preserving something else. The changes that do preserve the appropriate something define interterm equivalences.

The notion of recurrence or repetition without change that is presupposed by (is part of, as complementary to) the notion of change invoked in that definition is intraterm or *de jure* intersubstitutable equivalence. The fact that tokenings can have deictic and anaphoric significances means that the notion of recurrence presupposed by substitutional (and so inferential) relations cannot be reduced to that of lexical cotypicality. This raises three fundamental questions. First, once this issue has been distinguished from that of their being of the same lexical type, what does it mean to say that two tokenings are tokenings of the *same term* or sentence? Second, granting that expressions may exhibit a recurrence structure that can be represented by an equivalence class of cotypical tokenings, what *other* sorts of recurrence structure are there? Third, what difference does it make to the expressive power and function of a repeatable term (or expression of another grammatical category) which sort of recurrence structure it exhibits? That is, what sort of expressive impoverishment does an idiom suffer from if it lacks one or another of the different sorts of recurrence structure?

The terms 'term' and 'sentence' are usually thought of as picking out items that are lexically individuated—that is, by character rather than content, in Kaplan's typology. The *semantically* relevant recurrence classes of tokenings correspond not to this sort of repeatable but to what particular tokenings *express*. At the propositional level, it is possible to talk about different sentence tokenings that (in different contexts) express the same *claim*. It is in this sense, of shared content rather than character, that tokenings of "You are tired," uttered by me, and "I am tired," uttered by you, can under suitable circumstances be said to express the same claim. In the next chapter, where same-claiming is investigated in more detail in connection with the use of the 'that' clauses that specify the propositional contents of explicit ascriptions of propositional attitude, this semantic equivalence relation over sentence tokenings is grounded in practical proprieties of *deference* (see 8.4.3). This is a central structure of giving and asking for reasons as a *social* enterprise—of assertion as making premises available for *others* to use in inferences.

The semantically relevant recurrence classes of term tokenings considered here similarly correspond not to the characters of terms as lexical items but to the contents their tokenings express. Awkwardly, there is in ordinary philosophical parlance no generally agreed-upon expression that stands to 'term' as 'claim' or 'propositional content' stands to 'sentence'. 'Individual concept' and 'singular sense' are sometimes used, but each carries theoretical baggage it is best not to import here. The more neutral notion of the *conceptual contents* expressed by term use will be used here to indicate what is wanted. The use of such substantives is misleading in the present context in any case, for it suggests that what links the tokenings that make up the recurrence class of a term is their *shared* relation to some one content. This presupposes that recurrence classes are equivalence classes, and that is im-

portantly not always the case. It is best instead to look at how recurrence classes work and then to adjust talk of contents to fit their expressive function.

For an expression to be used as a singular term is for the inferential potential of sentences containing it to be determined in part by a set of material substitution-inferential commitments linking that term to others. This principle obviously applies to *repeatable* expressions, the tokenings of which can recur from sentence to sentence. Sharpening the focus to the level of unrepeatable tokenings requires that the idea of one term appearing in different sentences—presupposed by its being linkable to others by substitution-inferential commitments—be translated in terms of tokenings that count as recurrences of other tokenings. An unrepeatable tokening can occur in only one sentence tokening. So what it is for an unrepeatable *tokening* to have the significance of an occurrence of a singular term must be cashed out in terms of the inferential potential of the sentence tokening in which it occurs, and of other sentence tokenings in which tokenings occur that qualify as recurrences of the original tokening. For two tokenings  $/t_i$ ,  $/t'_i$  to be linked by a substitution-inferential commitment means that if  $/\Phi t_k$  is any sentence tokening containing a singular term tokening  $/t_k$  that is a recurrence of  $/t_i$ , what it expresses has an inferential consequence that can be expressed by a sentence tokening  $/\Phi t'_l$ , where  $/t'_l$  is a recurrence of  $/t'_i$ . Thus substitution-inferential commitments should be thought of as linking tokening recurrence structures. For a tokening to be used or treated as an occurrence of a singular term is accordingly for its significance to be determined by SMSICs relating the recurrence structure to which it belongs to the recurrence structures to which other tokenings belong.

These recurrence structures *may* be equivalence classes of term tokenings all of which share a single lexical type. In that case a substitution-inferential commitment linking an equivalence class of tokenings of the type  $\langle \text{Benjamin Franklin} \rangle$  and an equivalence class of tokenings of the type  $\langle \text{the inventor of bifocals} \rangle$  licenses inferences from what is expressed by tokenings of the form  $/\Phi(\text{Benjamin Franklin})_i$  to what is expressed by tokenings of the form  $/\Phi(\text{the inventor of bifocals})_i$ , and things look just the way one would have expected from the discussion in the previous chapter. But recurrence structures not only need not be restricted to lexically cotypical tokenings; they need not be equivalence classes at all. Given the role that recurrence structures of tokenings play in defining substitution-inferential relations, recurrence must be a reflexive relation; each tokening must (trivially) count as a recurrence of itself. After all, the significance of each tokening is guaranteed in advance to be governed by the same set of substitutional commitments that it is itself governed by. More substantively, it must be a transitive relation: if  $/t_k$  is a recurrence of  $/t_i$  and  $/t_j$  is a recurrence of  $/t_i$ , then  $/t_k$  is a recurrence of  $/t_j$ . For from the point of view of substitution, to say that  $/t_k$  is a recurrence of  $/t_i$  is to say that  $/t_k$  inherits the SMSICs determining its significance from

$/t_j$ . If  $/t_j$  in turn inherits the SMSICs determining its significance from  $/t_i$ , then  $/t_k$  inherits the SMSICs determining its significance ultimately from  $/t_i$ .

But it does not follow from the role of recurrence relations in defining substitution-inferential commitments that recurrence must be a *symmetric* relation.  $/t_j$  may inherit the SMSICs determining its significance from  $/t_i$  without its being the case therefore that  $/t_i$  inherits the SMSICs determining its significance from  $/t_j$ . If the inheritance runs in either direction, both tokenings have their significance determined in the end by the *same* SMSICs. Nonetheless, as will become clear, the direction of inheritance can make a difference in counterfactual situations: if the SMSICs governing  $/t_i$  were different, so would those governing  $/t_j$  be, but not in the same sense vice versa.<sup>70</sup> Where the recurrence structure is not symmetric, it has the form of a *chain* or *tree*, rather than of an equivalence class.

#### 4. Anaphoric Recurrence Structures

This asymmetric structure of recurrence (inheritance by one token of the substitution-inferential potential of another) is *anaphora*. For one tokening to be anaphorically dependent on another is for it to inherit from that antecedent the substitution-inferential commitments that determine the significance of its occurrence. Different interlocutors may disagree about what those commitments are, but to take it that there is an anaphoric connection is to take it that the use of the anaphoric dependent is correctly evaluated according to whatever substitutional commitments govern the use of its antecedent. One cannot settle the substitutional commitments determining the significance of the occurrence of a singular-term tokening without computing the recurrence class to which it belongs.<sup>71</sup> If one takes a certain tokening to be an anaphoric dependent, attributing a determinate substitutional significance to it accordingly requires identifying some other tokening as its antecedent.

Thus

- (p) Carlyle wrote his brilliant satire of Hegel, *Sartor Resartus*, in part to show that he was an important thinker

has two readings, depending on whether  $/he_p$  is taken to be anaphorically dependent on the tokening  $/Carlyle_p$  or on the tokening  $/Hegel_p$ . Those who understand the claim as involving the latter anaphoric commitment thereby take it to entail and be entailed by its substitutional variant

- (p') Carlyle wrote his brilliant satire of Hegel, *Sartor Resartus*, in part to show that Hegel was an important thinker,

for both of these tokenings of  $\langle \text{Hegel} \rangle$  are, as  $/he_p$  is on this reading, recurrences of  $/Hegel_p$ . It does not follow from this account, however, that dox-

astic commitment is always preserved by "replacing pronouns by their antecedents."

Whether it is or not depends precisely on how that phrase is understood. If "replacing a pronoun by its antecedent" means generating a substitutional variant in which another tokening of the same *type* as the antecedent is put in the place originally occupied by the pronoun, the principle is false. It is true only if it means generating a substitutional variant in which a tokening that is a *recurrence* of the antecedent is put in the place originally occupied by the pronoun. Thus

- (q) An influential British author wrote his brilliant satire of Hegel, *Sartor Resartus*, in part to show that he was an important thinker

is not, even on the first reading of the anaphoric commitment it involves, equivalent to

- (q') An influential British author wrote his brilliant satire of Hegel, *Sartor Resartus*, in part to show that an influential British author was an important thinker.

For the latter, but not the former is entailed by

- (q'') An influential British author wrote his brilliant satire of Hegel, *Sartor Resartus*, in part to show that his friend John Stuart Mill was an important thinker.

In light of the discussion (in Chapter 5, Section IV) of definitization transformations in connection with Chastain's similar examples, it is clear that what is wanted is rather

- (q''') An influential British author wrote his brilliant satire of Hegel, *Sartor Resartus*, in part to show that *the* influential British author [or *that* author] was an important thinker.

Though grammatically a definite description, *(the influential British author)* is not semantically a definite description, failing as it does of unique designation; and */the* influential British author/*q'''* should be understood as anaphorically dependent on, and hence as a recurrence of */an* influential British author/*q'''* (itself a tokening of a type other tokenings of which function quantificationally, rather than as singular terms). Similar remarks apply to the variant in which a tokening of *(that author)* is used as the dependent (again a tokening of a type other tokenings of which function differently, in this case deictically rather than anaphorically).

According to this account, understanding one tokening as anaphorically dependent on another is attributing (or in one's own case, acknowledging) a certain kind of commitment. In the language employed here, it is a commitment to the dependent tokening being a *recurrence* of the antecedent tokening. Recurrence commitments of this sort can be understood in terms of the

*inheritance* of substitutional commitments, which has already been discussed.<sup>72</sup> It should be emphasized that this is an account of what anaphoric relations consist in—or better, given the methodological phenomenism about normative statuses that governs the theoretical idiom employed here, about the practical attitudes that constitute *taking* or *treating* two expressions as anaphorically linked. It does not pretend to address the questions about anaphora that linguists and cognitive psychologists have been most concerned with—namely, when it is *correct* to adopt this attitude and treat one expression rather than another as the anaphoric antecedent of another, what *lexical* or *syntactic cues* there are for adopting this attitude, or how audiences *in fact* go about deciding which of various possible readings to adopt. The question of interest here is what it is to do the trick—what counts as doing that trick—rather than when it is called for or how it can be brought off.<sup>73</sup>

The basic claim is that tokenings can have their recurrence classes determined in different ways. The recipe for calculating the recurrence class may be one that looks first to the lexical *type* of the tokening, including as basic recurrences other possible tokenings of the same type. Anaphora, by contrast, is a way of computing the recurrence class of a tokening by reference to another *tokening*. It is from that tokening that the anaphoric dependent inherits or borrows its recurrence class (which in turn determines the substitutional commitments that determine its significance). The first recurrence mechanism is *symmetric*. It results in a core class of cotypical tokenings, each of which is related to the others in the same way: by means of their type similarity. The second is *asymmetric*. It depends on the distinct roles played by the anaphoric antecedent and the anaphoric dependent tokenings. The distinction crucial to understanding the way in which the notion of *substitutional* structure makes it possible to extend the notion of *inferential* significance from the sentential to the subsentential level turned out to be the distinction between the *symmetric* substitutional commitments that govern the use of singular terms and the *asymmetric* substitutional commitments that govern the use of predicates. Just so the distinction crucial to understanding the way in which the notion of *token-recurrence* structure makes it possible to extend the notion of *substitutional* commitment from the level of repeatable types to that of unrepeatable tokenings turns out to be the distinction between the *symmetric* recurrence structure governing the use of expressions such as (some occurrences of) genuine definite descriptions and the *asymmetric* recurrence structure governing the use of anaphoric dependents.

##### 5. *The Significance of Asymmetric Recurrence*

The asymmetry of recurrence reflected in the fact that interchanging the expressions playing the roles of anaphoric antecedent and dependent in general preserves neither the identity of claims nor doxastic commitment

pays substantial expressive dividends. Discursive practices that lack this structure are expressively impoverished in a variety of ways. The two most important concern the *empirical* dimension of discourse and its social or *communicational* dimension. Under the first heading, it is because they encompass noninferential reporting practices (with their attendant reliability structure of authority, as introduced above in 4.2–3) that languages can express empirical claims. Indexical or token-reflexive constructions, particularly deictic or demonstrative ones, play an essential role in such reports. But no language can have such constructions unless it also has asymmetric recurrence structures: *deixis presupposes anaphora*.

Under the second heading, the capacity to pick up another interlocutor's reference by using a pronoun is one of the central mechanisms by which communication is secured across the interpersonal gap created by differences in doxastic commitments (which induce differences in the inferential circumstances and consequences of application expressions are taken to have by different interlocutors). The significance of anaphora in this context is that it permits each interlocutor to produce utterances employing tokenings that have been stipulated to be recurrences of arbitrary tokenings by others. At both the level of sentences and the level of terms, such recurrence provides the basic points of contact between different repertoires of commitments (including inferential ones). The rest of this chapter is concerned with exploring these two sorts of expressive capacity that anaphora bestows.

The definition of anaphoric dependence allows many tokenings to be (treated as) anaphorically dependent upon the same antecedent. As the mention of the significance of anaphoric links among tokenings for securing communication among interlocutors suggests, the antecedent tokening may be uttered by someone other than the one who produces the tokenings that are anaphorically dependent upon it.<sup>74</sup> It is also possible for that antecedent itself to be anaphorically dependent on some prior antecedent. Since recurrence and inheritance of substitutional commitments is transitive, so is anaphoric dependence. It is in this way that anaphoric chains or trees are formed. They can be anchored or initiated by tokenings that are not themselves anaphorically dependent on other tokenings. These are *anaphoric initiators*.

This role can be played by expressions of two different kinds. Tokenings that acquire their substitutional significance symmetrically (recurring by type), such as some uses of definite descriptions and proper names, can serve as initiators; they are potential anaphoric antecedents that need not themselves in turn be anaphoric dependents. Also, indexical and demonstrative tokenings—which acquire their recurrent tokenings asymmetrically in the form of anaphoric dependents, without themselves being recurrences of any prior tokenings—can serve as initiators; indeed, if their occurrence is not to be cognitively and semantically idle, they *must* so serve. Being an anaphoric antecedent or dependent is a role that individual tokenings can play. Though

their lexical and syntactic types can suit them for those roles, tokenings are not *compelled* to adopt one or the other role by those types. It is crucial to the communicative function of anaphora that any term tokening (or sentence tokening—indeed any tokening that is not syncategorematic) whatsoever can function as an anaphoric antecedent.

Less obviously, with the exception of indefinite descriptions, there seem to be no expression types that preclude their tokenings from functioning as anaphoric dependents. That is, there are no other types *all* of whose tokenings function as anaphoric initiators—expressions *none* of whose tokenings can have the significance of anaphoric dependents. Definite descriptions, for instance, can be used anaphorically—typically as the result of definitization transformations of indefinite descriptions. Demonstrative constructions can also be used to form anaphoric dependents, as in:

Kant admired Rousseau, but that writer admired only himself  
and

Fichte fought for political tolerance in Germany, and partly because of his efforts this precious end was eventually achieved.

Anaphorically dependent uses of proper names are discussed below in 8.5.6. Since indefinite descriptions always also have quantificational and predicative uses (as in “Carlyle was an influential British author”), there are *no* expression types *all* the tokenings of which are recurrences of one another—or indeed (for that reason) even all coreferential. There can, of course, be sets of cotypical tokenings all of which are recurrences of each other (and hence coreferential). But they all belong to recurrence structures that can also contain anaphoric recurrences of other lexical types, and they all exclude some (possible) tokenings of that type.

#### IV. DEIXIS AND ANAPHORA

##### 1. *Demonstratives*

Substitutional commitments relate token repeatables. So unrepeatable tokenings must be sorted accordingly as some count as recurrences of others in order for any of them to have the sort of indirectly inferential significance in virtue of which their production can contribute to making a move in the language game. That sorting may be based only on the lexical or syntactic type instantiated by the tokenings, or it may be based on further features of the individual tokening. Within tokening-based approaches to explaining the significance in discursive practice of expressions and complex components of expressions, the two central phenomena appear to be anaphora and deixis.

In recent years demonstrative and indexical constructions have received a

lot of attention. Perry and Lewis and others have shown that such expressions are not eliminable in favor of expressions that recur by type, such as standard uses of definite descriptions (those that do not themselves contain indexical elements).<sup>75</sup> Evans has provided an elegant account of how the capacity to have demonstrative thoughts and the practical capacity to locate one's actions and perceptions simultaneously in egocentric and public space and time are interdependent.<sup>76</sup> The result of these discussions (and of the many others addressing the same topic) is a robust sense that deictic tokenings provide expressive resources that are essential to our conception of ourselves as empirically situated knowers and agents.

Discussions of the use of token-reflexive or indexical expressions usually do not include treatments of anaphoric dependency. Indeed, pronouns do not fit well with the paradigms that lead to this terminology for tokenings whose use depends on features other than their types. The idea behind talk of *token reflexiveness* is that what tokenings of types such as 'I' refer to depends on who utters the tokening, for types such as 'now' and 'here' it depends on when and where the tokening was produced, and so on. What a tokening of 'it' refers to indeed depends on features of the tokening itself (just which features depending on the type or character of the pronoun). But the relevant feature of the tokening is just its anaphoric antecedent, and this is not a feature that can be specified independently and in advance of semantic and pragmatic interpretation, the way speaker, time, and place can. (Indeed, whether it is to be counted a semantic or pragmatic feature on traditional ways of dividing up these issues is itself a nice question.) Similarly, thinking of pronouns as *indexical* demands that one be able to specify the index to which their semantic evaluation is relative, and once again the anaphoric relation of a tokening to an antecedent tokening is not happily assimilated to general indices such as speaker, time, and place, which are specifiable in nonsemantic terms. Anaphoric chains running through bits of discourse are not naturalistic features of them like which organism produces the tokening, or when or where it is produced. They are *normative* features attributed to the discourse by deontic scorekeepers, matters of conditional commitment or commitment inheritance—of the obligation that the significance assigned to, or score kept on, one part of the discourse answer in systematic ways to the significance assigned to, or score kept on, another.

It is interesting to consider the use of demonstratives in connection with this contrast. A common strategy is to assimilate demonstratives to the indexical paradigm. The semantic interpretant correctly assigned to a demonstrative tokening depends not only on the lexical type it instantiates but also on a further index, which is taken to function in a way analogous to the way the indices of speaker, time, and place function in the semantic interpretation of tokenings of 'I', 'now', and 'here'. That additional index in the case of demonstratives is the object indicated or demonstrated by the one using the

demonstrative. What sort of an index is this? Is what a speaker is indicating or demonstrating a naturalistic fact, specifiable in nonnormative, nonsemantic terms, as speaker, time, and place are? Or is it a normative, scorekeeping matter of how various commitments should be understood as related to one another, as anaphoric dependence is? Or is the dependence of the significance of demonstratives on what is demonstrated of yet a different sort, not to be assimilated to either sort of relation?

What is it to indicate or demonstrate an object? The idea that animates discussions of demonstratives in terms of the indexical paradigm is that the core phenomenon is *pointing*: the demonstrated object is the one pointed at. This idea encourages a picture of indication as a physical matter of picking out an object by extending the line formed by the knuckles of the index finger out until it intersects something opaque. Of course things are not so simple. Wittgenstein reminds us that even such a practice of pointing requires a great deal of social stage-setting—the untrained may be unable to transfer their attention beyond the tip of the pointing finger, or may perversely trace the line of indication in the wrong direction, from finger tip to base, and so take it that something behind the one pointing has been singled out. Again, he reminds us of the emptiness of ‘bare’ demonstration. The use of ‘this’ or ‘that’ must at least implicitly be connected with some sortal, for the same physical gesture can have the significance of pointing to a book or to its cover, its title, its color, its shape, and so on. Although everyone would concede that for reasons such as these the actual practice of pointing is a complex affair, the picture of the virtual line extending from the finger nonetheless exercises considerable force. It seems to promise a way of picking out objects by causal, rather than conceptual triangulation (though as the point about sortals indicates, one of the intersecting beams is quite hazy in the picture).

But this is a false promise. The one using a tokening of ‘that pig’ demonstratively in many cases need not *do* anything in order to have indicated or demonstrated one particular pig in the barnyard, provided that the unique salience of that pig has somehow already been established—whether through the efforts of the one employing the demonstrative or not. The requisite sort of salience is a motley; it can consist in distinguishing features of the perceptible environment, properties that are highlighted by their relation to psychological factors, conceptual stereotypes, background beliefs, the previous course of the conversation, and perhaps much else. As with anaphora, it is helpful to put to one side the difficult psychological question of how scorekeepers in fact determine salience and so often correctly discern what object is being indicated or demonstrated, and also the difficult conceptual question of what makes it *correct* to take or treat one object rather than another as the one being demonstrated. For the fundamental question concerns the scorekeeping attitude itself: What is it to *take* or *treat* some object as what is being indicated or demonstrated in connection with a demonstrative use

of some expression? What does taking or treating a singular-term tokening as having a demonstrative significance that ties it to a particular object consist in?

The function of demonstration is to pick out an object and attach a tokening to it—to settle that the tokening refers to that object (notice that the second ‘that’ is functioning anaphorically, not demonstratively in this sentence). In light of the general discussion of picking out or using terms to refer to objects, this means that the effect achieved by successful demonstration must be understood in terms of the *substitutional* role of the demonstrative tokening.<sup>77</sup> But a demonstrative tokening as such is unrepeatable; substitutional commitments govern repeatable expressions—those that can occur in more than one sentential context and that can be replaced in each context by others. Only accidentally and in very special circumstances would a speaker be in a position to repeat a demonstration of an object. Even where repeated demonstration is possible, it results simply in the production of another unrepeatable tokening, one that can be seen not to have the significance of a recurrence of the original by the fact that there is always the possibility that, unbeknownst to its author, it in fact picks out a different object. If demonstrative tokenings could not recur, then they could play no substitutional role, hence no inferential role, and so would be semantically and cognitively idle. They would in that case not be ways of picking out or talking about objects at all, but mere noises.

Of course the recurrence structure in virtue of which demonstrative tokenings can play a conceptual role is not far to seek. Demonstrative tokenings can be picked up anaphorically. Because they can have anaphoric dependents, demonstratives can figure in substitution inferences:

*That pig* is grunting, so *it* must be happy. I’m glad, because *it* is our champion boar, Wilbur.

Anaphoric chains that the demonstrative initiates are available to figure in substitutional commitments and the inferences they govern just as repeatable term types such as *{Wilbur}* are. Because they are, uttering the demonstrative can be understood as contributing to making a move in the language game, in particular as indicating an object. It follows that the capacity of pronouns to pick up a reference from an anaphoric antecedent is an essential condition of the capacity of other tokens (which can serve as such antecedents) to have references determined deictically. Deixis presupposes anaphora. No tokens can have the significance of demonstratives unless others have the significance of anaphoric dependents; to use an expression as a demonstrative is to use it as a special kind of anaphoric initiator.

One might choose to assimilate the use of indexicals generally to deictic uses in this regard. I can make a claim by uttering a token of the type *{John should be leaving the house now}*, only because I and another can later utter tokens of types such as *{If he had left then, he would have been at the*

meeting on time), in which the tokening of *<then>* should be understood as anaphorically dependent on the earlier tokening of *<now>*.<sup>78</sup> The claim would be that in such a context 'then' functions much as it does in contexts such as

Clothes were made differently in *the seventeenth century* than they are these days; *then* they were made by hand, while now they are made by machines.

The tokening /*then*/ here clearly is used anaphorically, with /*the seventeenth century*/ as its antecedent. It could be claimed that 'then' is *always* used in this way (though only at the price of allowing virtual or merely possible antecedent tokenings). In the same way, tokenings of 'you' could be understood as anaphorically dependent on (perhaps merely possible) tokenings of 'I' by others.

In these cases, however, by contrast to that of demonstratives, one can understand how recurrences of indexical tokenings are possible without invoking specifically anaphoric connections. The character *<I>*, for instance, is systematically linked to *<you>* (and *<he>* or *<she>*), in such a way that tokenings of the one can count as recurrences of tokenings of the other (so that the same content is expressed by their use) without that linkage having to be understood anaphorically. The way in which uses of *<here>* and *<now>* can be picked up by uses of *<there>* and *<then>* respectively seems to work in a way that is intermediate between the systematic interchange of speaker and audience that governs *<I>* and *<you>*, on the one hand, and the demonstrative element that is the essence of *<this>* and *<that>*, on the other. But in these cases, too, index-matching rules connecting the contexts of utterance can be formulated that specify which tokens of the one type count as recurrences of which tokens of the other. Nothing like this, though, is possible in the case of purely deictically significant tokenings. There recurrence can be understood *only* on the anaphoric model.

It has been argued that deixis presupposes anaphora. But if anaphora could be given a deictic analysis, then it would not follow that anaphora is the more fundamental phenomenon. Here one would look for a converse of the familiar cases in which apparently deictic or demonstrative expressions are actually playing the role of anaphoric dependents, as in "Kant was a Prussian Pietist, but *that philosopher* did not always think like one." It is hard to assess the promise of this idea in the absence of a detailed working-out of it. One idea would be to assimilate the relation between an anaphoric dependent and its antecedent to the relation between a demonstrative and what is demonstrated. Anaphoric dependents would be understood as indexical tokenings that referred to their antecedents. In a tokening /Hegel understood Kant's argument, but he did not refute it/<sub>i</sub>, the token /it/<sub>i</sub> would be understood as meaning what a token /that/ would mean, if it could be arranged that what was demonstrated was the antecedent tokening /Kant's argument/<sub>i</sub>.

But as it stands this cannot be how such a story would go. For in that case the tokening /it/<sub>i</sub> would be conceived of as intersubstitutable, not with other tokenings co-identified with /Kant's argument/<sub>i</sub>, but with tokenings that (presystematically) would be said to refer to the tokening /Kant's argument/<sub>i</sub>—such as /the very tokening of type (Kant's argument) that was just uttered (or tagged with the index i)/. That is not what Hegel understood but failed to refute, for he never heard of that tokening. Some mechanism would be required to get us from the demonstrated anaphoric antecedent as referent of the supposed demonstrative-anaphoric dependent to the referent of that tokening. It should be obvious from the discussion of Chapter 5 what sort of mechanism that is. It is precisely this job that defines some locution as meaning what 'refers' means. An operator that forms indirect definite descriptions of objects from direct definite descriptions (or other specifications) of term tokenings is just what 'refers' is. If the natural anaphoric analysis of such locutions is to be avoided (as pursuing this strategy would require), then an alternate demonstrative or more broadly deictic reading must be offered of 'refers' and locutions performing a cognate function, in order to offer a deictic analysis of other anaphoric locutions. If this is not forthcoming, then anaphora must indeed be seen as presupposed by but not presupposing deixis.

To summarize, then. The recent tradition has focused on the cognitive centrality and irreducible importance of deixis. Concern with "direct reference" has been developed in part by exploiting intuitions about the basic nature of the word-world link established by the demonstrative use of expressions—what we used to think of as the use of demonstrative expressions, before it became apparent that anything can be used demonstratively (just as anything can be used anaphorically). From this point of view it has seemed natural to make the point that linguistic significance is always the significance of a (possible) event or uttering, that is, of a tokening—which may then be connected to semantically relevant types according to various models, by insisting on the priority and irreducibility of unrepeatable demonstrative and indexical uses over descriptive repeatable ones. Anaphora, as another tokening-based phenomenon, has not seemed of the essence, for it deals only with intralinguistic continuations or preservations of something one must in principle be able to understand already without it. Thus if one cared about tokenings as well as types, it seemed natural to care about deixis first, and about anaphora, if at all, only later. It now appears, however, that such an attitude, natural as it seems, is not only strategically wrong-headed but actually incoherent.

## 2. Deictic Mechanisms Presuppose Anaphoric Mechanisms

Deictic uses presuppose anaphoric ones. One cannot coherently describe a language in which expressions have demonstrative uses but no pronominal uses (although the converse is entirely possible). For indexical

uses generally, like deictic ones, are essentially unrepeatable according to types. Different tokenings of 'this' or 'here' or 'now' are not in general recurrences of each other, or even co-identifiable. Yet it is only as repeatable (that is, as elements of recurrence classes) that they can be substituted for. Recurrence is presupposed by the possibility of substitution, and the possibility of substitution is presupposed by picking out occurrences as semantically significant (that is, as indirectly inferentially and so assertionally significant). Since deictic uses as such are not type-recurrent, that recurrence must be understood as token recurrence—in particular, as anaphoric.

In short, unless one could pick deictic uses up anaphorically to generate recurrence classes, one would not be able to involve such deictic tokenings in (undertaken or attributed) identificatory substitutional commitments, and so could not treat them as involving occurrences of singular terms. Without the possibility of anaphoric extension and connection through recurrence to other tokenings, deictic tokenings can play no significant semantic role, not even a deictic one. Deixis presupposes anaphora. Anaphora is the fundamental phenomenon by means of which a connection is forged between unrepeatable events and repeatable contents. No semantically significant occurrence of a subsentential expression can be discerned unless it is governed by substitution inferences, which requires token recurrence: no (semantically significant) occurrence without (the possibility of) recurrence.

One consequence of this claim is an argument that is similar in many ways to one of the most important arguments of "Empiricism and the Philosophy of Mind." There Sellars argues that it is not possible to conceive of a language consisting only of noninferential reports. Noninferential responsive reporting does not form an autonomous language fragment, does not constitute a set of practices one could engage in without also engaging in specifically inferential practices. Not every claiming or sentence tokening can be a noninferential reporting, the exercise of a reliable disposition to respond differentially to features of the nonlinguistic environment. For what distinguishes reports from any other response produced according to a reliable responsive disposition is precisely its possession of propositional content, that it means something that can be understood by the responder and by others. Possession of such content is in turn a matter of the inferential significance of the reporting response. Understanding reports requires being able to distinguish what further claims follow from them and what claims would provide evidence for them. This inferential articulation, which is the possession of conceptual content by the responses produced noninferentially, presupposes the possibility of making the inferences involved, drawing the conclusions or offering the justifications indicated. These will not be reporting uses of sentences, since they are arrived at inferentially. So not all claimings can be reportings, if even the noninferential language entries are to have the significance of reportings.

The conceptual dependence of deictic mechanisms on anaphoric ones

suggests a way of extending this argument from the level of sentences to that of singular terms. Just as reportings as events are not autonomously significant (they depend for their semantic significance on the possibility of connection with other claimings by inference), so demonstrations are not autonomously significant (they depend for their semantic significance on the possibility of connection with other term tokenings by anaphora). The same sort of mistake that Sellars diagnosed in those who were tempted to see claimings acquiring their content solely through the circumstances in which they are appropriately produced as noninferential reports or *Konstatierungen*—a content that is then only reflected in (consumed or presupposed by) inference—can then be discerned in those who would assign a similar priority (presupposing the possibility of autonomous significance) to deictic mechanisms of “direct reference” in securing references that are then merely reflected in or preserved by (and so are consumed or presupposed by) anaphora. The phenomenon here is exhibited for terms, but the significant extension of Sellars’s point is better thought of as one that moves from lexical-syntactic types to tokenings:

inference : anaphora	::	
semantic significance of types : semantic significance of tokenings	::	
reporting uses of sentences : deictic uses of terms.		

Just as it is their potential for inferential involvements that makes sentence repeatables bearers of contents, so it is the potential for anaphoric involvements that makes unrepeatable tokenings bearers of contents.

*Conceptual* articulation, it was claimed in Chapter 2, is in the first instance, *inferential* articulation. *Subsentential* expressions, it was claimed in Chapter 6, can be understood as *indirectly* inferentially articulated, and hence as conceptually articulated—in spite of the fact that they cannot play the role of premise or conclusion in inference—in virtue of the significance of their occurrence for *substitution* inferences. Unrepeatable tokenings, paradigmatically demonstratives, can now be seen to be *conceptually* articulated, for they can stand in *anaphoric* relations to other tokenings, and the chains thus formed can be involved in substitutional, and hence inferential, commitments. The use of a demonstrative may be elicited noninferentially as a response to an environing stimulus. What makes it a term referring to an object—rather than a mere conditioned response like “Ouch”—is its role as an anaphoric initiator of chains that can be the subjects of substitutional commitments. It is in virtue of those anaphoric connections that a demonstrative tokening can play a conceptual role.

Equipped with this thought, it is possible to address the question of what it is to take it that some particular object has been demonstrated (whether by an actual gesture or by implication). For now this attitude of taking it that

some particular object has been demonstrated appears just as a special case of taking the use of any expression at all to have picked out a particular object. It is to be understood in terms of the sort of cognitive, conceptual, or semantic triangulation discussed in Section I—recognizing an object as the same again by undertaking a substitutional commitment. Such commitments are made explicit in the form of identities employing expressions that are anaphoric dependents of the demonstrative tokening. The anaphoric relations may be implicit, as in:

*It is Wilbur (or the big one),*

or they may involve explicit use of what might be called (after the terminology of Chapter 5) anaphorically indirect demonstrative descriptions, as in:

*The pig he pointed to (indicated, demonstrated) saying 'that pig', is Wilbur (or the big one).*

Both sorts of tokenings, /It/ and /The pig he pointed to (indicated, demonstrated) saying 'that pig', should be understood as pronouns, anaphorically dependent on a prior utterance /that pig/.<sup>79</sup> Thus the concept expressed explicitly by locutions such as 'what one is *pointing to*', like what is expressed by locutions such as 'what one is *referring to*', must be understood anaphorically. Thinking of referring in the way that has been developed here (by appeal to the concepts of inference, substitution, and anaphora), one could say that referring cannot be explained in terms of pointing, because pointing must be understood in terms of referring.

Anaphora has been presented here as a kind of token recurrence—a relationship among tokenings that is presupposed by, and hence not analyzable in terms of, substitutional commitments. Taking one individual's tokening to be anaphorically dependent on another is not attributing a substitutional commitment; it is attributing a more primitive sort of commitment, one that determines which substitutional commitments regarding other tokenings are relevant in assessing the substitutional significance of the one treated as anaphorically dependent. It may seem that this is an unnecessary shuffle, that a further level of analysis need not be broached. For anaphorically related tokenings are coreferential, and treating two expressions as coreferential has been explained in terms of the practical deontic scorekeeping attitude of attributing substitutional commitments. The reason such an account will not do has already been indicated: substitutional commitments govern the use of *repeatable* expressions. Anaphora is required to generate repeatables from unrepeatable tokenings, paradigmatically deictic ones, where cotypicality does not carry even a defeasible presumption of coreference, hence not of (co-)recurrence.

Identity claims make substitutional commitments explicit as the contents of possible assertions and judgments. It was pointed out above that the use of these explicating locutions as inference licenses evidently depends on a

prior notion of term recurrence, since terms occurring on the two sides of the identity must be reidentifiable in the other premise and in the conclusion (respectively) of such inferences. It is only by means of anaphoric recurrence that deictic tokenings become accessible to identity claims at all:

That cat is watching television.

It is (=) the most spoiled cat in the room.

Therefore the most spoiled cat in the room is watching television.

The tokening /it/ here is not replaceable with another tokening of the type ⟨that cat⟩, saving the goodness of the inference. The closest thing along these lines that will do is to replace /it/ with a tokening of the type ⟨that same cat⟩ or ⟨that very cat⟩. These look like deictic uses but are in fact anaphoric dependents, as tokenings of ⟨the same cat⟩ or ⟨the cat just mentioned (or pointed at)⟩ would be in this context. The function of such locutions is precisely to make available tokenings that are recurrences of the one they (anaphorically) refer to.

### *3. Rigidity is an Anaphoric Phenomenon*

Recognizing that locutions such as 'that very *K*' play the role of operators that form anaphoric dependents sheds light on another aspect of demonstratives: their modal *rigidity*. Kripke introduced the term "rigid designator" to distinguish expressions (such as proper names) that pick out the same individual in all possible worlds, from those (such as definite descriptions) that do not. The pretheoretical phenomenon this theory-laden description addresses is the observation that even if Archie is the most spoiled cat in the room, Archie *might not* have been the most spoiled cat in the room (one even more indulged may just have stepped out for a snack); but it is not possible that Archie not have been Archie. Evaluated with respect to the actual situation, the claim that Archie is the most spoiled cat in the room is true. The coreference (that is, intersubstitutability) of these terms is a commitment actually undertaken by the evaluating scorekeeper. Evaluated with respect to other possible situations, the claim that Archie is the most spoiled cat in the room would not be taken to be true, since that description would then pick out another cat, say Ana; a different set of commitments undertaken on the part of the scorekeeper would entail that Ana, not Archie, is the most spoiled cat in the room. Against either background set of commitments, however, Archie would still be Archie. For the coreference of different tokenings of the proper-name type ⟨Archie⟩ is guaranteed by their being recurrences of one another. Definite descriptions are not rigid designators in that when modal contexts (which make explicit the inferential potentials of expressions when evaluated with respect to diverse sets of background commitments) are taken into account, different tokenings of the same definite description type

are not guaranteed to be coreferential. It follows that at least in these contexts, cotypicality of definite descriptions is not sufficient for tokenings to belong to the same token recurrence structure. Of course, cotypical tokenings of grammatically definite descriptions are not guaranteed to be coreferential in any case, because such descriptions have uses in which they play the role of anaphoric dependents; two tokenings of the type *<the author>* would not be expected to be coreferential if they were drawn from different anaphoric chains appearing in reviews of different books.

Rigidity is an anaphoric phenomenon. Instead of repeating the proper name, the modal claim above can be expressed by saying:

Archie is the most spoiled cat in the room, but *he* (or *that cat*) might not have been the most spoiled cat in the room.

The pronoun has as its antecedent, and so is a recurrence of, the tokening */Archie/* that appears in the claim about how things in fact are. The anaphoric chain to which it belongs is then available to specify that same cat in other possible situations. The anaphoric chain, in other words, denotes rigidly.<sup>80</sup>

Kaplan introduces a rigidifying operator 'dthat' with the stipulation that while  $\lambda x Dx$  is not in general a rigid designator (since its denotation varies from world to world), dthat( $\lambda x Dx$ ) is to be rigid, picking out in each world whatever  $\lambda x Dx$  picks out in the actual world.<sup>81</sup> His operator in effect forms a type all of whose tokenings are stipulated to be anaphoric dependents of a tokening in the actual world. Thus 'dthat' does systematically what expressions like 'he' and 'that very cat' do informally in specifying other possible situations involving the same objects that are picked out in the actual situation by contingently associated expressions. The idea that rigidity has something special to do with the use of demonstratives—which lies behind the choice of 'dthat' (a homonym of the demonstrative 'that') to express the rigidifying operator—arises precisely because demonstrative tokenings can recur *only* anaphorically, and hence rigidly.<sup>82</sup> Thus someone who says

*That bar* is the standard meter stick. If *it* had been heated, *it* would have been more than one meter long.

is using the anaphoric dependent to make the demonstrative recur, thereby displaying the same rigidity that would be made explicit in saying

*That bar* is the standard meter stick. If *it* had been heated, *Dthat(that bar)* (or *Dthat(the bar the speaker indicated)*) would have been more than one meter long.

Kripke's original discussion pointed to the distinction between the modal rigidity of proper names and the nonrigid behavior of the definite descriptions that often serve to fix the reference of those names.<sup>83</sup> In Kaplan's terms, proper names can be understood as rigidified, 'dthat'ed descriptions or dem-

onstratives. Understanding rigidity in anaphoric terms accordingly suggests that tokenings of proper names be themselves understood as anaphoric dependents—elements in an anaphoric chain that is anchored in some name-introducing tokening.<sup>84</sup> Causal-historical theories of proper names then appear as dark ways of talking about the sorts of anaphoric chains that link tokenings of proper names into recurrence structures. Concern about the nature of baptism or name introduction is concern about how this special sort of anaphoric chain or tree can properly be initiated, so that all the anaphorically linked tokenings it contains pick out one object. Concern about transmission of names is concern about what is involved in earlier tokenings being picked up as anaphoric antecedents by later ones. It is clear in these terms what is happening when different tokenings of the same lexical type are used as names of different people—how there can be more than one person called ‘George’ or ‘Aristotle’. In such cases there are just multiple anaphoric chains; the multiplicity of people who can be referred to as ‘George’ is a phenomenon to be understood by analogy to the way in which many people can be referred to as ‘she’. Investigations of the roles played by social-linguistic context and practices, or conventions, on the one hand, and individual name-user’s intentions, on the other, in determining what previous uses a particular tokening ought to be considered beholden to should be understood as investigations of which anaphoric chain a particular tokening ought to be considered to be part of.

The claim is not that assimilating these questions about the use of proper names to scorekeeping questions about recurrence commitments (commitments regarding the inheritance of substitutional commitments) solves at a stroke all the questions that have vexed causal-historical theories. On the contrary, those questions, when transposed into the anaphoric framework, typically address issues about what determines when it is correct to treat one tokening as anaphorically dependent on another. The present discussion pretends to address only the issue of what it is for a scorekeeper to take one tokening to be anaphorically dependent on another, not the specific practices by which such scorekeeping attitudes are assessed as correct or incorrect. Nonetheless, seeing these issues concerning proper names as special cases of general issues concerning anaphoric links pays certain explanatory dividends. (Some of these are exploited below in 8.5.5–6, in the discussion of Kripke’s puzzle regarding the behavior of proper names in ascriptions of propositional attitude.) For instance, it becomes clear that the issue of whether a name user remembers where the name was picked up from or is disposed to defer to some other individual more knowledgeable in its use is relevant to assessing uses of the name only insofar as they bear on the *commitments* that name user has undertaken; one may be obliged (according to a scorekeeper) by the circumstances under which one acquired a name even if one is ignorant or mistaken about them, and may be committed to defer to various authorities without being disposed to do so.

Fans of causal-historical theories of reference do not typically restrict these accounts to proper names. Natural-kind terms and many predicates are thought to function according to this model as well. Indeed, it is important to remember that the opposition between causal and descriptivist theories of reference that Kripke appealed to in his influential arguments against the latter arises specifically in addressing the reference of proper names. No one was ever a descriptivist about the reference of the predicates used to form descriptions—on pain of an obvious regress. Some other mechanism of reference was always envisaged for those predicates.

Thus it is natural to understand expressions such as “. . . is red” and “. . . has a mass of twelve grams” as having the denotations they do in virtue of their links to authoritative episodes of calling things red and measuring their mass in grams. According to the present suggestion, then, these expressions ought also to be understood as functioning anaphorically. Insofar as this is the right way to look at things, then, there is in fact only one primitive recurrence structure, namely the anaphoric one. Apparently type-recurrent expressions such as proper names and basic predicates should in fact be understood as having their tokenings linked by relations of anaphoric dependence. The relations between tokenings in these structures can be symmetric, by contrast to the asymmetry of paradigmatic pronoun-antecedent links, because and insofar as those tokenings owe allegiance to a common antecedent.

Such an approach, according to which *everything* works anaphorically, may seem to explain too much. If all these sorts of expressions have anaphoric recurrence structures, and such structures act rigidly in modal context, what room has been left for expressions that are modally flaccid, varying from context to context in the substitutional commitments determining their significance? Although basic predicates should be understood to recur anaphorically, compounds of them—in particular definite descriptions formed from them—need not. Although the recurrence structures determining the inheritance of substitutional commitments of their basic parts link them to antecedents in the actual world—that is why the answer to Lincoln’s question “If we agree to call the tail a ‘leg’, how many legs would horses have?” is still “Four”—the existential and uniqueness commitments involved in the use of definite descriptions are in modal contexts evaluated with respect to alternative situations.

Archie, the cat in front of the television, does exist; but he, that very cat, might not have—even though another cat much like him might have been named ‘Archie’ in circumstances much like those in which Archie was baptized, and might be sitting in front of the television. Evaluated with respect to a different set of commitments undertaken by the one keeping score (that is, with a different set of claims being taken-true, and so functioning as the facts), a description might pick out a different object than it does according to the actual commitments of the scorekeeper. For what it is

intersubstitutable with depends on what other substitutional commitments are in play. This is compatible with there being an underlying anaphoric recurrence structure governing the use of the components of the description, which determines which of the substitutional commitments in play are in fact relevant to determining the proper use of the description in question.

The foregoing remarks are not intended to present a theory of the use of demonstratives, proper names, and definite descriptions. They are meant rather to suggest reasons for recasting the standard problems regarding such expressions into the terms of the common framework provided by the notion of anaphoric structures of tokening recurrence. According to the picture being presented, taking someone to have used an anaphorically dependent tokening in making a claim is attributing an *anaphoric commitment*. An anaphoric commitment is a commitment to treating the dependent tokening as a recurrence of the tokening that is taken (by the one attributing the commitment) to be its antecedent. Recurrence commitments, the genus of which *anaphoric* commitments are a species, are commitments regarding the *inheritance* of *substitutional* commitments. In the same way, *inferential* commitments, the genus of which *substitutional* commitments are a species, are commitments regarding the *inheritance* of *doxastic* or assertional commitments.

The story accordingly has three layers. At the top, sentences can be understood as propositionally contentful in virtue of their use in expressing *claims*—that is, assertional commitments. The key concept at this level is *inference*, for what makes the contents expressed *propositional* is the role of sentences in giving and asking for reasons. Inferential connections among claims are understood in turn pragmatically, in terms of *consequential* relations among the attitudes by means of which score is kept on commitments and entitlements to commitments—how attributing one commitment entails attributing others, precludes entitlement to others, and so on. At the next level, subsentential expressions can be understood as indirectly inferentially contentful, in virtue of the significance their occurrence has for the inferential involvements of the sentences in which they occur. The key concept at this level is *substitution*, for taking subsentential expressions to be contentful consists in distinguishing some inferences as substitution inferences, some inferential commitments as substitutional commitments. The substitutional structure of the inferences sentences are involved in is what the contentfulness of their subsentential components consists in. At the lowest level, unrepeatable tokenings (paradigmatically deictic uses of singular terms) can be understood as involved in substitution inferences, and so as indirectly inferentially contentful, in virtue of their links to other tokenings in a recurrence structure. The key concept at this level is *anaphora*. For taking an unrepeatable tokening to be contentful requires associating it with a repeatable structure of the sort that can be the subject of substitutional commitments. Anaphoric inheritance by one tokening of the substi-

tution-inferential potential of another does just that. The articulation characteristic of specifically *discursive* commitments is to be understood most broadly in terms of *inference*, the details of which require attention to *substitution*, the details of which in turn require attention to *anaphora*.

## V. INTERPERSONAL ANAPHORA AND COMMUNICATION

### 1. *Communication*

Linguistic studies of anaphora typically distinguish between intrasentential anaphora and discourse anaphora.<sup>85</sup> The study of discourse anaphora is addressed not only to intersentential anaphora, where an anaphorically dependent tokening occurs in a sentence different from that in which its antecedent occurs, but also to interpersonal anaphora, where those tokenings are uttered by different interlocutors. The account of anaphora in terms of token recurrence—that is, as consisting in inheritance by one tokening from another of the structure that determines which substitutional commitments are relevant to its semantic assessment—applies to interpersonal (and hence intersentential) anaphora as well as to the fundamental kind of intrasentential anaphora. Indeed, certain features of that account stand out more sharply in the interpersonal case. Thus looking at examples of this kind a bit more closely provides an opportunity to clarify the sort of scorekeeping involved in attributing recurrence commitments, by specifying further the sort of inheritance they involve. More important, a new dimension of the expressive role of anaphoric connections among tokenings comes into play in the interpersonal case.

Anaphora as here conceived contributes two crucial sorts of expressive power to the idioms in which it is operative. First, as discussed in the previous section, anaphora makes possible the construction of repeatable expressions from unrepeatable tokenings. It is the mechanism by means of which unrepeatable tokenings are picked up and made conceptually—that is, ultimately, inferentially—significant. It is only because deictic and other indexical tokenings can recur anaphorically that their occurrence contributes to the inferential role played by sentences containing them, and hence that their occurrence can be counted as semantically significant at all; no occurrence without recurrence. Even where no overtly indexical expressions occur, this expressive capacity is crucial to the functioning of empirical languages.

Empirical languages are those that include noninferential reporting practices, and the authority of such reports is essentially tied to the particular unrepeatable tokening that is elicited as a response by the exercise of a reliable differential disposition. For such tokenings to have cognitive authority (for them to be available as expressing commitments to which interlocutors can be entitled and to which they can appeal in entitling themselves to further conclusions), their significance must be governed by substitutional

commitments. Such commitments relate recurrence repeatables, which in this case must be constructed anaphorically. Thus the report

The traffic light has just turned red

can serve as a premise from which to draw conclusions according to inferential patterns that can be made explicit in the form of such conditional principles as:

If it was red then, it will be green soon.

This expressive capacity is important even if it is the same interlocutor who makes the noninferential report and draws conclusions ("It will be green soon") from it; after the first glimpse the reporter may no longer be in a position to report the color noninferentially. But there is a second sort of expressive power anaphora contributes to discursive practice that arises *only* in the social context of interpersonal communication of information.

For information (whether true or false) to be communicated is for the claims undertaken by one interlocutor to become available to others (who attribute them) as premises for inferences. Communication is the social production and consumption of reasons. So communication (giving and asking for reasons) involves the interaction of the *inferential* articulation of *contents* that is at the center of the *semantics* presented here and the *social* articulation of discursive *commitments* that is at the center of the *pragmatics* presented here. The nature and significance of this interaction of the inferential and the social dimensions of discursive practice is a large and important topic. It is the subject of the next chapter, which argues that the representational dimension of propositional contents is a reflection of the essential role played in their specifically inferential articulation by differences of social perspective—that is, differences between the point of view of the one who undertakes a commitment and the points of view of those who attribute it. As a result, the contents of the claims that are deployed monologically in intrapersonal reasoning in soliloquy must be understood as having been conferred by public practices of deploying claims dialogically in interpersonal reasoning in conversation. Meditation is made possible by disputation.

In advance of that fuller discussion of communication, the treatment here of the specifically social dimension of the expressive function of anaphora can only be preliminary. Nonetheless, some of the cardinal points are sufficiently detachable to be available already at this point. Interpersonal anaphora plays an important role in securing the possibility of communication across the doxastic gap created by the differing commitments of speaker and audience. The capacity of those in the audience to pick up a speaker's tokening anaphorically, and so connect it to their own substitution-inferential commitments, is part of what makes it possible for them to understand the speaker's utterance by extracting information from it. Anaphoric connec-

tions among tokenings that are utterances by different interlocutors provide a way of mapping their different repertoires of substitutional commitments onto one another—a structure scorekeepers can use to keep track of how each set of concomitant commitments relates to the others.

Such correlation of the substitution-inferential commitments (and hence doxastic commitments) undertaken by a scorekeeper with those attributed to others is a necessary part of the *interpretation* that is the uptake by a scorekeeping audience of some speaker's claim. It is an essential part of being able to use others' judgments as reasons, as premises in the scorekeeper's own inferences (even just hypothetically) to assess their significance in the context of those collateral commitments. Interpretation in this sense is necessary even in the case where all parties share a language.<sup>86</sup> The reason communication requires interpretation of this sort is twofold. First, speaker and audience typically have *different* sets of collateral commitments—if they did not, communication would be superfluous. Second, the inferential significance of a claim (what its consequences are and what would count as evidence for it) depends on what *auxiliary hypotheses* are available to serve as collateral premises. So differences in background beliefs mean that a remark may have one inferential significance for the speaker and another for each member of the speaker's audience.

## 2. Frege and Kant on Fruitfulness

This point reaches deep into inferentialist approaches to semantic content. It will be recalled that in the definition of conceptual content that opens the *Begriffsschrift*, Frege acknowledges the role of collateral commitments serving as auxiliary hypotheses: two judgments have the same content if and only if “all inferences that can be drawn from the first judgment when combined with certain other ones can always also be drawn from the second when combined with the same other judgments.” The ‘always’ here signifies universal quantification over auxiliary hypotheses. It is not enough if there is *some* set of further judgments that yields the same set of consequences when combined with each of the candidates whose contents are being assessed. Such a requirement would obliterate distinctions of content, since for *any* two claims such a set of auxiliary premises can be found—in Frege's systems, for instance, any two claims have the same consequences when conjoined with a logical contradiction.<sup>87</sup> This quantification over possible sets of background beliefs accordingly is an acknowledgment that what follows from a claim depends on which further claims one is allowed to assume in extracting those consequences. (The dual point also holds, of course; for what constitutes evidence for a claim—its inferential circumstances, rather than consequences, of application—equally depends on the available auxiliary hypotheses.)

Frege makes more of this relativity of inferential significance to available

auxiliary hypotheses in the famous opening paragraph of "On Sense and Reference": " $a = a$  and  $a = b$  are obviously statements of differing cognitive value [*Erkenntniswerte*];  $a = a$  holds *a priori* and, according to Kant, is to be labelled analytic, while statements of the form  $a = b$  often contain valuable extensions of our knowledge and cannot always be established *a priori*. The discovery that the sun is not new every morning, but always the same, was one of the most fertile [*folgenreichsten*] astronomical discoveries." *Folgenreichsten* here is literally *richest in inferential consequences* (in what follows from it). The cognitive value of a statement is to be assessed by its inferential significance, by the difference that adding it to one's repertoire of endorsed judgments makes to what else one is committed or entitled to.<sup>88</sup>

This idiom, and in particular the understanding of analytic identities it expresses, is borrowed directly from Kant. In the section of his *Logik* entitled "Logical Perfection of Cognition as to Quantity,"<sup>89</sup> he says: "The magnitude of cognition may be understood in a twofold way, either as extensive or as intensive magnitude. The former refers to the extension of cognition and therefore consists in its volume and manifoldness; the latter refers to its content [*Gehalt*], which concerns the manifold validity [*Vielgültigkeit*] or logical importance and fruitfulness [*Fruchtbarkeit*] of a cognition, as far as it is considered as a ground for many and great consequences [*großen Folgen*] (*non multa sed multum*)."<sup>90</sup> Here content is understood in terms of fruitfulness in the sense of leading inferentially to many consequences. Kant's definition of analyticity to which Frege is appealing is similarly couched in these terms:

The identity of concepts in analytic judgments can be either explicit [*ausdrückliche*] (explicata) or non-explicit [*nicht-ausdrückliche*] (implicata). In the former case analytic propositions are tautological.

Note: Tautological propositions are virtualiter empty or void of consequences [*folgeleer*], for they are of no avail or use. Such is, for example, the tautological proposition Man is man. For if I know nothing else of man than that he is man, I know nothing else<sup>90</sup> of him at all.

Implicitly (implicite) identical propositions, on the contrary, are not void of consequences or fruitless [*folge- oder fruchtleer*], for they clarify the predicate which lay undeveloped (implicite) in the concept of the subject through development (explicatio).

Note: Propositions void of consequences must be distinguished from propositions void of sense.<sup>91</sup>

Implicitly identical propositions have an expressive role—namely developing the content of a term by making explicit some of its inferential consequences, as in analytic claims such as "The oldest living mammal is a vertebrate."

The circumstances under which claims of the form  $a = b$  "contain valuable extensions of our knowledge" are those in which, first, they are not

analytic (and hence *a priori*) in the explicative sense and, second, the knowledge they are extending includes further claims expressed using the terms *a* or *b*. For it is only in the presence of such auxiliary hypotheses that the identity licenses nontrivial substitution inferences. What it is for a claim to have a nontrivial cognitive value or content is accordingly defined by a *particular* quantification over possible sets of collateral commitments; there must be *some* context in which adding the claim has nontrivial inferential consequences. So the notion of content is being defined in terms of a more basic notion of the inferential significance of adding a claim to a set of antecedently endorsed claims. For a claim to have a nontrivial content at all is for the inferential significance of its endorsement to include nontrivial inferential consequences in *some* doxastic context, and for two claims to have the *same* cognitive or conceptual content is for their significances to comprise the same inferential consequences in *all* doxastic contexts. The primitive notion of inferential significance (of what follows from a claim and what is evidence for it) is explicitly relativized to a set of background claims—namely the set of those that are available as auxiliary hypotheses or collateral premises in extracting inferential consequences.

So even though Kant and Frege do not talk about the social dimension of inferential articulation, their elaborations of inferential conceptions of conceptual or cognitive content implicitly acknowledge that the inferential *content* of a claim manifests itself in different inferential *significances*—different claims counting as its consequences and potential evidence—from the perspectives provided by various sets of concomitant commitments. When Frege speaks of “extensions of our knowledge,” he is comparing the perspective available before a claim is added to the repertoire of commitments we undertake (and take ourselves to be entitled to) with the perspective available afterward. What is a fruitful (inferentially significant) addition from one point of view may not be so from another. Though Frege does not discuss the consequences this observation has for understanding synchronic communication connecting different doxastic perspectives, looking at that case is helpful in understanding the diachronic cases he does appeal to.

### 3. Quine, Communication, and Reference

The underlying point is that what a given endorsement of claim commits one to, is entitled by, and is incompatible with depends on what else one is committed to, on what collateral information is available as auxiliary hypotheses for the inferences in question. Quine appeals to this Duhemian relativity of evidential significance to total evidential context in the closing sections of “Two Dogmas” to enforce a constraint on theoretical concepts of meaning (that is, claim content). Transposed into the idiom in use here, his holist argument for relativizing the meaning of a claim to the “total theory” of which it is a part is that:

- the meaning of a claim is what must be grasped to understand it, and
- what is understood must at least determine the inferential significance of endorsing what is understood, but
- what follows from a claim depends on what other claims are available as auxiliary hypotheses, so
- any difference in collateral commitments means a difference in inferential significance, hence meaning.

The fact that the inferential significance of endorsements is always and in principle relative to collateral commitments available as auxiliary hypotheses shows just what Quine wants it to, and thereby gives a definite sense to the claim that “the unit of meaning” is the whole theory or set of concomitant beliefs.

What effect does this relativity of inferential significance have for understanding *communication*? Quine does not explicitly raise this issue because he systematically waffles on the question of whether his “webs of belief” or “total theories” are individual or communal, whether we each have a different “total theory” or all share one. The account being unfolded here of the social-practical structure of inferential articulation—and hence of propositional contents—is one way of trying to take account of the motivations that push him now to talk one way, now another. This is an important issue in the context of an argument for not distinguishing changes of meaning from changes of belief; what is one to make of the consequence Harman extracts from this theory, that when I notice a cloud pass in front of the sun, the meaning of all my words changes? It must be granted that the noninferential addition of this new commitment alters (at least slightly) the inferential significance of all the claims that I do endorse, and all those I might. At the very least, conditionals of the form “If there is a cloud in front of the sun, then *p*” clearly would come in this way to have a different potential for transforming my commitments, and this would in turn alter the inferential significance of any claim that could appear as the consequent of such a conditional (or the conclusion of the inference it makes explicit)—and that is any claim whatsoever.

But must this alteration of the inferential *significance* different claims have for me be understood as involving an alteration in the inferential *content* they express? The view developed in the next chapter is one according to which the *inferential* holism that requires the pragmatic significance of doxastically endorsing a propositional content to be relativized to a repertoire of concomitant commitments must be understood in the context of a *social* holism. That social holism requires the grasp of the semantic content whose endorsement has such a significance to depend on scorekeepers’ abilities to exploit relations among the different perspectives constituted by the different commitments undertaken by and attributed to those whose deontic scores

they keep track of. The significance for the understanding of communication of a holism that relativizes to a repertoire of background commitments either the inferential significance or the conceptual content of claims depends, like the significance of all commitments, on the auxiliary hypotheses that are available to serve as collateral premises in drawing inferential consequences from it. In this case, the background commitments it is important to be aware of take the form of a model of communication. Holism about inferential significances has different theoretical consequences depending on whether one thinks of communication in terms of *sharing* a relation to one and the same *thing* (grasping a common meaning) or in terms of *cooperating* in a joint *activity* (coordinating social perspectives by keeping deontic score according to common practices).

Communicating is naturally conceived of as *conveying* something. According to such a conception, before an episode of communication takes place only the communicating agent possesses what is to be conveyed; after successful communication the recipient possesses it as well. Overt performances serve as the *vehicles* by which what is communicated is transported from speaker to audience. In the Lockean version of this transportation model of communication, what is transferred is ideas—which are related to words by conventions, which are in turn reflected in the associations of the various interlocutors. Upon having an idea, the speaker associates a spoken or written word with it, and upon hearing or seeing the word, the audience associates the corresponding idea. Communication is a way for speaker and audience to achieve a shared idea.

The framework conception of communication as conveyance of something can be filled in by various particular notions of what is conveyed. Rather than ideas, it might be propositions, meanings, or information that speaker and audience are understood as sharing. Of course not every sort of performance that brings about a similarity between its maker and its taker provides a candidate for a conception of communication. The concept of communication involves that of *understanding*. What is to be communicated by an utterance is what its audience is to understand by it. What the producer of a meaningful performance has initially and what in the case of successful communication its consumers eventually acquire is something—a content or meaning determining the significance of the remark—that is understood by both parties. How the details of the conveyance model are filled in depends on a further conception of what it is to grasp or understand what is conveyed.

A problem arises if this commonsensical model of communication is combined with the inferentialist account of discursive practice in terms of deontic scorekeeping presented in these pages. According to that account, the fundamental communicative performance (making a claim) is acknowledging or undertaking a doxastic commitment. The sort of understanding or uptake of such a performance required for successful communication is for

the audience to figure that performance correctly in its score: to attribute the right commitment to the one making the claim. What makes the commitment a discursive commitment is its inferential articulation. In particular, the propositional contentfulness of a doxastic commitment consists in the material norms governing its role as premise and conclusion in inferences.<sup>92</sup> In the paradigmatic case of communicating by claiming, the audience's understanding of a claim must determine the inferential significance that adopting or believing that claim would have—that is, what one would be committing oneself to by endorsing it, what other commitments might entitle one to that endorsement, what other commitments are incompatible with it (and so preclude being entitled to such an endorsement), and so on. It is only insofar as the audience assigns some such significance to an utterance that a claim content is communicated or conveyed thereby.

Given the relativity of the inferential significance of a claim to the context of concomitant commitments available to serve as auxiliary premises, it follows that inferential significance is not preserved in communication—is not conveyed or transported from producer to consumer of communicational performances. For any difference in collateral commitment may involve a difference of inferential significance, understanding, and appropriated meaning. If I believe that Zoroaster is the sun and that its shining is his beatitude, then an utterance of "The sun is shining" means something different in my mouth than it does in your ears. If it is nonetheless possible for us to agree or disagree about that claim, that cannot be because it has the same significance for us. Inferential significance can be determined only relative to a total belief-set, so if what audiences understand must determine such significances, it cannot be independent of the context of collateral commitments. Since, as pointed out above, communication is superfluous in the case in which all commitments are shared (which alone would guarantee sameness of inferential significance), if inferential significances were what needed to be conveyed for communication to take place, communication would be impossible in all cases in which it was not otiose.

This is the line of thought that led theorists (such as Feyerabend<sup>93</sup>) who took Quine's inferential holism seriously to worry about the incommensurability of different theories or sets of commitments. Corresponding to the transportation model of communication is an accumulation model of progress—acquiring epistemic rights to more and more true claims fabricated out of a common stock of meanings or candidate belief contents. If these must at least include inferential significances, then since those significances can depend upon any collateral commitments, meanings are not shareable across theories, and so not establishable cumulatively as theory develops and changes. Given the very different background beliefs quantum theory has given us to govern our inferences involving the word 'electron', how can we so much as understand Rutherford's turn-of-the-century claim that electrons are particles with definite boundaries, orbiting atomic nuclei with definite

boundaries? Given that ‘electron’ meant something so different for him than it does for us, how is it possible for us to deny the very claim that he was making—as opposed to denying some *other* claim whose content appeals to *our* concept of electrons? ‘Incommensurability’ is the name given to this threat to our understanding of what communication is and how it is possible. It is a threat that arises for inferentialist approaches to meaning and understanding once the sensitivity of inferential significances to background beliefs is appreciated. Although this challenge is more pointed in the case of diachronic conceptual change in the history of scientific theories, the corresponding difficulty evidently confronts inferential role theories of the sort of content that is grasped and conveyed in synchronic, face-to-face, intralinguistic communication among interlocutors with different repertoires of doxastic commitments.

Quine himself quickly drew the conclusion that what matters semantically is not *meaning* but *reference*—what we are representing or talking about rather than just what we are saying about it. Although he does not put the point in the context of communication, this move reflects the realization that even if (in virtue of my Zoroastrian beliefs) the observation that the sun is shining means something different in my mouth than it does in your ears, you can still learn something from me that you can use in your own inferences—if and insofar as you understand me to be talking *about* the sun, and saying *of* it that it belongs in the class of shining things, that I am representing that thing as being in that class. Again, though Rutherford’s many false background beliefs make his claim that electrons orbit around an atomic nucleus mean something to him that is unintelligible in the context of post-quantum-theory background beliefs, we can still understand him to have been talking *about* electrons (the same things we refer to) and to have been representing them as having certain properties and standing in certain relations. The information communicated consists in the purely extensional content of the claim made. A difference in inferential significance and commitment is compatible with identity of referential commitment and achievement. Where common reference of terms and extensions of predicates can be secured in spite of inferentially different employments, progress is comprehensible as talking about more and more objects, invoking more and more predicate-extensions, and coming to say more and more true things about those objects, for instance by classifying them under the predicate extensions. This is the conclusion that comes out of the debate between Feyerabend and Sheffler.<sup>94</sup> It is clearly the lesson that the early, realistic Putnam drew and conveyed to students and admirers such as Field, Boyd, and Devitt. It even sets the terms for Putnam’s recantation of realism in “Realism and Reason.” Thus an important motivation for the emphasis on semantic extensions—the referential dimension of discourse—can be found in the concern with making intelligible the possibility of communication.

This strategy gives up on contents as inferential roles, in favor of a differ-

ent sort of primitive. Inference can then be reinstated at two levels. First, some inferential proprieties can be read off of inclusion relations among the extensions of expressions. More important, recognition of the relativity of extensions to various elements of context yields the notion of intensions, as functions from indices to extensions. Such intensions are a more robust sort of content, which can be seen to be shared by speaker and audience in favored cases. Also, more finely grained inferential proprieties can be read off of inclusion relations among the sets that serve as domains and ranges of the intension functions. This is a two-leveled scheme, starting with extensions and ascending to intensions as functions defined on them. Inferential significances (the inferential potentials of particular claims in particular doxastic contexts) play no systematic role. Instead of inferential significances varying from speaker to speaker, there are extensions varying from possible world (together perhaps with other indices) to possible world.

#### 4. Intensions

Analogy with this appeal to functions suggests that in the inferential case one might treat the inferential content expressed by a sentence tokening as a function, assigning to each repertoire of concomitant commitments an inferential significance. Such significances could be (crudely) thought of as ordered pairs of circumstances and consequences of application. The first element then might consist of sets of inferentially sufficient antecedent claims (those from which the claim in question can be inferred) and the second of a set of inferentially necessary consequent claims (those that can be inferred from the claim in question).<sup>95</sup> Since what is evidence for or commits one to a claim, and what it is evidence for or commits one to, depends on what background commitments are available as auxiliary hypotheses, inferential *contents* could then be thought of as *functions*. The content of each claim would be represented by a function that takes sets of concomitant background commitments as arguments and yields inferential significances as values.

The theoretical advantages of such a picture would accrue from taking inferential contents so construed as what is shared and communicated within a discursive community. Differences between the inferential significance that a claim has in the mouth of a speaker and the ears of an audience would then be compatible with a common understanding of what is being claimed. It would not then be necessary to concede the counterintuitive claim that the meanings of all one's words change (at least slightly) whenever one acquires a new belief. This explanatory advantage would be bought at a significant price, however. Unless the theorist is content with stipulative semantics—associating intension functions with expressions by fiat—an account must be offered of what it is about the way expressions are used that *confers* such contents on utterances and the states and attitudes

they express. What is required in this case is an answer to the question, What is it for an expression to be so used as to have associated with it one rather than another intension determining a function from doxastic context to inferential significance?

The generic difficulty with answers to this question stems from the very features that make an intensional response attractive in the first place. For functions of the sort in question are individuated so finely that it is hard to see how the use of an expression could determine that one rather than a slightly different one should be associated with it. In different forms this is the worry underlying Quine's rejection of intensions, Lewis's discussion of the relation of linguistic behavior to formal semantics for artificial languages in "Languages and Language,"<sup>96</sup> and Kripke's "finiteness" version of Wittgenstein's skeptical arguments concerning the underdetermination of use by meaning if meanings are conceived in standard ways. In each case the difficulty arises because one can in general construct a function that differs from a given one only for arguments that are in one way or another beyond the reach of behavioral dispositions. Where this is so it becomes difficult to see what is being envisaged (never mind how one could know that it is true) when it is said by the theorist that one rather than another of these behaviorally indistinguishable functions is nonetheless to be associated with a particular claiming.

This difficulty is particularly pressing in the case of communication across generations. Thinking of the communication of content in terms of shared intensions, functions from context of collateral commitments to significance, is most plausible as a response to worries about incommensurability for a synchronic linguistic community. Sharing intensions is speaking the same language in a strong sense. It is not clear how plausible such an account is in the diachronic case, where what is at issue is the possibility of incommensurability produced by conceptual change within a scientific tradition. Surely Rutherford or even Bohr did not and could not have shared the intensions contemporary physicists associate with such expressions as 'electron', 'mass', 'particle', and so on. It is not just that our views have changed substantially during the twentieth century, but that they have changed in ways unforeseeable by our conceptual ancestors of a few generations ago. It would require considerable argument to show that they had nonetheless used their expressions according to intensions that left room for all of our radical rethinkings, which could accordingly be represented just by differences in the context of assertional commitments with which each claim is conjoined. It is not possible to rule out such an approach a priori, but it is not surprising that it is hard to find a champion for an intensional transportation model of diachronic communication.

These cases provide one of the strongest motivations for adopting a different strategy: one that breaks with the conclusion Quine arrives at in "Two Dogmas" by distinguishing between a kernel of inferential (and perhaps also

doxastic) commitments that must be shared by those who count as grasping the content, concept, or meaning in question, and a shell of peripheral beliefs, which could differ without alterations of content. One privileges *some* of the inferences a concept is involved in as constitutive of it, treating the rest as warranted by collateral information. Grasping the concept then involves mastering only these essential inferences, and these are what interlocutors must share on pain of misunderstanding one another. Acquiring a new peripheral belief—for instance that a cloud now obscures the sun—would not then count as altering the concepts expressed by such words as ‘sun’.

The difficulty faced by this approach is just the one Quine emphasized: saying what it is about the practices of using expressions that deserves to be characterized as treating some claims and inferences involving a concept as essential to it, and others as providing merely ancillary information about what it applies to. In constructing artificial languages, one might simply stipulate that some commitments are to be in the first class, while others are in the second. Even then one would be obliged to say how the proprieties of using expressions then differ depending on how particular commitments are classified. But insofar as this apparatus aspires to contribute to the analysis of natural languages or languages in use, those features of discursive practice that confer such a distinction of status between conceptual and merely empirical commitments must be specified. In the present context, one would need to explain in scorekeeping terms the different roles played by the practical attitudes of taking or treating commitments as conceptual and empirical. Of course, the fact that Quine can find no trace in our discursive practice of an analytic/synthetic distinction by looking at such candidate attitudes as treating as unrevisable or as *a priori* is hardly decisive. Other possibilities are not far to seek. (One that has not gotten the attention it deserves is Sellars’s suggestion that the practical status that privileges concept-constitutive inferences is their counterfactual robustness.)<sup>97</sup>

### *5. A Three-Leveled Approach*

Nothing rules out such a strategy, but it is not the one pursued here. The present account substitutes a three-leveled approach for the standard two-leveled one.<sup>98</sup> Instead of beginning with extensions and defining intensions as functions from indices (including possible worlds, which provide a background of endorsed claims serving as the facts) to extensions, the story begins with the inferential significances of claims. The theory then moves down, defining the extensional dimension of discourse in terms of substitution-inferential commitments. Those commitments in turn determine equivalence classes of expressions corresponding to what is represented—what is talked and thought *about*. Various features of the interpretive scorekeeping practices appealed to in this move down from perspectival inferential significances to *extensions* then make it possible to

move *up* from those significances to propositional or conceptual *contents* (corresponding in some ways to *intensions*), which systematically relate the distinct perspectives responsible for the different significances claims have to different interlocutors. What has been left out of the traditional formalism of extension and intension, from this point of view, is precisely the *interpersonal communicational* dimension. Yet it is this that gives the discerning of extensions and intensions its connection to discursive practice, and hence the explanatory role in virtue of which alone it can be appropriate to call what is discerned *semantic* correlates.

The way in which concern with what is talked about arises in the process of mapping the repertoire of commitments of an interpreted interlocutor onto the repertoire of commitments of an interpreting interlocutor is discussed in the next chapter. That chapter also seeks to explain the sort of perspectival propositional contents that coordinated scorekeeping practices confer. The paradigm of communication as joint possession of some common thing is relinquished in favor of—or modified in the direction of—a paradigm of communication as a kind of cooperation in practice. What is shared by speaker and audience is not a *content-as-function* but a scorekeeping *practice*. Contents as functions from repertoires to inferential significances can be seen as implicit in such practices, but the practice can retain its identity even though the functions implicit in it are different (at different times, and from different doxastic points of view).

For what is implicit can be made explicit in various, not always compatible ways. From each doxastic point of view on a speech act there can be a content common to the one undertaking a commitment and the scorekeepers attributing it, but what is taken to be shared may be different from the points of view of different scorekeepers. Thus inferential *contents* are essentially perspectival—they can in principle be specified only from a point of view. What is shared is a capacity to navigate and traverse differences in points of view, to specify contents from different points of view.<sup>99</sup> Explaining this capacity is explaining what it is to take or treat (understand or interpret) someone's remark as representing or being about one thing rather than another. So what appear theoretically as distinct moves *down* from inferential significances to *extensions* by assimilating expressions as intersubstitutable (= coreferential), on the one hand, and *up* from those significances to *intensions* by relativizing them to repertoires of background commitments, on the other, correspond to aspects of a single interpretive activity of understanding, grasping a meaning—the cognitive uptake of communication that is deontic scorekeeping.

The perspectival nature of propositional contents and the way in which their essential representational dimension emerges from communicative scorekeeping practice is approached in the next chapter by considering what is made explicit in *de dicto* and *de re* ascriptions of propositional attitude. The role of anaphora in securing coreference across differences in perspective

can be considered here as an introduction, however. Anaphora serves to link the equivalence classes of expressions that are intersubstitutable according to one interlocutor to the classes generated by the substitutional commitments of others. The need for such a mechanism arises in the interpersonal context because the speaker may have different substitutional commitments from the audience. If the speaker believes that the first postmaster general of the United States is the inventor of bifocals, and the audience does not, the inferential significance of the claim "The inventor of bifocals spoke French well" is different from their various perspectives. The question then arises how those in the audience can manage to have an attitude toward the same claim the speaker is making, can agree or disagree with *it*, rather than some variant of it that they associate with the same noises or inscriptions. Given that speaker and audience disagree about whether the claim is about the first postmaster general of the United States, how can they nonetheless secure a common topic of conversation in order to argue about whether or not *he* spoke French well?

This way of putting the question contains the answer. (Compare the way the fact that one can assert the modal nonrigidity of the description 'the first postmaster general' by saying "Benjamin Franklin was the first postmaster general, but *he* [the man just referred to] might not have been" points to the central expressive function played by anaphoric relations in understanding that phenomenon.) Use of an anaphoric proform implicitly *stipulates* coreference with the anaphoric antecedent upon which it is semantically dependent. Thus differences in the substitutional commitments that determine the propriety of inferences involving 'the inventor of bifocals' according to speaker and audience can be bracketed and a common topic of conversation secured by using a tokening that is anaphorically dependent on the speaker's tokening. To respond to the speaker by saying "*He* did not speak French well" is to disagree with the claim made, *whoever* the inventor of bifocals might turn out to be. Indeed, if more than the object referred to is in question, the claim can be affirmed or denied by using an anaphoric dependent on the whole sentence, rather than just picking up one of its singular terms: the audience can say "That is true" or "What you claim is false." Interpersonal anaphora achieves just the effect that matters for securing communication in the face of differences in collateral commitments.

The capacity to use a pronoun that anaphorically picks up another's tokening is also a cardinal component of another important ability, one whose cognitive significance is often underrated. For pronouns enable us to talk without knowing what we are talking about. Thus a speaker can come late into a conversation in which someone is already being referred to as 'he' and can jump in—continuing that conversation with a remark such as "If he did that, he deserves whatever he gets." The speaker may under such circumstances have no idea at all of who it is that is being talking about. The form in which the later claims are expressed nonetheless commits the speaker

anaphorically to their being about whoever it is the others were already talking about. That is, a scorekeeper will assess the doxastic commitment the latecomer has undertaken according to whatever substitutional commitments that scorekeeper takes to govern the antecedent of the anaphoric tokening 'he'. Anaphora is a mechanism that permits undertaking and attributing commitments concerning objects that one need not be able to specify (nonanaphorically) if challenged. Thus one is not obliged to know or accept the descriptions by means of which the utterer of the anaphoric initiator might pick out the subject with respect to which both are undertaking and attributing commitments.

### *6. Speaker's Reference*

These cases of interpersonal anaphora show that one must be careful in thinking of anaphora as inheritance of substitutional commitments by one tokening from another. The anaphoric antecedent is what determines the substitutional commitments relevant to the assessment of the significance of its dependents. But in using a pronoun that is anaphorically dependent on a tokening uttered by another, one is not thereby bound by whatever substitutional commitments the other happens to acknowledge as governing that tokening. An interlocutor who disagrees with the speaker's assertion "The inventor of bifocals spoke French well" by saying "*He* did not speak French well" is not making an incompatible assertion by adding "And he was not the first postmaster general, either," even though the utterer of the antecedent of those pronouns *is* committed to the intersubstitutability (that is, coreference) of 'the inventor of bifocals' and 'the first postmaster general.' Although the divergence of perspective that makes the point evident did not arise in the case of intrapersonal anaphora,<sup>100</sup> the substitutional commitments to be inherited anaphorically by one token from another are assessed by the scorekeeper who attributes the anaphoric commitment, that is, who takes or treats the one tokening as anaphorically dependent on another. To take one tokening to be anaphorically dependent on another is to take it that it should be understood as governed by whatever substitutional commitments govern its antecedent.

Different scorekeepers may disagree about what these are, and they may disagree even with the ones producing the performances whose significance they are assessing. They may nonetheless all agree in attributing an anaphoric commitment, that is, in interpreting one tokening as being anaphorically dependent on (hence a recurrence of) the same antecedent tokening. A scorekeeper who takes it that the inventor of bifocals is the inventor of the lightning rod will take it that the first speaker claimed of the inventor of the lightning rod that he spoke French well, and that the second speaker claimed of that same individual that he did not speak French well. That is, a scorekeeper who undertakes such a substitutional commitment and attributes

that anaphoric commitment is obliged to take it that what the first speaker said is *true* just in case the inventor of the lightning rod spoke French well, and that what the second speaker said is true just in case he did not. For what a scorekeeper takes to be true is just what that scorekeeper endorses. The scorekeeping significance of attributing an anaphoric commitment is accordingly just that the significance of the dependent tokening is to be assessed according to the *same* substitutional commitments by which its antecedent tokening is assessed—whatever those are. Where the scorekeeper is concerned with when a given claim is true, it is the substitutional commitments that scorekeeper *undertakes* that matter, rather than those attributed to the utterer of the antecedent.

This is to say that according to a scorekeeper who undertakes a commitment to the intersubstitutability of ‘the first postmaster general’ and ‘the inventor of the lightning rod,’ one who asserts “The first postmaster general spoke French well” has thereby in a certain sense undertaken a commitment to the claim that the inventor of the lightning rod spoke French well. And this is true even in the case where the one making the original assertion would deny that the first postmaster general is the inventor of the lightning rod. The speaker is, according to such a scorekeeper, committed to that further claim just in the sense that what he has said is true if and only if the inventor of the lightning rod spoke French well. In this sense, what someone is committed to may (according to a scorekeeper) not only outrun, but even conflict with, what that interlocutor is prepared to acknowledge. The scorekeeper must keep two sets of books.

The necessity for this dual score follows from the fact that there are in principle two places a scorekeeper can draw auxiliary hypotheses from in extracting the inferential consequences of (and so the commitments consequentially undertaken by) a set of commitments some individual is taken to acknowledge. Those auxiliary hypotheses may be other commitments the individual acknowledges, or they may be commitments the scorekeeper *undertakes* (acknowledges), rather than attributing as acknowledged. Since these latter represent the *facts* (facts being just true claims), according to the scorekeeper, these latter consequences are those that, according to the scorekeeper, *actually* follow from the claims made (given how things really are), regardless of whether the one making the claims realizes that they follow or not. The relations between these two sets of books, and the way their interaction constitutes the representational dimension of propositional content, is the topic of the next chapter. As an introduction to the perspectival character of claim contents that is investigated there, it is helpful to consider the phenomenon of *speaker's reference* in terms of interpersonal anaphora.

What Kripke called “*speaker's reference*,” by contrast to “*semantic reference*”—a distinction closely allied to Donellan's distinction between “*referential*” and “*attributive*” uses of definite descriptions<sup>101</sup>—is a phenomenon that depends on the possibility of taking up identificatory or substitutional attitudes toward a tokening that is not treated as functioning in a type-recur-

rent way. It is a matter of the significance (substitution-inferential potential) that an audience attributes or ought to attribute to a particular tokening, by contrast to the significance that would otherwise be associated with it on the basis of its type. What is fundamental is the way an audience interprets or keeps score on the tokening. Once what it is to *take* someone to be speaker referring is understood, it will be possible to understand what it could be for some situation conventionally to call for or warrant the use of this sort of interpretation. So the account is in terms of audience uptake, not what the speaker does or intends. Like the cases just considered, the situation in which an audience counts as treating someone as having "speaker-referred to someone other than the one semantically referred to" by a remark is always one in which the identificatory commitments in the vicinity of the recurrence class of the uttering to be interpreted that the audience attributes to the speaker are different from those that the audience undertakes itself.

Adopting the subsentential forms a bit so as better to accord with tradition, a case might go like this. The speaker, Fred, says, "The man in the corner with champagne in his glass is very angry." According to Wilma, in the audience for this remark, Fred claims that Barney is the man in the corner with champagne in his glass. So according to Wilma (that is, the commitments she *attributes*), Fred might just as well have expressed his claim by saying, "Barney is angry." But according to Wilma (that is, the commitments she *acknowledges*), Barney is the man in the corner with ginger ale in his glass, and the man in the corner with champagne in his glass is Nelson. (She takes it that Fred does not see Nelson and does not believe that he is in the corner at all.) Then we can say that, according to Wilma, Fred has speaker-referred to Barney and attributed anger to him but has semantically referred to Nelson and attributed anger to him.

These two different ways of interpreting the claim that Fred has expressed by his utterance correspond to assessing his assertional commitments with respect to the identificatory commitments that Wilma attributes to him, and to assessing those commitments with respect to the identificatory commitments that Wilma herself undertakes. It is essential that, according to Wilma, there is some expression that Fred could (compatibly with the commitment Wilma attributes to him) have used to semantically pick out his referent, in order that he be able to speaker-refer to it by another expression. For what Wilma is doing when she assesses his remark as true by taking him to have speaker-referred to Barney is treating his tokening /the man in the corner with champagne in his glass/ as an anaphoric dependent whose antecedent is another tokening that Fred could have used (and would have used had he realized that there was a dispute about the matter): perhaps a tokening of the type <Barney>, or <the man in the corner with bubbly liquid in his glass at whom I am looking>, or just <that man>. Attributing speaker-reference rather than semantic reference is assessing the substitutional commitments a tokening owes its allegiance to anaphorically rather than by type.

Having this interpretive or scorekeeping strategy available is useful for

reasons of charity. Wilma can make more of what Fred says come out *true* (according to her) by taking some of his remarks this way. Charity of this sort is necessary only where there is a relevant difference in perspective between audience and speaker—that is, where it makes a difference whether the commitments taken to be available as auxiliary hypotheses in drawing inferential consequences from a claim are those *undertaken* by the scorekeeper or those *attributed* to the one whose performances are being assessed. When such a difference in social perspective becomes explicit in ascriptions of doxastic commitment (the fundamental propositional attitude), it appears as the difference between ascriptions *de re* and ascriptions *de dicto*.

### **Appendix: Other Kinds of Anaphora— Paychecks, Donkeys, and Quantificational Antecedents**

Section III offers an account of the practical attitude a discursive scorekeeper must adopt in order to count as treating one tokening as anaphorically dependent on another in the most basic sense—that is, taking the substitutional commitments that determine the significance of the dependent to be inherited from those that determine the significance of its antecedent. This suffices to show how anaphoric relations can be introduced into or diagnosed in the simplified discursive practices described here. In actual natural languages, anaphora is an immensely complex phenomenon; many more sophisticated tropes have been built up around the asymmetric token-recurrence structures identified here as the fundamental anaphoric phenomenon. Discussion of these goes beyond the scope of the present project, but perhaps a few signposts are in order.

One important issue that is put to one side here concerns the thorny problem of paycheck cases. Understanding a sentence like

- (a) The man who gives his paycheck to his wife is wiser than  
the man who gives it to his mistress

requires treating the anaphor /it/<sub>a</sub> as replaceable by another tokening of the same type as its antecedent, /his paycheck/<sub>a</sub>, even though these tokenings will be governed by different SMSICs (have different referents). /His paycheck/<sub>a</sub> will not have the same referent as /it/<sub>a</sub>, any more than the two tokenings of <the man> do. So this sort of anaphora cannot be dealt with in terms of the establishment of token-recurrence structures. Nonetheless, it is clear enough how to understand this sort of lexically ‘lazy’ anaphora: /it/<sub>a</sub> is replaceable by another token of type <his paycheck>, and the antecedents of the two tokens of type <his> are the different, noncoreferential tokens of type <the man>. The hard question (which is important to linguists and for some projects in artificial intelligence) is not understanding the correct reading but

telling when that sort of reading is called for. It is how to tell when one ought to understand anaphoric dependence in terms of token-recurrence, and when it should be understood rather in terms of the sort of type-recurrence that paycheck cases demand. The explanatory task undertaken here is finished, however, when the differences between the two sorts of readings have been made clear.

Another large issue passed by here concerns anaphoric dependents whose antecedents are *quantificational* expressions. In the simplest cases, the interpretation of claims formed in this way follows from the general account of quantifiers; the anaphoric chains determine what count as substitution instances of particular and universal quantifiers, the significance of those substitution instances is determined by the token-recurrence model, and the significance of the quantificational claim is determined disjunctively or conjunctively by those instances. This is how anaphoric dependents on quantificational initiators should be understood when those dependents behave like the bound variables of the predicate calculus. As Evans has pointed out, however, not all anaphoric dependents on quantificational antecedents are happily assimilated to this model.<sup>102</sup> Thus on the most natural reading,

John bought *some donkeys*, and Harry vaccinated *them*

entails that Harry vaccinated *all* the donkeys John bought, whereas the bound reading

[some x: donkeys x] [John bought x & Harry vaccinated x]

requires only that there be some donkeys that John both bought and Harry vaccinated. In his excellent discussion, Neale points out further that

John bought *exactly two donkeys*, and Harry vaccinated *them*.

*Few politicians* came to the party, but *they* had a good time.

*Just one man* drank rum, and *he* was ill.

entail that John bought exactly two donkeys, few politicians came to the party, and just one man drank rum, respectively—consequences that are lost on the bound reading.<sup>103</sup> As he concludes: “The upshot . . . is that among pronouns anaphoric on quantifiers we need to distinguish between those that function as bound variables and those that do not.”<sup>104</sup>

The interpretation wanted has already been alluded to, in the discussion above of ‘definitization transforms’ in connection with Chastain’s treatment of anaphoric chains (see 5.4.2).<sup>105</sup> The quantificationally unbound anaphoric dependents of quantificational expressions in examples such as those above go proxy for definite descriptions formed from their antecedents. So the sentences above make the same claims as:

John bought *some donkeys*, and Harry vaccinated *those donkeys*.

*John bought exactly two donkeys, and Harry vaccinated those donkeys.*

*Few politicians came to the party, but those politicians had a good time.*

*Just one man drank rum, and the man who drank rum was ill.*

Here the quantificational antecedent determines the class of relevant substitution instances, and the significance of the clauses in which the anaphoric dependent appears is determined by that class. The bound cases differ just in that the clauses in which the anaphoric dependent appears also function to constrain the class of substitution instances with respect to which *both* clauses are evaluated. The difference between the bound and unbound cases accordingly corresponds to a difference in the order of application of the two operations of determining a class of substitution instances and making anaphoric connections. In the bound case, the anaphoric connections govern the inheritance of substitution-inferential significance by one clause from another within each quantificational substitution instance; in the unbound case, they govern rather the inheritance of a class of quantificational substitution instances by one clause from another.

Thus what might be called 'definitizing' anaphora—which governs the inferential significance of quantificationally unbound anaphoric dependents having quantificational antecedents—is another sort of sophisticated anaphora. It is distinct both from lazy, type-recurrent ('paycheck') anaphora and from the basic case of token-recurrent anaphora discussed in the body of the text. It is clear, at least in broad outlines, how such anaphora should be understood in the discursive scorekeeping idiom developed here. As before, the difficult task is formulating rules codifying when it is appropriate to adopt one sort of reading rather than another. As before, no stand is taken here on this difficult problem—and so none on Evans's suggestion that a pronoun anaphorically dependent on a quantificational expression behaves like a variable bound by it just in case the pronoun is c-commanded by the quantifier.<sup>106</sup> In the idiom suggested in the text, these issues are all taken to concern when it is appropriate to do the trick (construe the substitution-inferential significance of anaphoric dependence according to one model rather than another) rather than what it is to do the trick (keep score according to one reading rather than another), which is all that is of concern here.

Geach's original donkey sentence was

Any man who owns a donkey beats it.<sup>107</sup>

Here one does not want the definitized reading, for those who own two donkeys are being accused of beating both of them. The trouble is that apparently then 'a donkey' must be understood as expressing a particular quantifier relativized to the universal quantifier expressed by 'any man'. But

neither of the two ways of putting this relation in standard first-order quantificational language seems right. Unlike the original,

[every  $x$ : man  $x$ ] [some  $y$ : donkey  $y$ ] (Owns ( $x, y$ ) → Beats ( $x, y$ ))

is compatible with

Some man who owns a donkey does not beat it,

while

[every  $x$ : man  $x$ ] ([some  $y$ : donkey  $y$ ] (Owns ( $x, y$ ) → Beats ( $x, y$ )))

is syntactically incoherent, having the second quantifier, which occurs in the antecedent of a conditional, binding variables that occur in the consequent of that conditional.

This causes a problem, however, only for those concerned to provide a uniform way of mapping quantificational expression-types in natural languages onto operators in the first-order predicate calculus. Those not concerned with rules determining when it is appropriate to interpret tokens of the type  $\langle a K \rangle$  or  $\langle \text{some } K \rangle$  one way rather than another can rest content with understanding Geach's donkey sentence as having the inferential role of

[every  $x$ : man  $x$ ] [every  $y$ : donkey  $y$ ] (Owns ( $x, y$ ) → Beats ( $x, y$ )).

Further anaphoric dependents on these quantificational expressions may then act either as quantificationally bound anaphors or as definitized ones.

A final sort of example that deserves mention is Bach-Peters sentences, such as

A boy who was fooling her kissed a girl who loved him.

The pilot who shot at it hit the MiG that chased him.

The difficulty here is that the anaphoric chains cross; each dependent inherits its substitution-inferential role from an antecedent that inherits its role in turn from the original dependent. So these anaphoric circles do not settle what is to be counted as the anaphoric initiator. As has long been recognized, however, these surface forms are ambiguous; they have two nonequivalent readings, depending on which expression is treated as an initiator (which includes a dependent), and which as a dependent.<sup>108</sup> With definitization, quantificational cases such as the first example reduce to those involving definite descriptions, like the second, and the two readings of those are not far to seek. As Neale puts it:

If 'the pilot who shot at it' is given wider scope, 'him' is bound and 'it' is D-type:

[the x: pilot x & [the y: MiG y & y chased x] {x shot at y}] {[the y: MiG y & y chased x] {x hit y}}.

If ‘the MiG that chased him’ is given wider scope, ‘it’ is bound and ‘him’ is D-type:

[the y: MiG y & [the x: pilot x & x shot at y] {y chased x}] {[the x: pilot x & x shot at y] {x hit y}}.<sup>109</sup>

What Neale calls ‘D-type’ anaphoric dependents are those to be interpreted by definitization transforms of their antecedents. Once again, there is no special problem with interpreting each of these readings in discursive score-keeping terms, so long as care is taken to distinguish anaphora determining the inheritance of substitution-inferential significance within quantificational substitution instances from anaphora determining the inheritance of classes of quantificational substitution instances (in the case of definite descriptions, singletons). The defining symmetry of the Bach-Peters sentences ensures that in this case there is no residual problem of determining when one reading rather than the other is appropriate; the *only* task is making sense of the two readings.

In conclusion, although there are other sorts of anaphora in play in natural languages besides the one taken as fundamental in the discursive scorekeeping semantics, there are strategies available for making sense of them within the model as developed in the text.

# Ascribing Propositional Attitudes: The Social Route from Reasoning to Representing

---

Quot homines tot sententiae: suo quoque mos.  
(So many men, so many opinions; his own a law to each.)

TERENCE, Phormio

---

## I. REPRESENTATION AND *DE RE* ASCIPTION OF PROPOSITIONALLY CONTENTFUL COMMITMENTS

### 1. *Introduction*

At this point the deontic scorekeeping model has been developed sufficiently to show how discursive practice can confer *propositional* conceptual content on the repeatable *sentential* expressions used in performances manifesting doxastic and practical normative statuses, in virtue of their proper use being governed by *inferential* commitments. It has also been shown how discursive practice can further confer conceptual content on the repeatable *subsentential* components of those expressions, in virtue of their proper use being governed by *substitutional* commitments. Finally, it has been shown how discursive practice can confer conceptual contents on unrepeatable *tokenings*, whether sentential or subsentential, in virtue of their proper use being governed by *anaphoric* commitments. In this chapter these raw materials are assembled to address several important bits of unfinished business.

The chief task is to explain the *representational* dimension of thought and talk. Such an explanation is required to redeem a promissory note that was issued in Chapter 2, when commitment to an inferentialist order of semantic

explanation was undertaken. The complementary representationalist order of explanation, dominant since the seventeenth century, presents propositional contentfulness in representational terms from the outset—appealing to the notion of states of affairs involving represented objects as the worldly conditions of the truth of judgments. This approach is objectionable if it is pretended that an account in these terms gives one an independent grip on what is expressed by the declarative use of sentences—as though one could understand the notions of states of affairs or truth conditions in *advance* of understanding claiming or judging. The representational semantic tradition embodies an undeniable insight: whatever is propositionally contentful does necessarily have such a representational aspect; nothing that did not would be recognizable as expressing a proposition. The point of the inferentialist order of explanation is not to *object* to using representational locutions to talk about semantic content. Inferentialism must be understood instead as a strategy for *understanding* what is said by the use of such locutions. The objection is only to treating representational locutions as basic in the order of semantic explanation.

The inferentialist idea is to start with a preliminary understanding of conceptual content in terms of inferential articulation—to approach semantic contentfulness by means of the functional role claims play as premises and conclusions of inferences. Working out this idea along pragmatist lines focuses attention on inferring as a *doing*. An account is then sought of the practice of giving and asking for reasons, an account that can be expressed without the use of explicitly representational locutions. The aim is to be able to explain in deontic scorekeeping terms what is expressed by the use of representational vocabulary—what we are doing and saying when we talk about what we are talking *about*. Thus a criterion of adequacy for the inferentialist program is that it be possible to say without using specifically representational vocabulary what would count as introducing into discursive practice locutions that make *explicit* the *implicit* representational dimension of the semantic contents that claims acquire in virtue of their role in the game of giving and asking for reasons.

The claim developed and defended here is that representational locutions should be understood as making explicit certain features of *communicating* by claiming—the *interpersonal* giving and asking for reasons. The context within which concern with what is thought and talked *about* arises is the assessment of how the judgments of one individual can serve as reasons for another. The thesis is that the *representational* dimension of propositional content is conferred on thought and talk by the *social* dimension of the practice of giving and asking for reasons. Logicians typically think of inference as involving only relations among different propositional contents, not as also potentially involving relations among different interlocutors. However, discursive practice, the giving and asking for reasons, from which inferential relations are abstracted, involves both *intercontent* and *interpersonal*

dimensions. The claim is that the representational aspect of the propositional contents that play the inferential roles of premise and conclusion is to be understood in terms of the social dimension of communicating reasons and assessing the significance of reasons offered by others. The conceptual contents employed in *monological* reasoning, in which all the premises and conclusions are potential commitments of one individual, are parasitic on and intelligible *only* in terms of the conceptual contents conferred by *dialogical* reasoning, in which the issue of what follows from what essentially involves assessments from the different social perspectives of interlocutors with different background commitments. Representationally contentful claims arise in the social context of communication and only then are available to be employed in solitary cogitation.

Because, first, there is no propositional (and hence no conceptual) content without this representational dimension, and, second, that representational dimension is the expression of the social articulation of inferential practice, it follows that propositional and hence conceptual content can be conferred *only* by *social* practice. The social analysis of representational content will accordingly vindicate the claim of Chapter 3, as to the essentially social nature of conceptual content. It provides the promised argument for the claim that *discursive* practice must be understood as a fundamentally *social* practice. Appreciating the intimate connection between the representational dimension of conceptual content and the interpersonal or communicative dimension of the inferential practice of giving and asking for reasons brings into relief features of the inferential approach to semantics and the deontic scorekeeping approach to pragmatics that depend on their interaction.

On the semantic side, what appears is the social-perspectival character of propositional contents, and hence of conceptual contents generally. The semantic contents of discursive commitments, attitudes toward those commitments, and the linguistic performances that express those attitudes can in principle only be specified from the perspective provided by some repertoire of background commitments and attitudes; how it is correct to specify any particular content varies from one such repertoire to another. This relativity to doxastic point of view makes conceptual contents fundamentally unlike ordinary nondiscursive things.

On the pragmatic side, what appears is the social-perspectival character of the distinction between normative status and normative attitude—between what someone is *really* committed or entitled to and what anyone, including even the subject of those statuses, *takes* that individual to be committed or entitled to. Putting these semantic and pragmatic phenomena together yields an account of the *objectivity* of conceptual norms—a way of understanding how our scorekeeping practices can confer conceptual contents about whose proper applicability and consequences we can not only *each* but *all* be in error. It takes the form of an account of the structure that must be exhibited by the practices a community is interpreted as engaging in for those practices

thereby to be understood (both by those who are interpreted and by those who are interpreting them) as socially instituting implicit norms according to which the truth of claims and the correct use of concepts answer to how things objectively—rather than how things subjectively, or even intersubjectively—are taken to be.

These issues concerning the relations between pragmatics and semantics can be brought into simultaneous focus by considering the use of explicitating locutions that express the adoption of the sort of deontic attitudes studied by the pragmatic theory of scorekeeping in the form of the sort of propositional contents studied by the semantic theory of inferential articulation. The attribution of discursive commitments is an attitude that is implicit in deontic scorekeeping practices. It is something that scorekeepers *do*. The introduction of a sentential operator that functions as “*S* believes that . . .” or “*S* is committed to the claim that . . .” does in English makes it possible, not merely implicitly or in practice to *take* someone to be committed to a claim, but explicitly to *say* that someone is committed to a claim, and to which claim. The explicit is the claimable, what can be given as a reason and have reasons demanded for it; ascriptional locutions make implicit attributions explicit as the contents of claims.

The notion of explicitating locutions was introduced in Chapter 2 with the example of conditionals, which make *inferential* commitments explicit in the form of *assertional* commitments. Negation was then explained as playing the expressive role of explicitating incompatibilities. The paradigm of subsentential explicitating vocabulary was provided by identity locutions, which make *substitutional* commitments explicit. The expressive role of quantifiers and the operators that form definite descriptions then appeared as making explicit various kinds and combinations of substitutional commitments. All these are *semantically* explicitating locutions, which make explicit the *inferential* involvements of the nonlogical expressions they operate on. Traditional semantic vocabulary, paradigmatically ‘true’ and ‘refers’, has been shown to play the expressive role of making explicit the *anaphoric* relations that make it possible for unrepeatable tokenings to play indirectly inferential roles and so to have conceptual content. These too figure in the present scheme as semantically explicitating locutions (though the question of whether anaphora should be classed as a semantic or pragmatic phenomenon is equivocal at best according to traditional ways of dividing up those domains).

Vocabulary for ascribing propositional attitudes plays a corresponding explicitating expressive role, but it is *pragmatic* rather than semantic features of discursive practice that are made explicit by its use. For it is the adoption of a deontic *attitude* (the attribution of a deontic status) rather than the inferential articulation of a semantic content that is in the first instance made explicit by using these locutions. For this reason the traditional denomination of them as locutions for ascribing propositional *attitudes* is

somewhat misleading in the present context; it is propositionally contentful commitments—a kind of deontic *status*, not a kind of deontic *attitude*—that is in the first instance ascribed (that is, explicitly attributed) by the use of these tropes. An attitude is *expressed* by asserting an ascription, but it is not in general an attitude that is *ascribed*. Indeed, as will emerge, two attitudes are thereby expressed—for in asserting an ascription one *attributes* (usually to someone else) one propositionally contentful attitude, namely the one ascribed, and undertakes (oneself) another, namely the ascribing one.

However, ascriptional claims can not only be undertaken but also attributed, and even ascribed. Thus once the expressive resources they provide are available, it becomes possible to do what one could not do without them—to attribute not just statuses but attitudes. Only by considering what ascriptional locutions express is it possible to understand the relation between the *status* of *being committed* and the *attitude* of *acknowledging* a commitment—a distinction that has been systematically kept out of sight up to this point. Thus like the vocabulary that makes semantic (inferential) features of discourse propositionally explicit, the vocabulary that makes pragmatic (deontic) features of discourse propositionally explicit not only makes it possible to *say* things one could not *say* before but makes it possible to *do* things one could not *do* before, *by saying those things*. In virtue of this explicating expressive role, propositional-attitude-ascribing locutions deserve to count as *logical* vocabulary, albeit in an extended or generic sense that includes pragmatic as well as semantic species.

## 2. De Re Ascriptions Are the Fundamental Representational Locution of Natural Languages

In natural languages there are two sorts of explicitly representational vocabulary—locutions whose expressive role is to make explicit the representational dimension implicit in the inferential articulation in virtue of which discursive social practice is conceptually contentful. Central to the first are expressions like ‘true’ and ‘refers’. The way these work was considered in Chapter 5, where they were analyzed as anaphoric proform-forming operators. The sort of anaphoric dependence they display was then explained in Chapter 7 in terms of asymmetric token-recurrence and the consequent inheritance of substitution-inferential commitments. These expressions of natural language provide the basis from which technical semantic vocabulary is elaborated. Philosophers’ uses of terms such as ‘denotes’ and ‘satisfies’ in formal semantic investigations, and ‘represents’ in less formal epistemological theorizing, originate in and—except where a misunderstanding of their grammar has led to commitment to hypostatized relations and properties—can be explained in terms of these anaphoric idioms.

There is another, and in many respects more important, sort of representational locution in ordinary language. It comprises the idioms we typi-

cally use to express the intentional directedness of thought and talk—the fact that we think and talk *about* things and states of affairs. Words such as 'of' and 'about' play their characteristic intentional, semantic, or representational expressive role in virtue of the way they figure in *de re* ascriptions of propositional attitudes. It is these ascriptions that we use to say what we are thinking and talking of or about. 'Of' is used in many ways that have nothing to do with intentionality or representational aboutness, for instance its use to indicate possession, as in 'the pen of my aunt' or 'the positive square root of two'. 'About' is used in many ways that have nothing to do with the ofness of thought and talk, as in "The book is about six hundred pages long" or "It is about time for the meeting to start." The core locutions in terms of which the intentionality-expressing senses of these words must be distinguished are *de re* ascriptions, such as "Thomas Jefferson believed of meteorites that they did not exist"<sup>1</sup> or transforms of them, such as "Henry Adams's claim that the first postmaster general did not invent the lightning rod is about Benjamin Franklin."

Thus in order to identify vocabulary in alien languages that means what 'of' and 'about' do when used to ascribe intentionality and describe its content—or to introduce such vocabulary into expressively impoverished languages—one must be able to recognize expressions of *de re* ascriptions of propositional attitudes. The explanatory strategy pursued in this chapter is to address the issue of what is expressed by representational vocabulary, by showing how expressions must be used in order to be functioning as *de re* ascriptions of propositional attitudes. This provides an account of what is necessary to introduce, into an idiom that previously lacked them, representationally explicating idioms with the expressive capacity provided by terms like 'of' and 'about'.

The tradition distinguishes two readings of, or senses that can be associated with, propositional attitude ascriptions. The usual way of introducing the vocabulary is to say that ascriptions *de dicto* attribute belief in (commitment to) a *dictum* or saying, while ascriptions *de re* attribute belief about some *res* or thing.<sup>2</sup> The distinction is not specific to sentential operators such as 'believes', which express ascriptions of propositional attitude. Consider, to begin with, the assertion, in 1994, of the sentence:

The president of the United States will be black by the year 2000.

Read *de dicto*, this means that the *dictum* or sentence

The president of the United States is black

will be true by the year 2000. Read *de re*, it means that the *res* or thing, the president of the United States in 1994, namely Bill Clinton, will be black by the year 2000. The difference between the two readings is naturally understood as concerning the order of application of the two operations involved:

tracking through time, as demanded by the tense operator, and picking out an individual, as demanded by the definite description operator. The more plausible reading corresponds to first shifting the context of evaluation of the sentence forward from 1994 to 2000, and then considering what expressions are in that context intersubstitutable with the definite description 'the president of the United States'. The less plausible reading corresponds to determining those intersubstitutabilities first, then seeing what sentences involving those terms will be true in the year 2000.<sup>3</sup>

The present concern, however, is with how this distinction applies to ascriptions of propositional attitude. There are two ways to read:

Voltaire believed the man from whom Napoleon learned the most about the relations between war and diplomacy was a philosopher-prince.

One reading makes this a false claim, the other true. It is false that Voltaire believed the dictum

The man from whom Napoleon learned the most about the relations between war and diplomacy was a philosopher-prince.

If you had asked Voltaire, he would have denied that he endorsed this claim;<sup>4</sup> after all, when Voltaire died in 1778, Bonaparte was only nine. In contrast, Voltaire would have endorsed the claim:

Frederick the Great was a philosopher-prince.

This is a belief *about* the same *res*, since Frederick the Great was in fact the man from whom Napoleon learned the most about the relations between war and diplomacy. That is, Voltaire really did believe *of* or *about* the man from whom Napoleon learned the most about the relations between war and diplomacy (namely Frederick the Great) that *he* was a philosopher-prince.

In ordinary parlance the distinction between *de dicto* and *de re* readings is the source of systematic ambiguity. Sometimes, as in the case above, one of the readings involves a sufficiently implausible claim that it is easy to disambiguate. It is best, however, to regiment usage slightly in order to mark the distinction explicitly. This can be done with little strain to our ears by using 'that' and 'of' in a systematic way. Consider:

Henry Adams believed the inventor of the lightning rod did not invent the lightning rod.

Thus expressed, the ascription is ambiguous in the same way as the example above. It is quite unlikely that what is intended is the *de dicto*

Henry Adams believed *that* the inventor of the lightning rod did not invent the lightning rod.

Adams would presumably not have endorsed the *dictum* that follows the 'that'. It is entirely possible, however, that the *de re* claim

Henry Adams believed of the inventor of the lightning rod that  
he did not invent the lightning rod

is true. For since the inventor of the lightning rod is the inventor of bifocals (namely Benjamin Franklin), this latter claim could be true if Henry Adams had the belief that would be ascribed *de dicto* as

Henry Adams believed that the inventor of bifocals (or Benjamin Franklin) did not invent the lightning rod.

In the rest of this chapter the regimentation suggested by these expressions is followed: 'of' marks ascription *de re*, 'that' without 'of' marks ascription *de dicto*, and the absence of 'that' and 'of' marks an undifferentiated, potentially ambiguous ascription.

Quine emphasizes that the key grammatical difference between these two sorts of ascriptions concerns the propriety of *substitution* for singular terms occurring in them.<sup>5</sup> Expressions occurring in the *de re* portion of an ascription—within the scope of the 'of' operator in the regimented versions—are in his terminology *referentially transparent*—that is, coreferential terms can be intersubstituted *salva veritate*, that is, without changing the truth-value of the whole ascription. By contrast, such substitution in the *de dicto* portion of an ascription—within the scope of the 'that' operator in the regimented versions—may well change the truth-value of the whole ascription. (These formulations carry over directly into the present scheme in the by now familiar way—in terms of substitutions that do and do not preserve doxastic commitment.) Syntactically, *de re* ascriptions may be thought of as formed from *de dicto* ones by *exporting* a singular term from within the 'that' clause, prefacing it with 'of', and putting a pronoun (or other anaphoric dependent) in the original position. Thus the *de dicto* form

S believes that  $\Phi(t)$

becomes the *de re*

S believes of t that  $\Phi(it)$ .

The important point is that, as the regimentation reminds us, it is *de re* propositional-attitude-ascribing locutions that we use in everyday life to express what we are talking and thinking of or *about*. This expressive function dictates the proprieties of substitution that govern expressions occurring in the scope of the 'of'. Since those expressions serve to specify *what* is represented by a belief rather than *how* it is represented, any singular term that picks out the right object is all right; one specification is as good as another. By contrast, expressions that occur within the scope of the 'that'

(either in a pure *de dicto* ascription or in the *de dicto* portion of a *de re* ascription) serve to specify how things are represented by the one to whom the belief is ascribed. Thus only expressions the believer acknowledges being committed to intersubstitute, and so treats as ways of recognizing the same object again, are intersubstitutable in these contexts, *salva veritate* (that is, preserving doxastic commitment).

In perspicuous, regimented-but-recognizable English, what marks this perspectival distinction within the content-specifications of ascriptions is the occurrence of the words 'of' and 'that'. Thought of in this way, the distinction between *de dicto* and *de re* should not be understood to distinguish two kinds of *belief* or even belief-contents, but two kinds of *ascription*<sup>6</sup>—in particular two different *styles* in which the *content* of the commitment ascribed can be *specified*. (To be more precise, this is how one ought to think of what will be distinguished as '*epistemically weak de re* ascriptions'. *Epistemically strong de re* ascriptions, discussed in Section V, do ascribe a special sort of attitude.)

One way of trying to understand the representational dimension of propositional content is accordingly to ask what is expressed by this fundamental sort of representational locution. What are we *doing* when we make claims about what someone is talking or thinking *about*? How must vocabulary be used in order for it to deserve to count as expressing such *de re* ascriptions? Answering that question in a way that does not itself employ representational vocabulary in specifying that use is then a way of coming to understand representational relations in nonrepresentational terms.

The basic strategy, as indicated above, is to show how the use of this paradigmatic representational locution expresses differences in *social* perspective. The raw materials needed for a deontic scorekeeping account of *de re* ascriptions are already onboard: the attitudes of acknowledging and attributing doxastic commitments, substitutional commitments, and anaphoric commitments. Ascribing a doxastic commitment is explicitly attributing it—that is, attributing it by acknowledging another doxastic commitment. The substitutional commitments that govern the expressions used to specify the content of the commitment ascribed can similarly either be attributed to the one to whom the doxastic commitment is ascribed or be undertaken by the one ascribing it; that social difference of deontic attitude turns out to be what determines whether the ascription is *de dicto* or *de re*. Communication requires that scorekeepers be able to move back and forth between the significance tokenings have as governed by the commitments they themselves acknowledge, on the one hand, and by the commitments they take the speaker to acknowledge, on the other. That is why the expressive power of interpersonal anaphoric recurrence chains is important for securing the possibility of communication. It also turns out that what is expressed by *de re* ascriptions is just the understanding of what is represented by a speaker's claims or beliefs—what they are *about*.

### 3. Deontic Scorekeeping Account of What Is Expressed by *De Dicto* and *De Re* Ascriptions

These points emerge most clearly from looking at the expressive role of ascriptions, for that makes it possible to see the abstract need for two sorts of ways in which expressions could function in specifying the content of an ascribed commitment. Ascriptions are propositionally explicit attributions. Ascriptional locutions such as "... claims that ..." or "... believes that ..." permit one to say that one is attributing a commitment with a specified content to a specified individual, something that otherwise one could only do. In virtue of playing this expressive role, ascribings essentially involve two different deontic attitudes, to commitments with two different contents. Making an ascription involves doing two different things. Ascribing is *attributing* one commitment (to another), while *undertaking* (acknowledging) a different commitment (oneself).

It follows that expressions occurring in the ascription must somehow specify both the content of the commitment attributed by the ascriber and (thereby) the content of the commitment undertaken by the ascriber. That a particular expression occurs in the specification of the content of the attributed commitment accordingly can have a significance (affect the deontic score) for two different commitments. This makes possible a special sort of ambiguity. It may on occasion be important to be able to resolve such ambiguity; natural languages provide mechanisms for doing so by making explicit the sort of significance the occurrence of a particular expression ought to be understood to have. The expressions occurring in an ascription must specify three sorts of information, which differ in how their occurrence bears on the commitments being attributed and undertaken by the assertion of the ascription. The individual who is the target of the attribution, to whom a commitment is being ascribed, must somehow be indicated. The sort of commitment being attributed—for instance 'claims', 'believes', 'intends', or 'prefers'—must also be indicated. Finally, the content of the ascribed commitment must be specified.

The content-specifying expressions themselves can play two quite different roles, however, and that is where the difference between *de dicto* and *de re* ascriptions arises. Propositional contents, such as those characterizing both the ascribing and the ascribed commitment, are essentially inferentially articulated. But what else a commitment with a particular content commits one to, the committive-inferential consequences of adopting a commitment with that particular content, depends on its deontic context, on what concomitant commitments are available as auxiliary hypotheses or collateral premises. So the occurrence of an expression in the specification of the content of an ascribed commitment might have one inferential significance if evaluated with respect to the collateral commitments of the one to whom the commitment is attributed, and might have quite another if evaluated

with respect to the collateral commitments of the one undertaking the ascriptional commitment. A perspicuous regimentation of ascriptional locutions must notationally encode and display the distinction between expressions (or features of expressions) playing these various roles. Thus the content-specifying expressions should be marked somehow to indicate which sort of significance they are to be understood to have.

In specifying the content of the claim that is attributed by an ascription, a question can arise as to whom the ascriber takes to be responsible for this being a way of *saying* (that is, making explicit) what is believed, the content of the commitment. Consider the sly prosecutor, who characterizes his opponent's claim by saying:

The defense attorney believes a pathological liar is a trustworthy witness.

The defense attorney may hotly contest this characterization:

Not so; what I believe is *that* the man who just testified is a trustworthy witness.

To which the prosecutor might reply:

Exactly, and I have presented evidence that ought to convince anyone *that* the man who just testified is a pathological liar.

If the prosecutor were being fastidious in characterizing the other's claim, he would make it clear who is responsible for what: the defense attorney claims that a certain man is a trustworthy witness, and the prosecutor claims that that man is a pathological liar. The disagreement is about whether this man is a liar, not about whether liars make trustworthy witnesses.

Using the regimentation suggested above, the way to make this explicit is with a *de re* specification of the content of the belief ascribed. What the prosecutor *ought* to say (matters of courtroom strategy aside) is:

The defense attorney claims of a pathological liar that he is a trustworthy witness.

This way of putting things makes explicit the division of responsibility involved in the ascription. That someone is a trustworthy witness is part of the commitment that is *attributed* by the ascriber; that that individual is in fact a pathological liar is part of the commitment that is *undertaken* by the ascriber.<sup>7</sup>

Ascription always involves attributing one *doxastic* commitment and, since ascriptions are themselves claims or judgments, undertaking another. The suggestion is that the expressive function of *de re* ascriptions of propositional attitude is to make explicit which aspects of what is said express substitutional commitments that are being *attributed* and which express substitutional commitments that are *undertaken*. The part of the content

specification that appears within the *de dicto* 'that' clause is limited to what, according to the ascriber, the one to whom the commitment is ascribed would (or in a strong sense should) *acknowledge* as an expression of what that individual is committed to. The part of the content specification that appears within the scope of the *de re* 'of' includes what, according to the *ascriber* of the commitment, but not necessarily according to the one to whom it is ascribed, is acknowledged as an expression of what the target of the ascription is committed to. (This is what the target should, according to the ascriber, *acknowledge* only in a much weaker sense of 'should'.) What else someone is committed to by a claim that would be expressed in one way is a matter of the substitution-inferential commitments taken to govern the expressions that occur in the sentences the one whose acknowledgments are in question is disposed to assert or otherwise endorse in practice. Thus the marking of portions of the content-specification of a propositional attitude ascription into *de dicto* and *de re* portions makes explicit the essential deontic scorekeeping distinction of *social* perspective between inferences that are underwritten by substitutional commitments attributed and inferences underwritten by substitutional commitments that are undertaken.

So the difference expressed by segregating the content specification of a propositional attitude ascription into distinct *de re* and *de dicto* regions (marked in the regimentation by 'of' and 'that') can be thought of in terms of *inferential* and *substitutional* commitments. Propositional (= assertible) contents are inferentially articulated. Grasping such a content is being able to distinguish in practice what should follow from endorsing it and what such endorsement should follow from. But, as was pointed out above, the consequences of endorsing a given claim depend on what other commitments are available to be employed as auxiliary hypotheses in the inference. The ascriber of a doxastic commitment has got two different perspectives available from which to draw those auxiliary hypotheses in specifying the content of the commitment being ascribed: that of the one to whom it is *ascribed* and that of the one *ascribing* it. Where the specification of the content depends only on auxiliary premises that, according to the ascriber, the target of the ascription *acknowledges* being committed to, it is put in *de dicto* position, within the 'that' clause. Where the specification of the content depends on auxiliary premises that the *ascriber* endorses, but the target of the ascription may not, it is put in *de re* position.

More particularly, the use of expressions as singular terms is governed by *substitution-inferential* commitments. The rule for determining the scorekeeping significance and so the expressive function of *de re* ascriptions being proposed is then the following. Suppose that according to *A*'s scorekeeping on commitments, *B* acknowledges commitment to the claim  $\Phi[t]$ . Then *A* can make this attribution of commitment explicit in the form of a claim by saying

B claims *that*  $\Phi(t)$ .

If in addition A acknowledges commitment to the identity  $t = t'$ , then whether or not A takes it that B would acknowledge that commitment, A can also characterize the content of the commitment ascribed to B by saying

B claims of  $t'$  that  $\Phi(it)$ .<sup>8</sup>

Again, the question is whose substitutional commitments one is permitted to appeal to in specifying the consequences someone is committed to by acknowledging a particular doxastic commitment. Where in characterizing the commitment the ascriber has drawn out those consequences employing only commitments the ascriptional target would acknowledge, the content specification is *de dicto*. Where the ascriber has employed substitutional commitments the ascriber, but perhaps not the target, endorses, the content specification is *de re*.

One counts as having *undertaken* a commitment wherever it is appropriate for others to *attribute* it. One may be committed to some claims by default, but there are two ways in which what commitments one undertakes can depend on what one *does* or is disposed to do. First, one may *acknowledge* the commitment, paradigmatically by being disposed to *avow* it by an overt assertion. Or one may acknowledge it by employing it as a premise in one's theoretical or practical reasoning. This last includes being disposed to *act on it practically*—taking account of it as a premise in the practical reasoning that stands behind one's intentional actions. Second, one may undertake the commitment *consequentially*, that is, as a conclusion one is committed to as an inferential consequence entailed by what one *does* acknowledge. Commitments one acknowledges and commitments one acquires consequentially in virtue of those acknowledgments correspond to two ways of talking about beliefs: one according to which one believes only what one takes oneself to believe, and the other according to which one believes willy nilly what one's beliefs commit one to.

According to one usage, I believe only what I think I believe, what I take myself to believe. I do not believe things behind my back; my sincere avowals are authoritative with respect to what I believe. According to another usage, however, I believe the consequences of my beliefs, whether I think I do or not. For my acknowledged beliefs can commit me to more than I acknowledge; so I can end up with beliefs I do not know I have. Also, my actions, perhaps together with avowed preferences, may commit me to certain claims.

There can be tensions between these two ways of talking about beliefs: a narrow one tied to empirical dispositions to avow, the other more normative and expansive, closing beliefs under a consequence relation not limited by the believer's acknowledgment of it. Indeed, the fact that people often move

back and forth between talk of belief in the empirical sense, which does not involve inferential closure, and talk of belief in the logical or ideal sense, which does, is one of the reasons to prefer talk of commitment to talk of belief—or simply not to believe in beliefs. *De dicto* and *de re* ascriptions correspond to two socially distinct perspectives from which the consequences of doxastic commitments can be extracted. Since the propositional content of a commitment depends on what it is a consequence of and what is a consequence of it, they accordingly also correspond to two different ways of specifying such contents.

## II. INTERPRETATION, COMMUNICATION, AND *DE RE* ASCRIBPTIONS

### 1. *Interpretation in Wittgenstein's Sense*

The representational dimension of propositional content (and of conceptual content more broadly) is made explicit—that is, expressed in the form of propositional content—by the use of *de re* ascriptions of discursive commitments. The use of such ascriptions turns essentially on *social* distinctions of doxastic perspective between the ascriber and the one to whom a commitment is ascribed—that is, between the deontic repertoire associated with the one by whom the ascriptional propositional commitment is *undertaken* and the one to whom a propositional commitment is thereby explicitly *attributed*. The social distinction between the fundamental deontic attitudes of undertaking and attributing is essential to the institution of deontic statuses and the conferral of propositional contents. This is, as was pointed out already in Chapter 1, an *I-thou* sociality rather than an *I-we* sociality. Its basic building block is the relation between an audience that is attributing commitments and thereby keeping score and a speaker who is undertaking commitments, on whom score is being kept. The notion of a discursive *community*—a *we*—is to be built up out of these *communicating* components.

Deontic scorekeeping is the form of understanding involved in communication. It is a kind of interpreting. But it is implicit, practical interpretation, not explicit theoretical hypothesis formation. It is presupposed by the capacity so much as to entertain the claims that would express a hypothesis, evidence for a hypothesis, or conclusions from a hypothesis. Looking at the sort of scorekeeping (and so interpretation) that is made explicit in the use of *de re* ascriptions of propositional attitude—the kind of understanding communication consists in—will make clear how the representational dimension of conceptual content that those ascriptions express is rooted in the social context of communication.

Thinking of ordinary intralinguistic understanding as essentially involving interpretation is objectionable if interpretation is thought of on the model of explicit hypothesis formation.<sup>9</sup> Linguistic understanding depends on inter-

pretation in this sense only in extraordinary situations—where different languages are involved, or where ordinary communication has broken down. Recall in this connection the rehearsal in Chapter 1 of Wittgenstein's argument that norms that are expressed explicitly as rules must be based on norms that are embodied implicitly as proprieties of practice. His pragmatist point is that the distinction between following a rule correctly and following it incorrectly must not be understood in every case as determined by the application of another rule—which could be thought of as an explicit theory or hypothesis concerning the first rule. The normative cannot be understood as rules all the way down. Since he is determined to “restrict the term ‘interpretation’ [Deutung] to the substitution of one expression of the rule for another,”<sup>10</sup> in his terminology it is accordingly a radical mistake to think of our ground-level practical mastery of linguistic proprieties as consisting exclusively in a capacity to interpret.

This argument is, as was indicated in Chapter 1, one of the fundamental insights from which the present approach proceeds. Yet the deontic score-keeping account that has been developed here identifies the inferentially articulated sort of understanding characteristic of specifically discursive practice as a kind of interpretation. This sort of interpretation includes, but is not exhausted by, Wittgensteinian interpretation—substituting one expression of a rule for another. Since what matters for the regress-of-rules argument is the propositional explicitness of rules, substitution of one expression of a *claim* for another may be thought of as interpretation in the relevant sense, regardless of whether the claims in question are functioning as rules.<sup>11</sup> The question to be addressed here concerns how much of ordinary skillful intralinguistic practice depends on the capacity to substitute one explicit expression of a claim for another. The answer is that a great deal of ordinary communication between individuals who share a language requires interpretation in the broadly Wittgensteinian sense just indicated. This point will be made by considering four linguistic phenomena, all of which involve cases where we can often just understand what others are saying (without theorizing about it), but where just understanding them crucially depends on being able to substitute one expression of a remark for another.

The general structure that this argument depends on is the fact that a sentence in one person's mouth does not typically have the same significance as that same sentence emerging from another person's mouth, even where there is as much sharing of the language and as much mutual understanding as one likes. The fundamental reason is the kind of things claims and concepts are. As was emphasized above, the inferential articulation of conceptual contents is such that what someone becomes committed to by uttering a certain expression can be assessed only against a background of collateral commitments available as auxiliary hypotheses that can be brought in as other premises in drawing the inference. Even where people share a language (and so their concepts), which is the standard case of communication, there

will still be some disagreements, some differences in the commitments that people have undertaken. We each embody different perceptual and practical perspectives and so will never have exactly the same doxastic and practical commitments. In any case, if two interlocutors did (per impossibile) have exactly the same beliefs and desires, communication would be superfluous; so that case need not be considered here. As long as there are differences in the collateral set of commitments with respect to which the content of the claim expressed by a sentence needs to be assessed, the sentence in one mouth means something different from what that same sentence means in another mouth. So even in the smooth untroubled cases of communication, if you want to understand what I say, you have to be able to associate with it a sentence that in your mouth expresses the claim that the sentence I uttered expresses in mine. For your understanding it (your knowing what I have committed myself to) involves your being able to trace out the inferences that claim is involved in, the evidential significance of what I have said, in order to know what I am committing myself to. This means knowing how it could function as evidence for you, as well as for me, what claims its endorsement would preclude you, as well as me, from being entitled to, and so on. Apart from that capacity, you cannot extract information from what I say and cannot be said to understand it.

## 2. Four Linguistic Phenomena That Involve Interpretation

There are four phenomena in linguistic practice where it is clear this sort of capacity to interpret is of the essence of smooth conversational coping and practical grasp of the meanings expressed by the utterances of others. The first is the use of *personal pronouns*. When someone else says, "I'm talking," for a member of the audience to draw inferences from that remark and check out what would be evidence for it and what it would be evidence for, that auditor must be able (no doubt without thinking about it and smoothly, as a matter of unconscious skill) to substitute 'you' or 'he' or 'John' for 'I'. For that last expression in the mouth of the auditor does not mean what it means in the mouth of the speaker. Those tokenings are not corecurrent and so are governed (at least counterfactually) by different substitutional commitments. Understanding the expression requires being able to make that substitution of one expression of the claim for the other. The use of personal pronouns is a fundamental part of our shared language. So here is a simple case where sharing a language, being able to understand someone, consists in part in being able to make this substitution—that is to *interpret*, in Wittgenstein's sense. For the words in the one mouth do not have the same significance that they would in the other mouth.

A similar case is that of *demonstratives*. If someone says, "This is blue," one to whom the remark is addressed *might* (if in a favorable position) be able to use another token of the same type to make the same claim. But

typically, members of the audience will not be in a position to do that. If the speaker uses a demonstrative, the audience will have to use a pronoun, a description, or some other demonstrative (even if it is only 'that'). There is no requirement that in order to understand the original remark the audience must be in a position to make demonstrative reference to the same object at all. If the speaker used some demonstratives yesterday and someone in the audience wants today to rehearse the inferences in the vicinity of the commitments undertaken thereby, exercising an understanding of what was said, that auditor is not typically going to be in a position to redemonstrate those things. Other expressions must be available to be used to pick them out.

As was pointed out in Chapter 7, it has to be possible to pick up the demonstrative reference with a pronoun. When someone says, "*This* is blue," the audience must be able to say something like, "No *it* isn't either"—where 'it' is a pronoun that is anaphorically dependent on the speaker's tokening. Such anaphoric dependents continue a recurrence tree that gives everybody a repeatable way of expressing the content the speaker expressed, even those who are not in a position to make the same demonstration. There cannot be demonstratives without anaphora; such 'bare' or inaccessible demonstratives would be of no use in communication (nor, indeed, even for the thought of an isolated individual). The audience, in order to make use of the speaker's demonstratives, has to be able to pick them up with a pronoun. That is the only form in which they are repeatable. The current point is that the one understanding the speaker's remark must be able to *substitute* a *different* expression for the one the speaker used. Interlocutors must be able to perform this sort of interpretation—in the Wittgensteinian substitutional sense—in order to understand even in the smoothest intralinguistic conversation.

Those two examples, personal pronouns and demonstratives, are very straightforward and ordinary, even trivial cases where intralinguistic interpersonal understanding requires the ability to interpret in Wittgenstein's substitutional sense. The next two cases are more substantive. The important thing about the first two is that for both personal pronouns and demonstratives, understanding them *consists*, at least in part, in being able to substitute other expressions for them. The other two cases to be considered are *speaker's reference* and being able to give *de re* specifications of the content of someone's claim (of the expression of the belief expressed).

These are intimately linked. Both of them arise precisely because there are differences in repertoires of collateral commitments or background assumptions in any communicational situation. Conversation has the significance of communication only where the commitments of speakers and audience differ. It is because of this difference in background commitments that it is possible for one interlocutor to become entitled to a new commitment by listening to what someone else says—or to challenge the speaker's entitlement to the commitment that speech act expresses. This is the communica-

tive point of the assertional practice in virtue of which we are able to entertain propositional contents at all. But often in assigning or even just hearing the content of what someone says, it is necessary to take account of those differences.

The previous chapter discussed the classic sort of speaker's reference case, which arises when the one overhearing the remark "The man in the corner drinking champagne is a good husband" both *attributes* to the speaker a collateral commitment that would be undertaken by asserting "John is in the corner and is drinking champagne" and *undertakes* a collateral commitment that would be expressed by a sentence such as "John is in the corner and is drinking ginger ale." In order to understand the commitment the speaker intended to be understood as undertaking—and so potentially to extract information from what was said—the audience must be able to substitute for the defective description 'the man in the corner drinking champagne' some expression that, according to the substitutional commitments endorsed by the audience, succeeds in picking out the same individual. That individual is the one the speaker is talking about. (Though of course there is also a sense in which willy-nilly the speaker may have made a claim about someone else, to whom his expression "semantically" refers.) The expression the audience substitutes could be an explicit, anaphorically indirect definite description, such as 'the man the speaker is talking about'; or the same effect could be achieved implicitly by using a pronoun such as 'he', if one can arrange for it to be understood as having as its antecedent some other tokening the speaker *could* have used to indicate the same man, one that is not subject to the same conflict of doxastic perspective, such as 'the man in the corner', 'John', or 'the man with bubbly liquid in his glass'; or one might use one of those expressions directly. An interpreter's capacity to understand the remark and extract information across the gap between the commitments undertaken and those attributed depends on that scorekeeper being able smoothly, and perhaps completely unconsciously,<sup>12</sup> to substitute an expression that in the scorekeeper's mouth would undertake the commitment being attributed to the speaker.

The substitutional commitments that are *implicit* in the sort of interpretive competence required to be able to take someone's remark as having its significance determined in the manner characteristic of speaker's reference are made *explicit* in *de re* ascriptions of commitment. Indeed, the expressive capacity provided by ascriptions of propositional attitude that specify the content of the attributed commitment in the *de re* way is precisely what is needed to *say* what the difference is between interpreting a remark according to its speaker's reference and interpreting it according to its semantic reference. Thus, suppose there is someone, Bob, of whom the speaker is unaware, who is also in the corner and actually is drinking champagne, unlike ginger-ale-guzzling John, whom the speaker takes to be drinking champagne. The two interpretations then correspond to taking the speaker to have said (with-

out meaning to) of Bob that *he* is a good husband and taking the speaker to have said (using a false description) of John that *he* is a good husband. The difference between these two *de re* specifications of the content expressed is a matter of which substitutional commitments undertaken by the ascriber are taken to govern the speaker's tokening 'the man in the corner drinking champagne'.

In the first case, that tokening is treated as corecurrent with lexically cotypical tokenings in the ascriber's mouth. In the second, as indicated earlier (see 7.5.6 above), it is treated as anaphorically dependent on other tokenings, perhaps demonstrative ones, that the speaker could (according to the interpreting ascriber) have used equally well to express the commitment being acknowledged. This latter case involves substituting expressions of other types for the expression used, and it is these expressions that occur in the *de re* specification of the content. Indeed, (weak) *de re* ascriptions are formed by substituting, for locutions the target of the ascription might use in expressing the content of the commitment, locutions that the *ascriber* is committed to being intersubstitutable with them. Thus being able to offer *de re* characterizations of the contents of the commitments of other interlocutors requires interpretation in Wittgenstein's substitutional sense.

### 3. *Communication, Truth Conditions, and De Re Specifications of Propositional Content*

Unless one has this substitutional interpretive capacity, which is expressed explicitly in *de re* specifications of the contents of ascribed commitments, one would not be able to understand what others were saying—even in languages that lacked the personal pronouns, demonstratives, and the other officially token-reflexive structures considered above. The cognitive uptake or grasp of content that is the measure of the success of communication consists in auditors being able to move between their own doxastic perspective and that of the speaker in just the way expressed by *de re* ascriptions—that is, by inferentially exploring the significance the *propositional commitment attributed* to the speaker has in the context of the *substitutional commitments undertaken* by the audience who attributes it. There is no communication apart from this sort of interpretation. Being able to understand what others are saying, in the sense that makes their remarks available for use as premises in one's own inferences, depends precisely on being able to specify those contents in *de re*, and not merely *de dicto*, terms.

If the only way I can specify the content of the shaman's belief is by a *de dicto* ascription

He believes *that* drinking the liquor distilled from the bark of that kind of tree will prevent malaria,

I may not be in a position to assess the truth of his claim. It is otherwise if I can specify that content in the *de re* ascription

He believes of quinine that malaria can be prevented by drinking it,

for ‘quinine’ is a term with rich inferential connections to others I know how to employ. If he says,

The seventh god graces us with his presence,

I may not know what to make of his remark. Clearly he will take it to have consequences that I could not endorse; so nothing in my mouth could *mean* just what his remark does. But if I come to believe (perhaps by being told) that the seventh god is the sun, and that his grace is sunshine, then I can specify the content of his report in a more useful form:

He claims of the sun that it is shining,

or both *of the sun* and *of shining*, that one is doing the other, or indeed as saying *of the claim* that the sun is shining *that* it is true. These are forms I can extract *information* from, that is, can use as premises that I can reason with. Again, suppose a student claims that

The largest number that is not the sum of the squares of distinct primes is the sum of at most 27 primes.

The student may have no idea what number that might be, or may falsely believe it to be extremely large. But if I know that

17,163 is the largest number that is not the sum of the squares of distinct primes,

then I can characterize the content of his claim in *de re* form as:

The student claims of 17,163 that it is the sum of at most 27 primes,

and I can go on to draw inferences from that claim, to assess its plausibility in the light of the rest of my beliefs. (It is true, but only because *all* integers are the sum of at most 27 primes.) Identifying what is being talked *about*, what is represented by it, permits me to extract information across a doxastic gap. Thus, being able to substitute *de re* specifications of the content of a claim for *de dicto* ones is something anyone must be able to do, at least implicitly, in order to understand people whose beliefs are different (not massively or radically different, but just different on particular points).

*De re* content specifications not only indicate what a claim represents or is about, they are the form in which the *truth conditions* of claims are expressed. The sense in which they present a claim’s representational content (the information it makes available to another interlocutor) is that they

express—from the point of view of the interpreter attributing the commitment—what would be required for the claim in question to be *true*. Recall the account (in Chapters 3, 4, and 5) of the scorekeeping practices that underlie truth assessments. The central context in which such assessments classically arise is attributions of *knowledge*, a normative status integral to the social practice of making claims and judgments. According to the traditional account, knowledge is justified true belief. Transposed into a specification of a normative status something could be taken to have by interlocutors who are keeping score of each other's commitments and entitlements, this account requires that in order for it to be *knowledge* that a scorekeeper takes another to have, that scorekeeper must adopt three sorts of practical attitude:

First, the scorekeeper must *attribute* an inferentially articulated, hence propositionally contentful, *commitment*.

This corresponds to the *belief* condition on knowledge.

Second, the scorekeeper must *attribute* a (perhaps, but not necessarily inferentially inherited, but necessarily heritable) *entitlement* to that commitment.

This corresponds to the *justification* condition on knowledge. What is it that then corresponds to the *truth* condition on knowledge? For the scorekeeper to take the attributed claim to be true is just for the scorekeeper to endorse that claim. That is:

Third, the scorekeeper must *undertake* the same propositional commitment attributed to the candidate knower.

Undertaking a commitment is adopting a certain *normative stance* with respect to a claim; it is not attributing a property to it. The classical metaphysics of truth properties misconstrues what one is doing in endorsing the claim as *describing* it in a special way. It confuses *attributing* and *undertaking* or *acknowledging* commitments, the two fundamental social flavors of deontic practical attitudes that institute normative statuses. It does so by assimilating the third condition on treating someone as having knowledge to the first two. Properly understanding truth talk in fact requires understanding just this difference of social perspective: between *attributing* a normative status to another and *undertaking* or adopting it oneself.

This is just the distinction that underlies the use of *de re* ascriptions. As the regimented form considered here emphasizes, they mark overtly the distinction between the *doxastic* commitment that is *attributed* and the *substitutional* commitments that are *undertaken* by the attributor. More specifically, they indicate whether responsibility for the words used being an appropriate expression of the content of the claim ascribed is attributed or undertaken by the ascriber. The connection between the account of “... is true” as an anaphoric prosentence-forming operator and the scorekeeping

account of truth assessments becomes evident in *de re* ascriptions in which the entire content-specifying clause is exported to *de re* position:

Senator McCarthy believed of the first sentence of the Communist Manifesto that *it was true*.

The first sentence of the Communist Manifesto says that the specter of Communism is haunting Europe. That is undoubtedly something McCarthy believed, but it is quite unlikely that he had ever read the Communist Manifesto, and also quite likely that if asked whether he believed any of the claims of that manifesto, he would have denied that he did. Responsibility for specifying the content of his belief in this way is accordingly properly undertaken by the ascriber of the belief, rather than attributed along with the belief. The regimented *de re* form requires then that the whole content-specifying expression be exported to the scope of the 'of', and that an ascription-structural anaphoric dependent of it be left in its place inside the scope of the 'that'. Since what is exported is a sentence nominalization, such an anaphoric dependent is a prosentence. No truth assessment is being offered by such a use of "... is true," for that prosentence appears embedded here, in a context in which it functions only to express the undertaking of a *substitutional* commitment; the doxastic commitment is attributed, not undertaken. Nonetheless, in this expressive role the prosentence-forming operator is used to specify what must be the case for the attributed belief to have been true (according to the attributor). It would have been true just in case the first sentence of the Communist Manifesto were true; that is, according to a well-informed ascriber, just in case the specter of Communism *were* haunting Europe. The ascription-structural anaphoric connection between the exported expression (in *de re* or undertaking position) and its anaphoric dependent (in *de dicto* or attributing position) mirrors within an ascription in one mouth the interpersonal anaphoric connections among expressions in different mouths that make possible the uptake of speaker's reference and the extraction of representational, truth-conditional content across doxastic boundaries (as discussed in Chapter 7).

The role of *de re* specifications of propositional content in expressing the truth conditions of ascribed commitments is not restricted to cases where the entire content-specification is exported to *de re* position. For ordinary *de re* ascriptions, in which singular terms are exported, specify what a belief (or the sentence expressing it) *represents* in the sense of what it must be *true of* if it is to be true at all. A scorekeeper who takes it that the inventor of bifocals is the inventor of the lightning rod will take it that believing that the inventor of bifocals spoke French is believing of the inventor of the lightning rod that *he* spoke French. If it is true that the inventor of bifocals spoke French, then, according to that scorekeeper, it is *true of* the inventor of the lightning rod that *he* spoke French. In saying which individual a belief represents or is about, *de re* content specifications indicate the individual

whose properties must be consulted in order to determine the *truth* of the belief (all from the point of view of the ascriber). Thus a scorekeeper who in addition endorsed the substitutional commitment made explicit by the (provincially tempting and occasionally enunciated) claim that French is the only language in which words appear in exactly the order in which the corresponding ideas are arranged in thought could specify the truth conditions of the belief that the inventor of bifocals spoke French by saying that it is true just in case it is true of the inventor of the lightning rod and of the only language in which words appear in exactly the order in which the corresponding ideas are arranged in thought that the first spoke the second. The point is that scorekeepers must use the auxiliary hypotheses provided by their own commitments in assessing the truth of the beliefs they attribute or entertain, for taking-true a claim is just endorsing it. *De re* specifications of their contents are just those that employ the substitutional commitments undertaken by the scorekeeper or interpreter, rather than those attributed to someone else, and so are the specifications that present propositional contents in a form apt for assessments of *truth*. Thus the *information* a claim potentially communicates, its *representational* content in the sense captured by its truth conditions, is what is expressed by *de re* specifications of propositional content.

Extracting information from the remarks of others requires being able to do the sort of substitutional interpreting that is expressed explicitly in offering *de re* characterizations of the contents of their beliefs—that is, to be able to tell what their beliefs would be true of if they were true. It is to grasp the *representational* content of their claims. The most important lesson of the deontic scorekeeping account of the use of *de re* ascriptions is that doing this is just mastering the *social* dimension of their inferential articulation—the way in which commitments undertaken against one doxastic background of further commitments available for use as auxiliary hypotheses can be taken up and made available as premises against a different doxastic background. Having the sort of representational significance that is analyzed here in terms of relations between the socially distinct perspectives of an interlocutor who undertakes a commitment and a scorekeeper who attributes it is an essential component of propositional content, and hence conceptual content generally. It is where the notion of propositional contents as truth conditions—as depending for their truth on the facts about the objects they represent—gets its grip. This is the sense of ‘represents’ that is fundamental to the intentionality of thought and talk. The analysis of the uptake of this sort of representational content in terms of the kind of social substitution-inferential interpretation deontic scorekeeping consists in what stands behind the claim that the discursive practice that embodies such intentionality must be understood as essentially a *social* practice. Propositional, and so conceptual, *understanding* is rooted in the *interpretation* that *communication* requires, rooted in the *social* practice of deontic scorekeeping.

#### 4. Two Other Senses of 'Represents'

The intentional aboutness of thought and talk—the representational dimension that arises out of the interaction between the social and the inferential articulation of discursive practice—should be distinguished from two other ways in which things can be treated in practice as having a broadly representational content. These are two ways in which some things or states of affairs can be treated as representing others. The first does not presuppose grasp of the specifically propositional contents conferred by distinctively discursive practice; the second does.

The preconceptual variety can be implicit in the practical capacity to navigate by a map. It is a matter of reliable differential responsive dispositions—altering one's behavior with respect to the *terrain* in response to features of the structure being employed as a *map*. In its simplest form, this sort of taking or treating as a representation need be no more complicated than seeking shelter from impending rain upon sighting a certain sort of cloud. It is based on treating one thing as a sign of another in the sense of acting in a way appropriate to the latter in response to the former. Navigating by a map need be no more than a systematic constellation of such differential responsive dispositions. Creatures evidently need not be able to talk in order to treat things as representations in this sense. This is a primitive ability, without which it would not be possible to engage in the more sophisticated practices linking inferentially and socially articulated propositionally contentful perception and action that provide the indispensable framework of the empirical and practical aboutness of our judgments. There is no reason to quibble about whether the implicit practical preconceptual attitudes of treating something as a representation by suitably systematic dispositions to differential responses embody some sort of intentionality, so long as the distinction between such attitudes and propositionally contentful ones is kept firmly in mind. The account of socially and inferentially articulated deontic scorekeeping practices, including the discussion of the substitutional and anaphoric substructures of that articulation, have been aimed precisely at saying what must be added to the primitive preconceptual representational ability in order to arrive at the full-blooded conceptual one.

Within the context of genuinely discursive practice, which institutes intentionality properly so-called, a propositionally explicit form of the primitive, practically implicit taking or treating of something as a representation becomes possible. At this level it is possible to endorse inferences from claims about a representing thing to claims about a represented thing. This is how we concept users typically use maps: *inferring* from the noninferentially acquired claim that there is a wavy blue line between the two black dots on the map *that* there is a river between the two cities in the terrain. This capacity explicitly to use one thing or state of affairs as a representation of another is parasitic on our capacity to make claims or judgments *about*

maps and pictures, on the one hand, and *about* what is mapped or pictured, on the other. On pain of an infinite regress we cannot understand what it is for our thought and talk to represent objects and states of affairs as a matter of our taking or treating it *as* representing them in this sense. One way of understanding the strategy of agent semantics, discussed in Chapter 3, is as assimilating the representational content of *talk* to this model of *explicit* treating as a representation, while appealing to the prior representational content of *thought*, which must be otherwise explained. A prominent candidate for the account that is then required of the more fundamental intentionality of thought is then to assimilate it to the sort of primitive preconceptual taking or treating as a representation by noninferential differential responsive dispositions considered in the previous paragraph. The deontic scorekeeping strategy, by contrast, interposes an intermediate level of socially and inferentially articulated practices that provide the context for the propositional and so conceptual contentfulness of both thought and talk—practices that presuppose the lower level of implicit, merely differentially responsive taking or treating as a representation, and are in turn presupposed by the upper level of explicit inferential taking or treating as a representation.

In the previous section it was claimed that the primary representational locution in ordinary language—the one we use to talk about the representational dimension of our thought and talk, to specify what we are thinking and talking *about*—is *de re* ascriptions of propositional attitude. It is the role they play in such ascriptions that gives their meanings to the ‘of’ or ‘about’ we use to express intentional directedness. It has also been claimed that the expressive role of these locutions is to make explicit a distinction of social perspective involved in keeping our books straight on who is committed to what. The social dimension of inference involved in the communication to others of claims that must be available as reasons both to the speaker and to the audience (in spite of differences in collateral commitments) is what underlies the representational dimension of discourse. Beliefs and claims that are *propositionally* contentful are necessarily *representationally* contentful because their inferential articulation essentially involves a *social* dimension. That social dimension is unavoidable because the inferential significance of a claim (the appropriate antecedents and consequences of a doxastic commitment) depends on the background of collateral commitments available for service as auxiliary hypotheses. Thus any specification of a propositional content must be made from the perspective of some such set of commitments.

One wants to say that the *correct* inferential role is determined by the collateral claims that are *true*. Just so; that is what *each* interlocutor wants to say—each has an at least slightly different perspective from which to evaluate inferential proprieties. Representational locutions make explicit the sorting of commitments into those attributed and those undertaken, without which communication would be impossible given those differences of per-

spective. The *representational* dimension of propositional contents reflects the *social* structure of the *inferential* articulation conferred by their role in the game of giving and asking for reasons.

### III. *DE RE* ASCIPTIONS AND THE INTENTIONAL EXPLANATION OF ACTION

#### 1. *From Theoretical to Practical Reasoning*

The intentional states characteristic of sapience are distinguished by their propositional contentfulness. When attributions of such states are made propositionally explicit, those contents must be specified as part of the ascription. There are two different forms in which the propositional contents of ascribed intentional states can be expressed: *de dicto* and *de re*.<sup>13</sup> The different expressive roles played by the vocabulary employed in these two sorts of content-specifications reflect differences in social perspective that are already implicit in the underlying practical attitudes of attributing and acknowledging propositionally contentful deontic statuses. To understand intentional states or deontic statuses as having specifically *propositional* semantic contents is to understand the norms governing their pragmatic significance as *inferentially* articulated. The *de dicto* and *de re* styles of specifying such contents are the explicit expression of the *social* dimension of that inferential articulation.

The two ways of expressing those contents arise because the inferential role of a claimable content appears differently from the various points of view provided by the background beliefs on the basis of which the one acknowledging a commitment and those attributing it assess its significance by projecting consequences, antecedents, and incompatibilities. This social articulation of inferential role—that endorsing a semantic content has systematically different pragmatic significances for different parties to a conversation—is essential to the discursively fundamental communicative practice of giving and asking for reasons. Since *de re* ascription of propositional attitude is the primary locution serving to make representational relations explicit in natural languages—providing the use in virtue of which words like ‘of’ and ‘about’ come to express intentional directedness—this social articulation is also what underlies the representational dimension of the conceptual content of intentional states. *Representational* content can be understood in terms of *inferential* content, provided the *social* dimension of inferential practice is properly taken into account.

Doxastic commitments, deontic statuses of the sort that corresponds to the intentional state of belief, are caught up in two kinds of reasoning. The discussion of what is expressed by *de re* ascriptions of doxastic commitments has so far been directed at *cognitive* or theoretical reasoning: reasoning whose premises and conclusions are doxastic commitments. Presystematically, this

is reasoning that leads from beliefs to further beliefs. But there is also *practical* reasoning: reasoning whose conclusions are practical, rather than doxastic, commitments. Practical commitments are commitments to *act*, to make-true a certain propositional content. Chapter 4 discussed how to make sense of such practical discursive commitments and the practical inferences they are involved in, in the context of a prior story about the practices that constitute doxastic discursive commitments, and how these notions can then be used to explain the concepts of *action* and *intentional explanation*. Practical discursive commitments are deontic statuses of the sort that corresponds to the intentional state of intention, rather than belief. Pre-systematically, basic practical reasoning accordingly leads from beliefs to intentions. Inferential commitments that are endorsements of the propriety of patterns of practical reasoning articulate the significance both of individual preferences or desires and of norms of various sorts. These can appear as further premises provided the language in question has sufficient expressive power to make explicit the endorsement of those patterns of practical inference.

To treat a performance as an action is either to take it to be the result of exercising a reliable differential disposition on the part of the agent to respond to acknowledgment of a practical discursive commitment with a certain content by producing a performance with a certain character or to take it to be itself the acknowledgment of such a commitment. As Anscombe and Davidson have taught us, for a performance to be an action is for it to be intentional under some description; the specification under which the performance is intentional is determined by the practical commitment the performance acknowledges. To offer an intentional explanation of an action is then to attribute a sample piece of practical reasoning, whose conclusion is the practical commitment or intention with which the action is taken to be performed. Action depends on noninferential capacities reliably to *make-true* claims of various kinds as, or in response to, the acknowledgment of practical commitments, as perception depends on noninferential capacities reliably to *take-true* claims of various kinds, to respond to enveloping situations by acknowledging doxastic commitments.

Reliability in the practical sphere of action, as in the cognitive sphere of perception, is assessed by a scorekeeper on the basis of the match between the contents of noninferentially eliciting or elicited commitments *attributed* to the agent or perceiver, on the one hand, and the contents of the commitments regarding what the agent is making- or taking-true that are *acknowledged* by the scorekeeper who is assessing such reliability, on the other. Attributing cognitive reliability involves assessments of the *truth* of judgments that are noninferentially elicited; attributing practical reliability involves assessments of the *success* of performances that are noninferentially elicited (the truth of the claim one sought to make true). Assessments of both of these sorts involve hybrid deontic attitudes on the part of the score-

keeper—*attributing* one commitment and *undertaking* another. Such attitudes are articulated according to the fundamental *I-thou* social structure of deontic scorekeeping practices. *De re* ascriptions of propositionally contentful intentional states (or deontic statuses) express just such hybrid scorekeeping attitudes; they incorporate the undertaking of substitutional commitments as well as the attribution of doxastic ones. That is why the *truth* conditions of ascribed beliefs are to be determined, according to the ascriber, by the substitutional commitments that govern expressions occurring transparently (that is, inside the scope of the ‘of’ when the ascriptions are suitably regimented) in *de re* specifications of the content of the ascribed belief. For these are the substitutional commitments *undertaken* by the scorekeeper, who is ascribing, and so *attributing*, a doxastic commitment.

It might accordingly be expected that *de re* ascriptions of commitments would serve an expressive role in making explicit the representational dimension of assessments of the *success* of intentions and actions that corresponds to their role in making explicit the representational dimension of assessments of the *truth* of beliefs and claims. So they do. There are two kinds of intentional explanation, roughly those that explain what an agent *tried* to do, and those that explain what an agent *succeeded* in doing. In each case the performance is made intelligible by attributing to the agent a constellation of commitments structured as a piece of practical reasoning. That structure is provided by the attribution of an implicit, practical-inferential commitment linking premises the scorekeeper takes the agent to be committed to with a conclusion consisting of a practical commitment whose acknowledgment is taken to have been, or resulted noninferentially in, the performance in question. The difference between the two kinds of intentional explanation lies in how the action being explained is specified, and so in the content of the governing practical commitment.

## 2. De Dicto and De Re Ascriptions Underwrite Two Different Kinds of Intentional Explanation

The first form of intentional explanation is the kind discussed in Chapter 4, which yields a specification under which the performance being explained is taken by the scorekeeper to be intentional. Why, it might be asked, did Nicole discharge her rifle in the direction of the animal in front of her? She believed that the animal in front of her was a deer, and she desired to shoot a deer, and she believed that if she discharged her rifle then and in that direction, she would shoot the animal in front of her, and so shoot a deer. Those commitments, in the absence of competing ones, provided sufficient reasons for the formation of an intention that she shoot the animal in front of her by discharging the rifle then and in that direction. This last specification is one under which her performance was intentional, and hence an action. For this is a specification under which she *acknowledges* the

practical commitment, and it is that attitude that noninferentially elicits the performance.

But Nicole's performance admits of other descriptions. As Davidson reminds us under the heading of "the accordion effect," any consequence of the performance can be appealed to in order to specify what Nicole did—although most of these will be descriptions under which it was not intentional. Suppose that another such description of Nicole's performance is that she killed a cow. That is something she succeeded in accomplishing, something she did, not merely something that happened, though it is not (under that description) something she intended to do. Nonetheless, her action can still be made intelligible under that description by offering an intentional explanation of it. The animal in front of her was in fact a cow, not a deer. So Nicole believed *of a cow* (namely the animal in front of her) that it was a deer, and she desired to shoot a deer, and believed that if she discharged her rifle then and in that direction, she would shoot the animal in front of her, and so shoot a deer. Those commitments, in the absence of competing ones, provided sufficient reasons for the formation of an intention of, about, or directed at a cow (the animal in front of her) that she shoot a deer by shooting it (that animal) by discharging the rifle then and in that direction.

As the awkward expression of the intention appealed to by this intentional explanation is supposed to suggest (via the ascriptional regimentation put in play earlier), the difference between the two sorts of intentional explanation is just whether the contents of the attributed commitments it appeals to are specified *de dicto* or *de re*. *De dicto* ascriptions specify the contents of attributed commitments in terms that, according to the ascriber, the one to whom they are ascribed would acknowledge as specifications of the contents of commitments undertaken. From Nicole's point of view (according to the ascriber), her intention is that she shoot a deer and her belief is that the animal in front of her is a deer. She does not acknowledge these as states that are of, about, or directed at a cow. The claim that they are cow-representing states, that she represents a cow as a deer, is part of the commitment the ascriber *undertakes*, not part of what the ascriber *attributes*. Thus an intentional explanation whose point is to show things from Nicole's point of view—to show what she was *trying* to do, under what specifications she acknowledged the performance as her own, under what descriptions it was intentional—employs ascriptions that express the contents of the attributed commitments in the *de dicto* style.

In contrast, an intentional explanation whose point is to show why the action turned out as it did—why what Nicole intentionally did had the *consequences* it did, accomplished what it did, succeeded or failed as it did—is provided from the point of view of the ascriber. Intentional explanations of this sort employ ascriptions that express the contents of the attributed commitments in the *de re* style. That Nicole's belief that the animal in front of her was a deer was a belief *of a cow* that *it was a deer* in front of her

is just what is needed to explain why Nicole shot a cow. Corresponding to these two specifications of the content of Nicole's belief are two specifications of the content of Nicole's intention. She intended that she shoot a deer, but in the actual circumstances, where she intended to do that by shooting the animal in front of her and the animal in front of her was in fact a cow, she thereby (without realizing it) intended of a cow that she shoot *it*.

The basic form of intentional explanation is, on this account, the one that employs *de dicto* specifications. For according to this way of understanding the distinction, behind every *de re* ascription there must be a *de dicto* ascription (though the ascriber need not always be in a position to produce it in order to be entitled to the *de re* ascription). For to be committed to the claim that *S* believes of *t* that  $\Phi(\text{it})$  is to be committed to there being some *t'* such that  $t = t'$  and *S* believes that  $\Phi(t')$ . Since the distinction has to do with how the *content* of the ascribed commitment is specified, rather than with what sort of a commitment it is, the story about intentions is no different from that about belief.<sup>14</sup> In this case it is because the animal in front of Nicole was a cow and Nicole intended that she shoot the animal in front of her that she intended of a cow that she shoot *it*. Shooting a cow is something she *succeeded* in doing, although it is not something she was *trying* to do—that is, this *de re* way of specifying the content of her practical commitment is not one under which she acknowledges it. Shooting a deer is something she was *trying* to do but *failed*.

Shooting the animal in front of her is something that she both tried to do and succeeded in doing. Of course she might have failed even in this, if her aim had been bad. But then she would still have succeeded in firing the gun. If she failed even in this because of wet powder, she would have succeeded in pulling the trigger, or if it had jammed, in moving her finger. Davidson takes it that in each case of intentional action, there will be *something* the agent both intended to do and succeeded in doing. That is, he insists that some description under which a performance is intentional be *true* of that performance. The characterization of intentional action offered here does not involve such a commitment. As long as the performance was produced by exercising a reliable noninferential capacity to respond differentially to the acknowledgment of a practical commitment with a given propositional content by producing performances of which that content is true,<sup>15</sup> it does not matter if that particular case is one in which the usually reliable capacity misfired. If intending to pick up the bread I spill the wine, there is on this line no theoretical commitment to insisting that there must be something I intended to do that I succeeded in doing, for instance moving my arm. In this case my performance is intentional under a description, 'picking up the bread', that is not true of it.<sup>16</sup> Such a characterization makes no sense for intentions-in-action, where the performance *is* the acknowledgment of the practical commitment. But even an *unsuccessful* performance that is the result of exercising a usually reliable differential disposition to respond to the

acknowledgment of a prior commitment by producing a suitable performance can, just because of its provenance, qualify as intentional, according to this non-Davidsonian line.

Picking up the bread entails moving my arm, so a practical commitment to doing the one entails a practical commitment to doing the other. Undertaking commitment to picking up the bread accordingly involves undertaking a commitment to move my arm. But what corresponds to intentions is practical commitments not just *undertaken* but *acknowledged*, just as the beliefs that are efficacious in deliberation are not all those consequentially undertaken but only those acknowledged. In each case the attitude of acknowledgment need not be preserved by committively good inferences. Picking up the bread also entails contracting a variety of muscles in complicated ways, sending electrochemical signals through various synapses, and so on, but doing so need not in any ordinary sense be part of my intention. Even though in one sense I can move my arm only *by* doing these things, I can *intentionally* contract those muscles in just that way and so on only *by* reaching for the bread. Since in these circumstances picking up the bread *is*, in fact, contracting my muscles in those particular ways, in intending that I pick up the bread, I am intending *of* something I can do only by contracting my muscles in those particular ways that I do *that*. The identity is not one that I need to acknowledge in order to have the intention. A theorist who does acknowledge it, however, is in a position to offer a *de re* specification of the content of my intention. Just such specifications may need to be appealed to in explaining why my attempt to pick up the bread succeeded—or failed because of the unnoticed interposition of the wine bottle. Underlying such *de re*, or attributor-centered, ascriptions, however, must be some *de dicto*, or agent-centered, ascription. These specify the content of the intention in terms the agent would (according to the ascriber) acknowledge, and so cannot rely on inferential commitments undertaken by the ascriber but not attributed to the agent.

For individuals can be trained to have reliable noninferential dispositions to respond differentially only to their own *attitudes* of acknowledging commitments, not to their *statuses* of being committed, insofar as these outrun those they acknowledge. In this sense only such acknowledgments are causally efficacious in an agent's behavioral economy. In another sense, of course (a sense that the analysis of *de re* ascriptions as expressing hybrid deontic attitudes based on *de dicto* ascriptions seeks to explain), actions are causally responsive to features of the contents of intentions that are expressed by *de re* specifications, which outrun the agent's dispositions to acknowledge commitments. For what Nicole does, shooting a cow, depends on the fact, unacknowledged by her, that the animal in front of her is a cow. The two sorts of intentional explanation correspond to reconstructing the context of *deliberation*, in which only what the agent is aware of or acknowledges is relevant, and reconstructing the context of *assessment* of action, in which various

factors that are not accessible to the agent can be invoked in explaining how things turned out. Beliefs ascribed *de dicto* answer to the believer's dispositions to acknowledge, avow, or act according to the belief. Beliefs ascribed *de re* answer to what would make the belief true—which is to say, to what it in fact (according to the ascriber) represents.

It is tempting to formulate this difference in terms of a distinction between *narrow* and *wide* aspects of intentional content, the first of which depend only on the individual whose intentional states are in question, the latter of which depend also on the environment in which the individual with those states is embedded.<sup>17</sup> This is misleading if the narrow sense is identified with a methodologically solipsistic or individualistic point of view.<sup>18</sup> For the expressions that are used to specify the contents of both ascriptions *de dicto* and ascriptions *de re* are drawn from the same public vocabulary. There is no reason to think that any of this vocabulary can be semantically funded on the basis of the activities of a single individual. The content specifications of *de dicto* ascriptions, owing allegiance as they do to the acknowledgments of the one to whom they are attributed, do express (according to the ascriber) the agent's *subjective* perspective on intention and action. The content specifications of *de re* ascriptions, owing allegiance as they do to the facts bearing on actual outcomes or consequences of performances, do express (according to the ascriber) the *objective* aspect of intention and action. But these are complementary points of view that arise within the essentially social (in an *I-thou* sense) practice of deontic scorekeeping that alone is capable of conferring genuinely conceptual content.

What is expressed by *de dicto* ascriptions is conceptually prior to what is expressed by *de re* ascriptions *locally*, but not *globally*. That is, each true *de re* ascription is on this account true in virtue of some underlying true *de dicto* ascription: *S* believes of *t* that  $\Phi(\text{it})$  just in case for some *t'*,  $t = t'$  and *S* believes that  $\Phi(t')$ . It is also true (as will be seen in the next section) that *de dicto* ascriptions can be introduced into a language that previously had no way of making attributions of intentional states explicit, in advance of introducing on that basis also *de re* ascriptions; and the converse is not true. But what is expressed (where suitable explicating vocabulary is available) by *de dicto* specifications of the content of intentional states cannot be made sense of apart from what is expressed (where suitable explicating vocabulary is available) by *de re* specifications of the content of intentional states. Propositional contents (and so conceptual contents more generally) are essentially the sort of thing that can appear from both perspectives. Just as one cannot understand the content of a claim or judgment unless one has a practical mastery of the social distinction of attitude involved in assessments of it as *true* or *false*, so one cannot understand the content of an intention or an action unless one has a practical mastery of the social distinction of attitude involved in assessments of it as *successful* or *unsuccessful*.

### 3. Successful Action, Truth, and Representation

The success or failure of an action concerns the relation between the intentional *antecedents* of a performance and its actual *consequences*. It is assessed (as the truth of a belief is assessed) by comparing the commitments the assessor takes the agent to acknowledge with those the assessor acknowledges. Where suitable expressive resources exist, such assessments are made explicit by *de re* ascriptions. Understanding action accordingly requires that scorekeepers have binocular vision—that they be able to move back and forth between what is expressed by the *de dicto* ascriptions that figure in intentional explanations of one kind and the *de re* ascriptions that figure in intentional explanations of another kind. Without the capacity to produce and understand the first kind of intentional explanation, one cannot understand the difference between action and mere behavior: between shooting the gun and the gun's just going off. Without the capacity to produce and understand the second kind of intentional explanation, one cannot understand how shooting the cow can be something genuinely *done*: an action and not merely behavior, even though it was not done intentionally. (Turning on the light, which one intends, may be alerting the burglar, which one does not intend.) That is, one cannot understand the success and failure of intentions. To do that one must grasp the practical scorekeeping relation between specifications of the performance that are privileged by being the ones under which the agent is prepared to acknowledge or take responsibility for it, on the one hand, and the specifications (often in terms of consequences) under which it is available to anyone interpreting (keeping score on, attributing commitments to) the agent, on the other.

Explanations of the success or failure of an action—of why one's trying to do what one did had the consequences it did—appeal to what is true (according to the one doing the explaining), in addition to what (according to the one doing the explaining) the agent believed and so intended. To explain the success of an action, one must initially be able to ascribe beliefs and intentions *de dicto*—in order to say what the agent was *trying* to do, which sets the criteria of adequacy for *success*. *De re* recharacterization of those beliefs then puts them in a form in which their truth can be assessed. Typically the *success* of an action is to be explained by citing the *truth* of various of the agent's motivating beliefs. Nicole would have been successful if her belief that the animal in front of her was a deer had been true, that is if that belief had been a belief of a deer—rather than of a cow, that it was a deer. This is not the invariable form of an explanation of success, for I might succeed in my intention to put (that I put) oregano in the stew because of my false belief that the bottle you just carried to the counter contains oregano, in combination with my false belief that it is the bottle on the left, in case the oregano happens to be in the bottle on the left. Explanations of an agent's *reliable*

capacity to act successfully in some domain, however, cannot depend on such accidents and must accordingly appeal to true beliefs: Nicole believes of the animal in front of her that it is a deer, and it is true of the animal in front of her that it is a deer (that is, the animal in front of her is a deer). So the *success* of actions is to be assessed and explained by appealing to what is *true of* what the various intentional states leading to those actions represent or are *about*.

The analysis of what is expressed by *de re* specifications of the content of intentional states shows what this sort of representational contact with reality consists in—what is being *done*, in scorekeeping terms, when an assessor attributes such contact. Representational content is an essentially *social affair*—a consequence of the role claims play in interpersonal communication. *De re* ascriptions express the potential one individual's remark has as a premise from which *others* can draw conclusions. That is why they are the ones that figure in assessments of *truth*, which involve what the assessor endorses (not just attributes).

In diachronic communication and assessment of truth—for instance when explaining the (limited) practical successes of earlier theories and assessing the truth of their claims—the same structure is evident. Cognitive progress consists in saying more and more *true* things *about* the objects we talk about. Only such progress can explain increases in practical reliability—improvements in the long-term likelihood of success of actions based on those theories. The deep connection between our tautological *subjective* interest in the *success* of our practical endeavors (in getting what we want) and the *objective truth* and representational adequacy of our theoretical commitments (how we take things to be) is accordingly rooted in the social structure of hybrid deontic scorekeeping attitudes that is expressed in *de re* specifications of the contents of ascribed intentional states.

Though thoroughly social and linguistic, these representational contents are not *merely* linguistic—for they are not entirely up to us in the way in which what noises we use is entirely up to us. Assessing the success of action, for instance, involves, not only consideration of noninferential *exit* moves (from acknowledgments of commitments to performances) on the part of agents, but also noninferential *entrance* moves (from performances to acknowledgments of commitments) on the part of perceivers of those actions. The discursive practice that confers representational content is empirically and practically solid. It is a matter not just of hollow words and other linguistic expressions but also of mastery of their concrete circumstances of appropriate application and appropriate consequences of application, not all of which are purely inferential. Indeed, assessments of the success of actions often involve perceptual responses to intentional performances: Nicole and her scorekeepers can tell whether what she has shot is a deer by looking at it. This is one way in which agents themselves can come to be in a position to offer *de re* characterizations of their own previous beliefs, which can then

be used inferentially to alter the plans and intentions adopted in aid of some goal. Acknowledgments of doxastic commitments adopted noninferentially in response to performances produced noninferentially in response to acknowledgments of practical commitments can accordingly provide the feedback needed to complete the loop of a Test-Operate-Test-Exit cycle controlling goal-seeking behavior. The process by which commitments and concepts develop in a community over time through this sort of collaboration of the practical and empirical aspects of discursive practice is the expressing (in an inferentially articulated and hence conceptually explicit form) of the concrete constraints supplied by the fact that we think and act in an objective world.

The aim of the chapter so far has been to delineate the essential *social* structure that underlies the representational dimension of discursive practice, by offering an account of the hybrid deontic attitudes that are expressed by *de re* specifications of conceptual content—the primary representational idiom in natural languages. The objectivity of conceptual content—the way in which its proper applicability is determined by how things are in such a way that anybody and everybody might be wrong in taking such a content to apply in particular circumstances—is not by these means to be explained away. Rather the way in which the social and inferential dimensions of discursive practice combine in conferring, securing, and deploying such contents is to be investigated.

#### IV. FROM IMPLICIT ATTRIBUTION TO EXPLICIT ASCIPTION

##### 1. *Prospectus*

The fundamental claim made in this chapter is that the *representational* dimension of discourse is rooted in *social* differences of inferential perspective, which correspond to different repertoires of discursive commitments. Understanding that representational dimension is the key to understanding the objectivity of conceptual norms—how assessments of the objective correctness or incorrectness of the application of concepts (and so of the making of claims) can be explained in terms of social practices of deontic scorekeeping. As indicated in the opening chapters, it is a critical criterion of adequacy of any account of sapience or discursive practice that it be able to make intelligible this objectivity of the norms that govern the use of concepts. In the first three sections of this chapter, the crucial underlying representational dimension of the semantic contents conferred on states, attitudes, and performances by the role they play in discursive practice has been picked out as what is expressed by (weak) *de re* ascriptions of propositional attitude, and a deontic scorekeeping account of the use of such ascriptions has been sketched.<sup>19</sup>

This section examines more closely the explicating expressive function

of ascriptions employing *de dicto* and *de re* specifications of propositional content. It focuses not, as heretofore, on what they make explicit but rather on how they do so. This account of ascriptional locutions takes the form of a structured recipe for introducing such locutions into an idiom that previously lacked their characteristic expressive power. The next section then moves from concern with epistemically weak *de re* ascriptions—which have not traditionally received much attention, though they are philosophically central according to the line of thought pursued here—to discuss epistemically strong *de re* ascriptions (and indeed, beliefs), which have standardly been taken to be of the greatest theoretical interest. With these accounts in hand, the final section returns to the issue of objectivity—of what it means for claims and the application of concepts to answer for their correctness to what is *true of* what they in fact represent or are about, rather than to what anyone or everyone *takes* to be true of whatever they *take* them to represent or be about.

The leading idea of the treatment of ascriptions presented here is that propositional-attitude-ascribing locutions are a species of *logical* vocabulary. Logical vocabulary has been distinguished in this work by its expressive role in making *explicit*, as something that can be *said*, some constitutive feature of discursive practice that, before the introduction of that vocabulary, remained *implicit* in what is *done*. In the paradigmatic case of introducing conditional locutions, where before one could treat inferences only as good or bad in practice, one comes by their means to be able to claim explicitly that one claim follows from another—to say that if *p* then *q*. Thus interlocutors can express the implicit inferential commitments in virtue of which (along with noninferential empirical and practical circumstances and consequences of application) sentences have the contents they do as the contents of explicit assertional commitments. Conditional, negating, identity, and quantificational locutions make explicit the inferential and substitutional commitments that articulate the *semantic contents* of states, attitudes, and performances. Ascriptional locutions, by contrast, make explicit *pragmatic attitudes*—paradigmatically attributions of commitment—which articulate the scorekeeping practices that confer those contents. In this sense they are pragmatic, rather than semantic, explicating vocabulary—a different species within the genus of locutions qualifying as logical in virtue of their explicating role.<sup>20</sup>

## 2. Introducing ‘Says’: Direct Quotation

The simplest form of ascription makes the attribution of an assertional speech act explicit in the form of a claim:

Wittgenstein says, “Wovon man nicht sprechen kann, darüber muß man schweigen.”

Since what is being considered is assertionally explicit ascription of specifically *assertional* speech acts, 'says' here has the sense of 'claims', or 'says claimingly'—that is, it means *utters with assertional force*, rather than simply *utters*. (So in this sense in saying a conditional, one does not say the antecedent.) How must an expression be used in order to play the ascriptional role of 'says' here in making explicit the attribution of assertional speech acts?

One of the implicit practical capacities interlocutors must be able to exercise in order to keep score on each other's commitments (and so to engage in discursive practice at all) is to treat each other as producing performances that have the pragmatic significance of claimings or assertions, which express explicit acknowledgment of doxastic commitments.<sup>21</sup> So part of being a competent discursive scorekeeper is being a reliable perceiver (under appropriate circumstances) of claimings, in that one is able differentially to respond noninferentially to the assertional performances of others by attributing to them those speech acts and the commitments they have thereby acknowledged. The response that is reliably elicited is then just the adoption of a practical scorekeeping attitude.

One essential expressive job of the assertion-ascriptional 'says' is to add an assertional response that is elicited by the same range of performances interlocutors are trained to respond to noninferentially by adopting such scorekeeping attitudes. 'Says' is used to form *reports* of assertional speech acts. Part of what it is for those reports to make explicit in the form of assertions what is implicit in attributions of doxastic commitments is that the assertional ascriptive commitments they express be *undertaken* or acknowledged noninferentially as responses to the same performances that elicit *attributions* of the commitments they ascribe. This expressive condition provides these ascriptions with noninferential circumstances of application. Like any perceptual reports, however, once these claims are available in the language, they can be relied upon as *testimony*.<sup>22</sup> Thus someone who has never read Wittgenstein can come to be committed (and entitled) to the assertional claim made above by overhearing someone else making that claim. And that individual may have acquired entitlement to the claim either directly, as a response to the speech act itself (in this case, an inscriptional one), or by testimony. Thus it follows from the general account of assertional practice that the circumstances of application corresponding to these ascriptions include not only assertional speech acts of the sort ascribed but also type-similar assertional assertions. Since the latter class of testimonial circumstances of application is common to all assertions as such, it is only the former class that is distinctive of specifically assertional assertions.

One essential consequence of application of these ascriptions has already been mentioned: asserting an ascription involves attributing a commitment as well as undertaking one. The distinctive consequences of application of these ascriptions also include their use in making explicit *deferrals* of the

responsibility to redeem one's *entitlement* to a claim. Recall that the basic model of assertional practice (introduced in Chapter 3) is structured by a constellation of complementary assertional *authority* and *responsibility*. The testimonial authority of assertional speech acts (their pragmatic significance as licensing others to undertake the same doxastic *commitments* they express) is balanced by a responsibility to exhibit, if appropriately challenged, one's *entitlement* to those commitments. This may be done in any of three ways: offering an inferential justification, exhibiting the commitment (and so the corresponding entitlement) as the result of exercising a reliable non-inferential perceptual ability, or invoking the testimony of an informant, to whom one defers the responsibility to redeem the entitlement in question.

Inferential justification is always a matter of making further explicit claims, which are thereby implicitly put forward as inferentially related to the claim in question, as premises to conclusion. The other two sorts of entitlement vindication are conceived of, to begin with, as nonassertional performances—of any kind whatever that might be invested with the appropriate practical scorekeeping significance. Thus in the most primitive cases, to invoke perceptual authority one might point to one's eyes or ears. To invoke testimonial authority one might point to one's informant, implicitly deferring responsibility the way someone in a theater line does by passing the ticket taker with a thumb jerked over the shoulder to indicate that the one next in line has tickets for both. But tropes may be introduced that permit these sorts of authority to be claimed explicitly. In the case of perception, these could range from Wittgenstein's suggestion "I speak English" to "I am a reliable reporter of red things under circumstances like these." Explicit deferrals are ascriptions.

In the case of the basic ascriptions formed using 'says', inferential justifications are minimal. As with any assertions, in languages with sufficient Boolean logical vocabulary, ascriptions follow from conjunctions in which they figure as conjuncts, and so on. Since as just indicated they can play the role of noninferential reports, they will share this mode of entitlement vindication with all other reportables. Their distinctive consequence of application, however, is that they can function to make deferrals explicit. That is, they are assertions that can have the pragmatic scorekeeping significance of deferrals. Instead of implicitly attributing an assertion to an informant by a wordless, (or at any rate assertionally inarticulate) deferral, one can say that so-and-so made the claim in question. Thus someone might attempt to vindicate entitlement to an assertive tokening of "Wovon man nicht sprechen kann, darüber muß man schweigen" by explicitly ascribing it to Wittgenstein. The assertion of the ascription offered as an example above then would function as a petition to inherit his authority for that view: anyone interested in the speaker's credentials for the commitment should investigate the reasons Wittgenstein has to offer for his. This role in making deferrals explicit is distinctive of ascriptions and is not inherited simply from their role as assertions, nor from their role as possible noninferential reports.

These ascriptions can accordingly be picked out by three fundamental features of the assertional commitments they express:

- the undertaking of those commitments involves attributing further assertional commitments acknowledged by others,
- their acknowledgment can be noninferentially elicited by speech acts by others that acknowledge further assertional commitments, and
- speech acts that acknowledge them can have the significance of deferring responsibility to vindicate entitlement to further assertional commitments to other interlocutors who have acknowledged commitment to those further assertional commitments.

To play this role, ascriptions must specify who these others are—the others to whom commitments are attributed, whose speech acts are being reported, to whom the ascriber can defer (should the ascriber not only attribute, but also undertake commitment to the claim ascribed). Even before locutions with the expressive power of ascriptions are available, interlocutors must already be able to keep score on the deontic statuses of different individuals. This implicit practical capacity to distinguish interlocutors is then exercised by the explicit use in ascriptions of singular terms that play the role that 'Wittgenstein' does in the example above. The use of singular terms (in 'referentially transparent' contexts such as this) has already been described.

More important, however, each of these distinctive features of the use of ascriptions of assertions appeals to a relation between an ascribing assertional commitment or its acknowledgment and the acknowledgment of some further assertional commitment, possibly on the part of another.<sup>23</sup> Ascriptive expressions must accordingly somehow specify those further assertional commitments. They do this by containing an embedded declarative sentence, one whose freestanding use would be the undertaking of an assertional commitment. The interesting differences between the various kinds of ascriptions concern the relation between this embedded tokening, on the lips or the pen of the ascriber, and various possible tokenings on the lips or the pen of the one to whom a commitment or its acknowledgment is ascribed. The basic ascriptions of assertional speech acts (overt expressions of acknowledgments of assertional commitments) whose expressive role has been considered so far are *direct discourse* or *quotational* ascriptions. This means that the relation between the ascribed sentence-tokening and the embedded sentence-tokening that specifies what is ascribed is one of lexical *cotypicality*. Acknowledgment of the commitment expressed by the ascription "Joanne says '*p*'",<sup>24</sup> involves attributing to Joanne the assertional commitment she would be acknowledging by assertively uttering a tokening of that very same sentence *p*—that is, one of the same lexical-syntactic type as that appearing in the ascription. Similarly, acknowledgment of the commitment expressed by that ascription is to be appropriately elicited noninferentially by a tokening by Joanne that is of the same type. And that ascription

is in order as an explicit deferral (though not necessarily successful in redeeming entitlement, for that depends on Joanne's entitlement) just in case Joanne has acknowledged an assertional commitment by uttering a sentence of that type. Any linguistic device for segregating the embedded sentence that has the significance of indicating this relation of cotypicality between reported and reporting tokening in ascriptions is playing the expressive role of quotation marks.

### *3. Introducing 'Says That': Indirect Discourse*

Ascriptions accordingly can be thought of as having three parts: one specifying the *target* of the ascription by a singular term ('Wittgenstein', 'Joanne'); one specifying what sort of state, attitude, or performance is being ascribed ('says'); and one specifying its propositional content ("Wovon man nicht sprechen kann, darüber muß man schweigen," [ $p$ ]). It has been shown how to understand in deontic scorekeeping terms the role played by each of these, in the case of the direct discourse or quotational ascriptions considered so far. The next step is to move from direct to indirect discourse, from *oratio recta* to *oratio obliqua*. Examples are such sentences as:

Kant says that the essence of Enlightenment is responsibility.

Hector says that he is hungry.

John says that I am confused.

The circumstances and consequences of application of this sort of ascription are structurally much the same as for ascriptions of sayings, with one important exception. Direct discourse ascriptions can serve (as well as indirect discourse ascriptions) as inferential grounds capable of justifying indirect discourse ascriptions.

For example, the ascription

Wittgenstein says, "Wovon man nicht sprechen kann, darüber muß man schweigen"

can be used to justify the ascription

Wittgenstein says that what one cannot speak about one must pass over in silence.

Similarly, Wittgenstein's inscriptive tokening of

Wovon man nicht sprechen kann, darüber muß man schweigen

could be responded to noninferentially by that same indirect ascription and could be the basis for a deferral achieved by asserting it. What is special about these indirect discourse ascriptions—and what must be understood if the asymmetric inferential relation between direct and indirect varieties of as-

cription is to be intelligible—is the use of the content-specifying clause, now marked in the regimentation by ‘that’ rather than quotation marks.

The fundamental considerations needed to understand these indirect discourse ascriptions have been laid out by Davidson, in his seminal essay “On Saying That,” and have been clarified and developed further by McDowell in his commentary “Quotation and Saying That.”<sup>25</sup> The approach they elaborate, which Davidson calls “the *paratactic theory*,” conjoins three basic ideas. First, it focuses on tokenings, rather than on types.<sup>26</sup> Second, the speech act being ascribed is specified by *displaying* a sentence tokening related to the one being ascribed. The essence of the account accordingly consists in its understanding of the relation between the *reporting* tokening and the *reported* tokening. It is worth elaborating this point briefly.

Of course, direct discourse ascriptions also specify the ascribed saying by displaying a tokening, namely the one that is quoted, but all that matters about that tokening is its type. Indirect discourse loosens the constraint of cotypicality. Davidson emphasizes that the reported and reporting tokenings need not be in the same language, as in the specification in English above of what Wittgenstein said (or wrote) in German. McDowell acknowledges the significance of *translation* in indirect ascription, but he is concerned also to emphasize that *indexicals* behave very differently in ‘that’ clauses than they do when sealed inside quotation marks. Quite different remarks are reported by the ascriptions:

John says, “I am confused”

and

John says that I am confused.

The first has John characterizing himself, the one to whom the claim is ascribed, as confused, while the second has John characterizing the ascriber as confused.

They differ as to whether the indexical is to be evaluated according to the context of the reported tokening or according to the context of the reporting tokening. The demonstration and context relevant to understanding the indirect

John says that *that man* is confused

is that of the ascriber, while the demonstration and context relevant to understanding the quotational

John says, “That man is confused”

is that of the one to whom the speech act is ascribed. Although McDowell does not make the point, expressions that are not canonically indexical but whose recurrence classes are also not determined simply by cotypicality—

paradigmatically *anaphoric* proforms—also behave differently in the two sorts of ascriptions.

Hector says, “He is hungry”

and

Hector says that he is hungry

admit quite different readings. The interpretation of the first ascription turns on the anaphoric antecedents that are made available by the discursive context of the reported tokening, while the interpretation of the second ascription turns on the anaphoric antecedents made available by the discursive context of the reporting tokening.

The third idea underlying the paratactic approach is that the relation between the reported and the reporting tokenings is what Davidson calls “*samesaying*.” Davidson is content to leave this as the rough-and-ready, more-or-less notion of sameness of content as what is preserved by good translations. This characterization of *samesaying*, however, is not of much help in thinking about the use of indexical and anaphoric tokenings in indirect ascriptions. A more useful way to put the condition is that the displayed tokening specifies the content of the ascribed uttering by being a *samesaying* with it, in the sense that its assertional utterance *would be* an undertaking of commitment to the very same claim that the ascriber undertook commitment to by the assertional utterance of the ascribed saying.<sup>27</sup>

In the context of the model of discursive practice developed here, an important feature of the expressive role of ascriptional contents is that their assertion can have the pragmatic significance of an assertorally explicit *deferral* of justificatory responsibility for a claim to the one to whom it ascribes a claim. The notion of sameness of claim appealed to by this formulation can be picked out in terms of the pragmatic notion of deferral potential. For a scorekeeper to treat a displayed sentence-tokening as standing in the appropriate *samesaying* relation to some other possible ascribed tokening is to take it that if the displayed tokening *were* uttered assertorally, deferral to the ascribed tokening would be in order.<sup>28</sup> Thus understanding what ascriptions make explicit offers a way of clarifying the notion of *samesaying* that governs the relation between reported and reporting tokenings in indirect ascription. For understanding the relation between reported and reporting tokening in terms of what claim *would* be made by the reporting tokening were it used assertively entails that the content of indexical expressions be evaluated according to the circumstances surrounding the ascriber’s utterance, not the ascribed one.

This way of doing the work Davidson invokes *samesaying* for also deals with the difficulty that leads McDowell, in a later article,<sup>29</sup> to retreat from the *samesaying* formulation in favor of a notion of “corresponding senses”—a

relation stronger than coreference but weaker than identity of sense. The difficulty, discussed in more detail in the next section, arises from the fact that the special role that expressions such as 'I' and 'now' play in practical reasoning and the formation of intentions, particularly intentions in action, determines a sense in which no one else can have a thought with just the same content that John expresses by saying, "I am confused." "You are confused," in my mouth, at best expresses a corresponding thought, not the very same one. For present purposes it is enough to notice that understanding the relation between reported and reporting tokenings in terms of scorekeeping proprieties of deferral of justificatory responsibility permits these cases to be taken in stride; I can justify my assertion of "You are confused," addressed to John, by deferring to his assertion of "I am confused." The issue of whether the significance of the special features of the use of 'I' and 'now' is best addressed by individuating claim contents more finely accordingly need not be dealt with here.

Thus the paratactic approach to indirect discourse explains the expressive role of 'that' clauses (which specify the contents of ascribed speech acts and commitments) by focusing on the relation between a possible tokening in the ascriber's mouth and a possible tokening in the mouth of the target of the ascription. What primarily distinguishes this style of content specification from the directly quotational is that quotation marks seal off content-specifying expressions from the context of utterance of the ascriber. By contrast, in indirect discourse the content expressed by words appearing in the scope of the 'that' depends on the context of evaluation of the ascriber who actually utters them: the personal pronouns the ascriber uses, the demonstrations the ascriber makes, the anaphoric antecedents available to the ascriber, even the language the ascriber speaks. Davidson sets out his account as a response to a particular problematic; its task is explaining why substitution within 'that' clauses of sentences with the same truth-value does not preserve the truth of the ascriptions containing them. His official resolution of the problem is that the sentence tokening that occurs within the 'that' clause is not strictly part of the ascription at all. Properly understood, the ascription has the form "*S* said that." Another sentence tokening has been appended—without benefit of punctuation, which is why Davidson calls the theory "paratactic"—simply to display the referent of 'that', which is used demonstratively.<sup>30</sup> Thus substituting another sentence tokening for the one occurring in the scope of the 'that', whether or not the substituend has the same truth-value, alters the *referent* of 'that', and so should not be expected to preserve the truth-value of the whole ascription. McDowell is right to point out the affinities such an analysis has to an assimilation of indirect discourse to quotational discourse (in spite of the sharp distinction Davidson wants to make between his approach and quotational ones)—though the quotational model needs just the sort of commentary McDowell provides about differential semantic permeability to pragmatic context. For it is just because the

corresponding substitution alters the reference of quote-names that it is not expected to be truth preserving in direct discourse.

The lesson of the presentation of the fundamental insight of the paratactic analysis presented in the last few paragraphs, however, is that *all* the work here is being done by the emphasis on the relation between the *tokening* displayed in the ascription and a possible tokening by the one to whom a speech act or commitment is ascribed. So long as this point is insisted upon, it is otiose—a matter of being cute at the cost of being misleading—to treat the ‘that’ in indirect quotations as a demonstrative referring to (or an anaphoric dependent whose antecedent is) the sentence tokening that follows it. Although this is the most striking feature of Davidson’s approach, it is just as well that this feature can be shown to be dispensable. For treating the role of ‘that’ as subordinating conjunction in ascriptions as derivative from its role as a demonstrative is both an etymological howler in English and depends on a coincidence that is not repeated in many other languages. It can simply be jettisoned, however, and Davidson’s genuine insights put in the form of a direct description of the expressive role of the sentence tokenings displayed in the ‘that’ clauses of indirect discourse ascriptions.

The account offered here consists of two suggestions for filling in Davidson’s analysis of saying-that. The first would expand his formula, along the following lines.

Galileo said that the earth moves

is to be understood as having the sense of:

Galileo said (something that *in his mouth then* committed him to what an assertional utterance of this *in my mouth now* would commit me to): The earth moves.

This way of putting it entails that the content of indexical (and anaphoric) expressions should be evaluated according to the circumstances surrounding the ascriber’s utterance, not the ascribed one, which is what is distinctive of indirect quotation. The second suggestion is that the notion of identity of commitment undertaken that does duty here for Davidson’s samesaying be understood in terms of the (*ceteris paribus*) scorekeeping propriety of *deference* to the first performance in vindicating entitlement to the second (merely possible) one. Each of these ways of developing Davidson’s idea expresses an understanding of the paraphrase as explicating the rules by which the reporting tokening, occurring in the content-specifying clause, relates to the reported tokening.

A common way of botanizing approaches to what is expressed by propositional-attitude-ascribing locutions distinguishes *sentential* and *propositional* approaches. Sentential approaches understand belief and similar attitudes as essentially involving relations to sentencelike entities,<sup>31</sup> where propositional approaches understand them as essentially involving relations

to more abstract objects. Sentential approaches have the advantage of relating the possession of propositionally contentful states to something intentional agents can unproblematically *do*, namely use sentences, but they are widely believed to have difficulty accommodating the sort of dependence of content on context exhibited by indexicals. Propositional approaches have the advantage that they can easily accommodate such contextual dependences, but they face a greater challenge in explaining how what believers *do* connects them with the abstract objects in virtue of which their beliefs are contentful.<sup>32</sup> The present approach, by focusing on tokenings, displays some features in common with each of these approaches but avoids their difficulties. On the one hand, it shows how indexicality and anaphora can be accommodated in an approach to intentional states that is centered on relations to linguistic items. On the other hand, its semantics is rooted directly in its pragmatics: the notions that do duty for abstract propositional contents are explained in terms of linguistic scorekeeping practices.

Davidson's invocation of samesaying suggests that the relation between the tokenings that matter for reports of intentional states and the linguistic performances that express them is to be explicated by appeal to an antecedent notion of expressing the same content. The idea here is rather that the theoretical explanatory work usually done by appeal to notions of semantic content is done instead by the notion of constellations of tokenings, structured by the commitments (inferential, substitutional, and anaphoric) that link those tokenings. What it is for the significance of the occurrence of a tokening to be governed by a commitment is in turn explicated in terms of scorekeeping practices. For some purposes it is then acceptable to speak loosely about tokenings linked in such ways as "expressing the same semantic content." For other purposes, for instance where the asymmetric structure of anaphoric linkages is to the fore, assimilating these constellations of tokenings to equivalence classes by talk of each being related to one and the same content, the one it expresses, is not appropriate. Structures of tokenings governed by the same commitments bear a certain resemblance to classical notions of sentence and term, and they bear a certain resemblance to classical notions of the contents expressed by the use of sentences and terms, but the differences must be borne in mind. It is just these differences that distinguish the account offered here of propositional-attitude-ascribing locutions from both the purely sentential and the purely propositional approaches.<sup>33</sup>

#### 4. Introducing '*Claims That*'

The indirect discourse locutions considered so far are ascriptions that make propositionally explicit the attribution of speech acts that have the pragmatic significance of acknowledging propositionally contentful commitments. It is a small step to indirect discourse locutions that ascribe the

propositionally contentful commitments acknowledged by those speech acts. Thus a third form of regimented ascription can be introduced:

Ruskin claims that medieval Gothic architecture is aesthetically and morally superior to Renaissance architecture.

'Claims' here should be understood as short for 'is committed to the claim'. Just as ascriptions formed using the regimented 'says' (claimingly) attribute a pragmatic force or significance that distinguishes sentences said from those merely uttered, so one need not actually say (claimingly) that *p* in order nonetheless in the regimented sense to claim that *p*. For one can adopt the attitude of acknowledging a commitment without having occasion overtly to avow it. (For this reason the status being ascribed here by the technical 'claims' would be expressed more naturally in English by something like 'holds').

The most straightforward version of such an ascription makes it appropriate just in case the individual to whom the commitment is ascribed is *disposed* to acknowledge it assertorially if suitably queried or prompted. The sentence tokening following the 'that' then purports to specify the content expressed by a tokening with assertorial significance that the ascriptional target *would* produce under appropriate conditions, rather than the content expressed by an assertorial tokening the target actually *has* produced. One way this expressive role shows up as a feature of the regimented use of 'claims' is in the propriety of inferences of the form

*S* says that *p*,

therefore

*S* claims that *p*,

while the converse form of inference is not in general valid (that is, commitment preserving). 'Claims' used in this way ascribes doxastic commitments that are, according to the ascriber, acknowledged by the one who is committed. Thus it corresponds to one way in which 'believes' is used (in accord with the overall explanatory strategy of having deontic status and attitude do much of the theoretical work generally done by appeal to the notions of intentional state and propositional attitude): namely, the use in which attribution of belief is tied most closely to dispositions to avow.

The difference between indirect discourse ascriptions regimented with 'says' and those regimented with 'claims' consists in a difference in how the reported tokening is to be understood as related to the target of the ascription. In the former case, what is attributed is an actual avowal of a doxastic commitment—the overt expression in an assertorial speech act of the deontic attitude of acknowledging such a commitment. In the latter case, what is attributed is merely the disposition to such an avowal, and the tokening in question may be merely virtual—one that *would* be produced if the disposi-

tion *were* actualized. But the relation between these two tokenings—the reporting tokening in the scope of the 'that' clause uttered by the ascriber and the reported tokening whose assertional utterance, on the part of the one to whom the commitment is attributed, *would* be the avowal of the ascribed commitment—is not different in the case of 'claims' than of 'says'. That is, the expressive role of terms occurring in the scope of the 'that' clause is not different in these two cases.

The major difference in the circumstances and consequences of application of ascriptions of commitment that follows from relaxing the requirement from actual to merely dispositional avowal of the ascribed commitment concerns inferential relations among such ascriptions. Inferential relations among the propositional contents of claims are reflected in inferential relations among the ascriptions of commitments. Thus one may offer as justification for the ascriptional claim that

*S* claims that  $\Phi(t')$

the further ascriptional claims that

*S* claims that  $\Phi(t)$

and

*S* claims that  $t = t'$ ,

where these claims might be backed up in turn by ascriptions (regimented using 'says') of overt avowals of those commitments. To say this is not to say that interlocutors always acknowledge commitment to what follows from commitments they acknowledge, even where the inference involved appeals only to auxiliary premises to which they acknowledge commitment. They *ought* to acknowledge commitment to such consequences of their acknowledged commitments, and that means that in the absence of incompatible commitments, entitlement to

*S* claims that  $\Phi(t)$

and

*S* claims that  $t = t'$

entitles one as well to

*S* claims that  $\Phi(t')$ .

These are *entitlement*-preserving inferences, not *commitment*-preserving inferences, and are defeasible as such. Thus the claim above, together with the claim that

Ruskin claims that medieval Gothic architecture is the purest outward expression of admiration for the concrete, natural, vari-

able, and changeful, and Renaissance architecture is the purest outward expression of admiration for the abstract, artificial, uniform, and unchanging

give good reason for endorsing also the ascription

Ruskin claims that the purest outward expression of admiration for the concrete, natural, variable, and changeful is aesthetically and morally superior to the purest outward expression of admiration for the abstract, artificial, uniform, and unchanging.

If, based on things Ruskin actually says, the ascriber were to find equally good reasons to deny this latter ascription (without thereby also finding good reason to withdraw the two previous ascriptions on which it is based), the conclusion of the ascriptional inference should indeed be withdrawn—but the fault would lie with Ruskin, rather than with the ascriber.

### *5. Introducing De Re Ascriptions*

Three forms of regimented ascriptions of doxastic commitments and the performances that express them have been introduced in this section:

1. Direct discourse ascriptions of assertional speech acts, of the form:

*S says [p],*

as in

Wittgenstein says “Wovon man nicht sprechen kann, darüber muß man schweigen.”

2. Indirect discourse ascriptions of assertional speech acts, of the form:

*S says that p,*

as in

Kant says that the essence of Enlightenment is responsibility.

3. Indirect discourse ascriptions of acknowledged doxastic commitments, which are avowed overtly by assertional speech acts, of the form:

*S claims that p,*

as in

Ruskin claims that medieval Gothic architecture is aesthetically and morally superior to Renaissance architecture.

There are two distinctions in play in these sorts of ascriptions. There is the distinction between attributions of a speech act and attributions of deontic status or attitude expressed by such acts, which is marked in the regimen-

tation by the use of 'says' or 'claims'. There is also the difference between specifying the content of what is ascribed in the mode of direct or of indirect discourse, which is marked in the regimentation by the use of quotation marks, or of 'that' segregating a reporting sentence tokening. These are independent distinctions, so there is no reason not to recognize as well a fourth category:

4. Direct discourse ascriptions of acknowledged doxastic commitments, of the form:

$S$  claims  $\lceil p \rceil$ ,

as in

Mill claims, "The internal sanction of duty is . . . a feeling in our own mind: a pain, more or less intense, attendant on violation of duty."

Claims of the form " $S$  claims that  $p$ " express in propositionally explicit form the attribution of endorsement of what is expressed by a propositionally contentful *dictum*. They correspond to *de dicto* ascriptions of propositional attitude. The next step in reconstructing the way *implicit* scorekeeping attitudes of *attribution* of performances and statuses can be made *explicit* as *ascriptions* is to look at how *de re* content specifications are related to these *de dicto* ones. These can be formed from either ascriptions of speech acts or of commitments, yielding two further sorts of regimented ascription:

5. *De re* ascriptions of assertional speech acts, of the form:

$S$  says of  $t$  that  $\Phi(it)$ ,

as in

Johnson says of a man now otherwise forgotten that he was obscurely wise and coarsely kind.

6. *De re* ascriptions of doxastic commitment, of the form:

$S$  claims of  $t$  that  $\Phi(it)$ ,

as in

Gibbon claims of his most famous work that its English text is chaste, since all its licentious passages are left in the decent obscurity of a learned language.

This use of 'of' has already been discussed. Expressions occurring within its scope are those for whose use in specifying the content of the ascribed commitment or speech act the ascriber *undertakes*, rather than attributes, responsibility. For a scorekeeper to endorse  $S_2$ 's *de re* ascription:

$S_1$  claims of  $t$  that  $\Phi(it)$ ,

the scorekeeper must take it that there is some appropriate *de dicto* ascription

$S_1$  claims that  $\Phi(t')$ ,

and that it is proper to attribute to  $S_2$  a commitment that could be ascribed by endorsing

$S_2$  claims that  $t = t'$ .

The point is that in specifying the content of the doxastic commitment  $S_2$  takes it that  $S_1$  acknowledges, its inferential consequences have been extracted by conjoining it with substitutional commitments acknowledged by  $S_2$ , not necessarily  $S_1$ . The principle behind the regimentation is that one moves from the *de dicto*

$S$  claims that  $\Phi(t)$

to the *de re*

$S$  claims of  $t'$  that  $\Phi(it)$ ,

by adding the nonascriptional premise

$t = t'$ ,

while one moves to the *de dicto*

$S$  claims that  $\Phi(t')$

by adding instead the ascriptional premise

$S$  claims that  $t = t'$ .

Ascriptions that specify the content of the doxastic commitment (or the speech act that is the overt acknowledgment of such a commitment) in the *de re* way accordingly let the ascriber express explicitly the *attribution* of a doxastic commitment and also the *undertaking* of the substitutional commitments that license specifying its content by using the particular expressions that occur in *de re* position. The ascriber is not claiming that the one to whom the claim is ascribed would *acknowledge* commitment to the content as specified *de re*—only that the one to whom the claim is ascribed has in fact *undertaken* commitment to that claim in virtue of what that individual *would acknowledge*. One undertakes, not only those commitments one acknowledges and so is disposed to avow, but also their consequences. The expressions used to specify the contents of commitments ascribed *de dicto* determine (perhaps requiring suitable shifts of language, indexicals, and anaphoric chains) tokenings the one whose commitments are in question would (according to the ascriber) recognize as expressing commitments that individual acknowledges. The expressions used to specify the content of commitments ascribed *de re* determine tokenings the ascriber recognizes as expressing commitments that the one whose commitments are

in question would (according to the ascriber) acknowledge as expressed by some other (unspecified) tokenings. The difference is just whether the *expressive commitment*, to a particular form of words being a way of expressing the commitment in question, is *attributed* along with the doxastic commitment being ascribed or *undertaken* along with the ascriptional commitment itself. These expressive commitments can be thought of as hybrid deontic statuses compounded out of attributed doxastic (or practical) commitments and substitutional commitments, which can be either attributed (in the *de dicto* case) or undertaken (in the *de re* case). The appendix to this chapter shows how the regimentation presented here can be extended to include *iterated* ascriptions that mix *de dicto* and *de re* styles of content specification, as in

This textbook claims that Russell claims of one of my favorite philosophers that he claimed of the planet Herschel discovered that it did not exist,

and shows that the ascriptional idiom that results is *expressively complete* with respect to possible constellations of scorekeeping attitudes toward doxastic and substitutional commitments.

It is worth noticing in passing that the hybrid deontic attitude made explicit by *de re* ascriptions has a dual, and that locutions have been introduced to make this attitude explicit as well. *De re* ascriptions *attribute* a doxastic commitment, while *undertaking* or taking responsibility for the expressive commitment involved in using a certain form of words to specify its content. One can also *undertake* a doxastic commitment, while *attributing* rather than undertaking the expressive commitment involved in using a certain form of words to specify its content. The locution used to make this attitude explicit is *scare quotes*. Consider this dialogue:

A: The freedom fighters succeeded in liberating the village.

B: Those <sup>s</sup>'freedom fighters<sup>s</sup> butchered two-thirds of its inhabitants.

The second interlocutor is undertaking responsibility for a claim but is disavowing responsibility for using the term 'freedom fighters' to express that claim. (Much the same effect could be achieved by prefixing 'so-called' to the offending expression.) In this context, the scare-quoted expression (regimented by superscripted 's's) functions *anaphorically*, picking up the previous interlocutor's use of the expression (just as the tokening /it/<sub>B</sub> is anaphorically dependent on /the village/<sub>A</sub>). The point of scare quotes is to stipulate that one is talking about the same thing that someone else was (using an expression intersubstitutable with it, in the extended sense that includes expressions that are only repeatable as elements in anaphoric structures), while disavowing responsibility for the propriety of using the expression that individual used to pick out that common topic of conversation. The sense in which scare quotes are performing an expressive function comple-

mentary to that performed by *de re* locutions is evident in the fact that in these circumstances *A* should characterize *B*'s claim by an ascription such as

*B* claims of the freedom fighters [that is, of the ones who according to me, *A*, are freedom fighters] that they butchered two-thirds of the village's inhabitants.<sup>34</sup>

It may also be remarked that there is no necessity to choose between conveying the information that is the specialty of *de dicto* specifications of ascriptional contents and conveying the information that is the specialty of *de re* specifications of ascriptional contents. *De dicto* specifications convey terms that the one to whom the doxastic commitment is attributed would (according to the ascriber) acknowledge as expressing that doxastic commitment. *De re* specifications convey terms that the ascriber takes responsibility for as expressing that doxastic commitment. Both sorts of information can be conveyed by using an ascription such as

*A* claims of the ones who butchered two-thirds of the village's inhabitants, *as* freedom fighters, that they succeeded in liberating the village.

Here the effect of scare quotes is implicit in the significance of the 'as' clause: the ascriber is attributing, not undertaking, responsibility for the use of this expression. This 'of' . . . 'as' . . . 'that' . . . regimentation segregates the various words used to specify the content of the ascribed commitment, so as to mark clearly the expressive role being played by each—whose commitments determine the admissible substitutions that settle the significance of the occurrence of each of those expressions.

The regimentation can accordingly be of use in construing and disambiguating ascriptions in which a single expression is called on to perform both *de dicto* and *de re* expressive roles, where in "Alfred believes the man in the corner is a spy," 'the man in the corner' may "be doing double duty at the surface level—both characterizing Alfred's conception and picking out the relevant *res*."<sup>35</sup> That dual role can be indicated by regimenting this ascription as

Alfred believes of the man in the corner *as* the man in the corner that he is a spy.

It is significant that *B* can characterize *A*'s claim, not only by an ascription such as

*A* claims of the ones who butchered two-thirds of the village's inhabitants that they succeeded in liberating the village,

but also by one such as

*A* claims of the ones that he calls 'freedom fighters' that they succeeded in liberating the village.

Here *B* achieves the same effect as scare quotes: securing a common topic without undertaking expressive responsibility for using the words 'freedom fighters' to characterize the ones who are nonetheless being talked about. In this case that expressive power is achieved by using the *explicitly* anaphoric indirect description 'the ones that he calls (refers to as) "freedom fighters"'. Recall that Chapter 5 described the use of these anaphorically indirect descriptions and argued that what makes an expression mean *refers* is its use in forming them. This sort of definite description makes explicit what is implicit in the use of scare quotes.

Descriptions can also be formed from occurrences of singular terms whose expressive role is complementary to that marked in the regimentation by occurrence within scare quotes, namely terms occurring within the scope of the 'of' in *de re* ascriptions.<sup>36</sup> Thus both *A* and *B* can use the *ascriptionally indirect description*

the ones of whom *A* claims (believes) that they succeeded in liberating the village,

and both can endorse its intersubstitutability with

the ones of whom *B* claims (believes) that they butchered two-thirds of the inhabitants of the village.

They can then use these terms to establish the common communicational ground necessary for a debate about the warrant for a commitment to intersubstituting either of these expressions with 'the freedom fighters'. Since the use of the terms that occur in *de re* position in ascriptions is governed by the ascriber's substitutional commitments in the same way that the use of terms having primary occurrence in the expression of nonascriptional claims is, the commitments involved in using ascriptionally indirect descriptions are just those outlined in Section II of Chapter 7, in the original discussion of definite descriptions. Putting this discussion together with that of *de re* ascriptions yields conditions under which

$$t = !x(S \text{ bel } \Phi(x))$$

that is, *t* is the one of whom *S* believes (claims) that  $\Phi(\text{it})$ .

## V. EPISTEMICALLY STRONG *DE RE* ATTITUDES: INDEXICALS, QUASI-INDEXICALS, AND PROPER NAMES

### 1. From Weak *De Re* Ascriptions to Strong *De Re* Beliefs

In one fundamental regard the account offered in this chapter is located firmly within the mainstream of contemporary philosophical thought. The last three or four decades have seen the development of a consensus that *de re* ascriptions of propositional attitude express something of fundamental importance for understanding the intentionality or repre-

sentational dimension of thought and talk—what it is for them to be *of*, *about*, or *directed at* objects and states of affairs that are not (in general) themselves in the same sense of, about, or directed at anything else. In other respects, however, the discussion so far is bound to appear simply to have missed the point of much of the discussion that has taken place within the scope of that consensus. Quine initiated the modern phase of interest in *de re* ascriptions by distinguishing, in terms of the substitutional behavior of singular terms, *notional* from *relational* senses of ‘believes’—attributed by the use of *de dicto* and *de re* ascriptions respectively.<sup>37</sup> The distinction that has been reconstructed here concerns rather two ways of specifying the content of ascriptions employing one unitary sense of ‘believes’.<sup>38</sup> One of the root intuitions that Quine begins with is that one does not put oneself in a position to have genuinely relational beliefs about an object simply by mastering the use of some singular term that in fact refers to that object. Someone who believes only that the shortest spy is a spy has not yet managed to have a belief that is *of* or *about* a particular object in the sense Quine is after with his notion of relational belief. Yet according to the reconstruction offered here, if Rosa Kleb is the shortest spy, then anyone who believes that the shortest spy is a spy thereby counts as believing *of* Rosa Kleb that *she* is a spy. Quine thinks of the latter ascription as appropriately used only to report someone who is in possession of information that might be of interest to the FBI—believing of someone in particular that that individual is a spy—a distinction that someone who believes only that the shortest spy is a spy surely does not achieve.

In short, the notion of *de re* belief that has exercised philosophers since Quine has more epistemic oomph to it than the thin notion reconstructed here. Indeed Kaplan, in “On Quantifying In”—which represents the first great milestone after Quine in the development of thought on this topic—formulates what is in its technical form essentially the reading of *de re* ascriptions offered here, only to discard it immediately as inadequate on the basis of the considerations Quine had raised about the shortest spy. The phenomenon of belief *de re* taken to be of interest is one that involves being en rapport with a particular object in a stronger sense than merely having some way or other of denoting it. Still, Kaplan does formulate what Dennett aptly calls this “denotational” sense of ‘about’ as a general strategy for moving from a notional sense of belief, conceived in terms of relations to linguistic items, to a relational sense of belief, conceived in terms of relations to the nonlinguistic items they denote. The idea is in essence that *S* believes<sub>Rel</sub> of *t* that  $\Phi(\text{it})$  just in case there is some expression  $\alpha$  such that *S* believes<sub>Not</sub> [  $\Phi(\alpha)$  ] and  $\alpha$  denotes *t*.<sup>39</sup>

The present account derives from this by three developments. First, a more sophisticated account is offered of the underlying notional or *de dicto* belief ascription—an account that appeals to linguistic tokenings rather than types and is not directly quotational. Second, the notion of substitutional

commitments is used in place of an explicit denotation relation. Third, the whole account is set in a social-perspectival framework that explicates the shifts of perspective from one repertoire of commitments to another that are required to deal properly with the *iteration* of mixed *de dicto* and *de re* ascriptions. Having introduced the basic strategy that underlies a generically denotational reading of ‘about’, however, Kaplan rejects it, and the subsequent tradition has followed him in this regard.

Kaplan suggests that what is required to strengthen the flabby denotational sense of ‘about’ into a genuinely relational form is what he calls “a frankly inequalitarian attitude toward various ways of specifying” objects.<sup>40</sup> He explores the epistemically stronger varieties of rapport with an object involved not just in being able to use an expression that in fact refers to it but in knowing what object one is talking about. Not all terms that can be used to refer to an object are equally serviceable from the point of view of providing this sort of epistemic access. Even though the use of a particular term, say “the shortest spy,” counts in some sense as picking out a particular object, knowing what one is talking or thinking about by using that term requires being able to pick out the object by the use of one’s own resources in some narrower sense. In this connection Kaplan points to numerals and quote-names as privileged designators (“standard names”) of numbers and linguistic expressions respectively.<sup>41</sup> Though the expressions

the number of spheres that can touch any one sphere in a  
24-dimensional Leech lattice

and

the eleventh through thirteenth lines of the first of Rilke’s  
*Duino Elegies*

pick out a number and a linguistic expression respectively, and so can be used in the denotational sense to talk and think *about* them, in a relatively clear sense one nevertheless knows what one is thereby talking and thinking about only if one understands them as intersubstitutable with the expressions

196,560

and

und die findigen Tiere merken es schon / daß wir nicht sehr ver-  
läßlich zu Haus sind / in der gedeuteten Welt.

respectively.

Although Kaplan tries to develop a specific notion of an expression  $\alpha$  being a “vivid name” of object  $t$  for believer  $S$  that accords with broadly Quinean intuitions about relational belief, the general lesson that should be drawn from his efforts is that different relational senses of belief, different sorts of epistemic rapport with the objects of belief, correspond to different restric-

tions on the expressions believers have available to them to pick those objects out. Just which privileged class of designators corresponds to "knowing who" or "knowing which" turns out to vary with particular situations in a way that defies antecedent systematization.

Recall the discussion (in Section I of Chapter 7) of the *Grundlagen* criterion that for an expression in the language to be properly understood as playing the role of a genuine singular term, and so as picking out a particular object, it must be understood as intersubstitutable with some other term. It was then argued (in Section II of that chapter) that to take the expression to pick out an object that exists in a particular sense (for example numerically, physically, or fictionally) is to take it that it is intersubstitutable with some term that is privileged as canonical with respect to that sort of existence. One can adapt the *Grundlagen* criterion applied there to the language as a whole and apply it to individual believers. The result is a minimal criterion of being able to use a singular term to put oneself cognitively en rapport with a particular object. To do so the believer must endorse *some* nontrivial identifying belief about it. That is, for some singular term ' $t$ ' to figure for a particular believer as a name of an object in a sense stronger than the merely denotational, the believer must endorse *some* (true) nontrivial identity claim relating it as intersubstitutable with another term.

This minimal inegalitarian attitude would rule out ascriptions of strong *de re* belief regarding the shortest spy that are based only on the belief that the shortest spy is a spy. According to this criterion, however, if Holmes believes the murderer left the footprints, and so has two ways of picking out that individual, he will count as having beliefs that are of that individual in a stronger sense than the inspector, who can refer to him only as "the murderer"; yet such a situation surely falls short of Holmes knowing who the murderer is. The claim is that beyond this minimal criterion there are many stopping places, many sorts of demands that can be made on the other termini of nontrivial identities that have as one of their termini tokenings that the individual whose attitudes are being reported is willing to use. In an important article Sosa follows out the idea that matching intuitions about relational belief requires that the substituend denoting the object be a "distinguished term." His persuasive and influential conclusion is that in order to serve this function, the notion of distinguished term will have to be a "wholly pragmatic matter, which can change radically from one occasion to the next."<sup>42</sup> One stronger demand has been of particular interest: requiring that one of the terms the believer can use to pick out the object—one of the expressions occurring in the nontrivial identity claim that identifies it for the believer—be used (or be anaphorically traceable back to one that is used) in the demonstrative or indexical way.

In the fifties, then, Quine distinguishes the attribution of merely notional belief from the attribution of genuinely relational belief. In the sixties, Kaplan shows how various sorts of relational belief can be understood in

terms of notional belief, by adding to the thin denotational model of aboutness restrictions to privileged classes of expressions available for use in specifying the content of the underlying notional belief. Sosa then showed how radically the appropriate restrictions must vary with circumstances of belief attribution and the interests of the attributors. In the seventies Burge, Perry, Lewis, and Kaplan himself, among others, focus attention on the kind of beliefs that are *de re* or relational in the epistemically strong sense that results from insisting that the believer be in a position to pick out the object of belief by the use of *demonstratives* or, more generally, *indexical* expressions.<sup>43</sup> Since that time much effort has been invested in the notion that "directly referential" expressions, paradigmatically indexical ones, make possible a fundamental sort of cognitive contact with the objects of thought, a kind of relational belief that is not conceptually mediated—in which objects are directly present to the mind, rather than being presented by the use of concepts.<sup>44</sup> According to this line of thought, what is of prime philosophical interest is this variety of epistemically strong *de re* beliefs. There are of course dissenters. Davidson, for instance, finds no use for anything but ordinary beliefs ascribed *de dicto*, and Dennett thinks (along lines somewhat similar to those pursued in the last two sections) that *de dicto* and (epistemically weak) *de re* ascriptions are essentially notational variants of one another, different styles of ascription of one kind of belief, and that insofar as the notion is intelligible at all, strong *de re* beliefs are of little significance.<sup>45</sup>

The view put forward here belongs to neither camp—neither to that of the boosters of nor to that of the scoffers at strong *de re* beliefs. It is distinguished from the scoffers in taking seriously the notion of epistemically strong *de re* beliefs. On the one hand, one should distinguish *de dicto* and *de re* styles of specifying the contents of ascribed propositional attitudes, as discussed thus far in this chapter. On the other hand, one should also recognize the existence of a distinct kind of epistemically strong *de re* beliefs—which deserve to be so called because of the way their distinctive contents derive from the sort of acquaintance with their objects that is expressed by the use of demonstratives and other indexicals. While sharing with them an acknowledgment of the existence of a distinct class of essentially indexical beliefs, the point of view adopted here nonetheless differs from that typical of the boosters of strong *de re* beliefs in its assessment of their theoretical significance. First, these beliefs do not form an autonomously intelligible sort or stratum of beliefs; one cannot coherently describe a situation in which this is the first or only kind of belief that is in play.<sup>46</sup> Second, although essentially indexical beliefs have a special sort of object-involving content that other beliefs do not, that object-involvingness should not be thought of as a *nonconceptual* element in their content; rather, the special sort of access to the objects their contents are about that the use of indexicals makes possible is a special kind of *conceptually articulated* access (though it is not correct to think of the role of the conceptual as *mediating* mind and its objects).<sup>47</sup> Finally, important

as essentially indexical beliefs are for our empirical knowledge and practical activity, what is of primary significance for understanding the representational dimension of thought and talk—its intentionality or aboutness in general—is the combination of doxastic perspectives expressed by (weak) *de re* specifications of the conceptual contents of beliefs of any sort, *not* the special sort of content possessed by strong *de re* beliefs, nor the special sort of rapport with objects they embody.

## *2. Essentially Indexical Beliefs: The Use of 'I'*

The content of a belief or doxastic commitment may be termed "essentially indexical" if it cannot be expressed in the form of a claim without using (or appealing to the use of) some locution as an indexical. Following Castañeda, Perry has argued persuasively that there are things that can be said using expressions such as '*I*', '*now*', and '*here*' that cannot be said by using nonindexical vocabulary.<sup>48</sup> I can of course refer to or describe myself in many ways, but no other term can do the expressive job done by '*I*'. For any other term '*t*' I might use to pick myself out, there would always be some possible circumstances in which I could believe that *t* had a property—was about to be eaten by a bear, was standing in the Stanford library, was leaving a trail of sugar on the floor of the supermarket, and so on—with thereby believing that *I* have that property. This difference can be manifest in my actions, since a belief I would express using '*I*' can be relevant to my practical reasoning and action in a way that no belief that can be otherwise expressed can match across the whole range of counterfactual situations. There will always be some of those situations in which I fail to realize that *I* am *t*. In those cases, I will fail to form intentions and to be motivated to act on the basis of beliefs about what *t* should do in the same way I would on the basis of beliefs about what *I* should do. Similarly, the belief that the meeting is starting *now* plays a different role in my behavioral economy than the belief that the meeting is starting at ten o'clock (or at any other time specified without the use of token-reflexives), since conditions can always be described under which I would fail to realize that the time so specified was *now*, as evidenced by my failure to form an intention-in-action.

In the idiom developed here, the key feature of the use of '*I*' that is not reproduced by other coreferential expressions (even those that are *necessarily* coreferential, since one could always fail to realize that they were) is its use in *expressing the acknowledgment of a commitment*. What '*I*' expresses is a potentially motivating acknowledgment of a commitment. The only expression that cannot counterfactually be separated from this motivational role is '*T*'. Acknowledgment of a practical commitment is the deontic attitude that corresponds to forming an intention—what is expressed by a sentence of the form "*I shall . . .*"

One of the central sorts of practical capacity involved in rational agency

consists of reliable dispositions to respond noninferentially to acknowledgments of some range of practical commitments by producing suitable performances.<sup>49</sup> But what one is trained to respond to in this way is what is expressed by sentences of the first-person form "I shall . . .," not what is expressed by sentences of the third-person form "*t* should . . ." The practical reasoning that leads to such efficacious acknowledgments of commitments to act also makes essential use of first-person pronouns:

*I am in danger of being eaten by a bear,  
so I shall run away.*

The identity of the terms appearing in the premise and the conclusion is essential to the commitment-preservingness of this form of practical substitution inference.

*The bearded man (or N.N.) is in danger of being eaten by a bear,  
so *I shall* run away*

is a good inference only insofar as it implicitly involves a commitment to the auxiliary premise

*I am (=) the bearded man (or N.N.).*

The fact that Perry points to—that no other expression can in general be substituted for '*T*' while preserving the role of the resulting sentences in practical reasoning and rational agency—is a consequence of the possibility of describing circumstances that ensure the agent will either lack the commitment expressed by this identifying auxiliary hypothesis or have commitments incompatible with it.

'*T*' finds its home language-game in acknowledgments of commitments to act, and secondarily in the expression of beliefs and desires that are directly relevant, as premises, to bits of practical reasoning that have formations of intention as conclusions. The central defining uses of '*T*' are not its uses in such sentences as "I can run the mile in five minutes" but its uses in "I shall open the door," as expressing the conclusion of practical deliberation, and therefore as used in the expression of the premises. Its primary use is as acknowledging a commitment. The lesson of Perry's examples is that one cannot replace '*T*' in its role in practical reasoning by any other designator, while preserving its motivating role in deliberative practical reasoning, under counterfactual circumstances. That is, although in particular circumstances '*T*' may be replaceable by other expressions, that is always defeasible by altering collateral beliefs and desires, in a way that is not paralleled for '*T*'. This is because what it expresses, the *undertaking* by *acknowledging* of a commitment, is something that no one but I can do. Locutions that do not function in practical reasoning in the way characteristic of first-person locutions (such as names and descriptions) can be used to *attribute* commitments to someone who may or may not (in different possible situations) turn out

to be, or to be known by the agent to be, that very agent. But *acknowledging* commitments is the basic way of *undertaking* them, and undertaking commitments cannot be reduced to *attributing* them, even to oneself. In the deontic scorekeeping social practices that institute commitments, these are two different kinds of practical doing. The phenomena that Perry displays in arguing for the ineliminability in favor of other thoughts of those expressed using 'I' is a manifestation of the irreducibility of the deontic attitude of acknowledging commitments to that of attributing them.

These same considerations can be seen to underlie the special features of the expressive role of 'I' that lead Anscombe, in her important meditation "The First Person,"<sup>50</sup> to deny that 'I' is a referring expression at all. From the point of view developed here, that dark claim appears to be an overstatement; the important thing is to specify how the use of 'I' differs from that of other referring expressions.<sup>51</sup> The line of thought that leads to the stronger conclusion could be put into the terms of this book something like this. If 'I' were a referring expression, it would refer to someone such that *attributing* a commitment (for instance, an intentional action or a belief) to that individual by its use would have the same practical significance as *undertaking* that commitment. One's *self* (= df. the object referred to by 'I') would just be the one, such that this sort of attributions to it *are* acknowledgments. Since one can show that no expression that could figure in attributions in this way could achieve the effect of acknowledgment, the conclusion is drawn that 'I' does not function as a referring expression at all. A less dramatic conclusion would be that it does not play a role analogous to that played by expressions that can be used merely to attribute commitments. The considerations Anscombe advances to distinguish the use of 'I' from that of other singular terms are somewhat different from, though related to, those Perry invokes.

The features she focuses on are:

1. One can have a special sort of nonobservational knowledge of one's own doings, of what one would express using 'I', and
2. The central sorts of claim one expresses using 'I' involve immunity from the possibility of misidentification.

These are closely related. Thus, the first point is that though I can be fooled in exceptional circumstances, in the usual run of things I do not need to *perceive* (visually, kinesthetically, or otherwise) my intentional bodily movements in order to know what they are. The second point is that when I acquire in this way a belief about what I am doing (about "actions, postures, movements, intentions," as Anscombe says),<sup>52</sup> there is no room (as there would always be were I merely observing) for a question as to *whose* it is I know about—no room for questions of the form "*Someone* is talking, sitting, perceiving a red triangle, trying to open the door, but is it *I*?" (Of course there are *also* cases in which questions of these kinds *are* in order, but they are precisely the ones where I have only observational knowledge of what I am

doing and how it is with me.) I cannot in these cases *misidentify* myself, accidentally pick out the wrong object, using 'I', as I can with *any* other way of referring to myself. (It should be clear that contact is made at this point with the considerations and counterfactuals that Perry points to.) It is easy to become mystified, thinking about what sort of peculiar object selves must be, and what sort of peculiarly intimate contact we must have with them, in order to be able to use first-person expressions that have these special properties. And it can be tempting to think that these features of 'I' show something deep about the difference between selves or subjects and about the objects that we can merely observe and refer to. But it is not obligatory to take this Cartesian, ontological road. The question is how best to understand the expressive role of 'I'.

Anscombe teaches us that the possibility of nonobservational knowledge expressed using 'I' that precludes the possibility of misidentification shows the inadequacy of an account of 'I' as the "word each of us uses to speak of [refer to] himself."<sup>53</sup> Anscombe exhibits the difference by considering a language (the 'A' language) in which an expression ('A') is used by each interlocutor just to refer to that very interlocutor. In the 'A' language, as opposed to our 'I' language, one must observe in order to know who ('A', or someone else) is doing something, and one is for that reason always liable to be mistaken about who really is doing it. One way to understand the particular sort of epistemically strong *de re* belief that can be expressed by using 'I' is to add practices to the 'A' language in stages, until 'A' has the fundamental features characteristic of 'I'. Anscombe describes the starting point: "Imagine a society in which everyone is labelled with two names. One appears on their backs and at the top of their chests, and these names, which their bearers cannot see, are various: 'B' to 'Z' let us say. The other, 'A', is stamped on the inside of their wrists, and is the same for everyone."<sup>54</sup> Individuals use chest-and-back names for each other and learn to respond to them as their own, but observations of one's own doings are reported by using the wrist-label 'A'. By hypothesis 'A' is used only on the basis of observation: tracing a bodily connection between the limbs whose motions are being reported and a wrist with an 'A' stamped on it. Thus misidentification is possible, for sometimes it will be someone else's token of 'A' that one glimpses. (It may be supposed that where one is in this way entitled to use 'A' in a report, one thereby becomes *inferentially* entitled to make the report that differs from the original in having the chest-and-back label one has learned to respond to as one's own substituted for 'A'.)

These labels can be used in keeping deontic score. So suppose that Anscombe's 'A' speakers engage in the sort of linguistic practices described in Chapter 3 and 4. They make claims, and they can make some of them noninferentially, as expressions of, or reports on, perceptions. These they can sometimes take-true by exercising a reliable differential responsive disposition, responding to a fact or state of affairs by a performance that has the

pragmatic scorekeeping significance in the community of acknowledging a doxastic commitment. A scorekeeper treats another interlocutor as a reliable observer (of a particular range of states of affairs under certain circumstances) by attributing entitlement to reports with those contents issued by that interlocutor under those circumstances. But at this stage there need be no way explicitly to *claim* reliability, for instance as a justification in response to a challenge by someone who is committed (noninferentially or not) to an incompatible claim.

One of the things they can observe each other doing is making observation reports. Suppose then that explicit ascriptions of claims are added to their vocabulary. Here it will be helpful to consider a special variety: ascriptions of reports of observations. These ascriptions will have inferential uses, but the uses of primary concern are noninferential. Thus *T* may report: "*S* claims-perceptually that *p*." A deontically hybrid ascribing locution could then be introduced so that if *T* also is disposed, noninferentially or otherwise, to endorse the claim that *p*, *T* could say, "*S* perceives (sees) that *p*." *T* could now respond to a challenge by another to *S*'s observation report that *p* by saying, "*S* perceives that *p*." This response is also open to *S*, who in the basic case will phrase the same response as "*A* perceives that *p*" (for by hypothesis *S*'s self-application of '*S*' is always inferential, based on a noninferential disposition to use '*A*').

At this stage, scorekeepers learn to treat the claim "*A* perceives that *p*" as in order whenever a perceptual claim that *p* is in order—commitment or entitlement to the '*A*'-ascription of perception that *p* is attributed in just those cases in which commitment or entitlement to the underlying perceptual report that *p* is attributed. (Note that a scorekeeper can take a reporter to be *entitled* to the claim that *p*, and equivalently to the self-ascription of a perception that *p*, even if the scorekeeper does not endorse the claim that *p*—that is, does not take it to be true; attribution of entitlement depends on taking the reporter to be reliable, not to be infallible.) "*A* perceives that *p*" commitment-entails "Someone perceives that *p*" and is commitment-entailed by "Everyone perceives that *p*," in just the same way that "*S* perceives that *p*" does.

For the next stage, suppose that this functional equivalence of perceptual reports that *p* and self-ascriptions of perceptions of the form "*A* perceives that *p*" is extended so that "*A* perceives that *p*" becomes available as an alternate *noninferential* response to the same situation that would previously have prompted only a report to the effect that *p*. By hypothesis, the folk in question can already respond differentially to enveloping situations by producing the claim that *p*, so all that is being imagined is that they add as another possible response the claim "*A* perceives that *p*." Although clearly possible, this addition to the grammar of '*A*' makes a substantial difference. For the first time, '*A*' is given a use that does *not* depend on observation of a labeled wrist. In its initial use, "*A* perceives that *p*" still required tracing

in observation a connection between the mouth issuing the perceptual report that  $p$  and a wrist marked with an 'A'. It was accordingly liable to failures of identification on those occasions when reporters could see someone else's wrist. That possibility does not exist for the noninferential *responsive* use of 'A' in ascribing perception. Here 'A' is applied nonobservationally—in the sense that while one must observe that  $p$  in order to be entitled to the claim "A perceives that  $p$ ," one need not *in addition* observe a wrist inscribed with an 'A'. Perceptual reports of the form "S perceives that  $p$ " always leave room for misidentification, but perceptual reports of  $p$  in the form "A perceives that  $p$ " do not. They do not simply because there is no question about whether the dispositions one is exercising or evincing are one's own—any more than when an environing situation noninferentially elicits the report that  $p$  there can be a question as to whether claiming that  $p$  is the performance one is disposed to produce in response to that situation.

What has been done for reliable differential dispositions to respond noninferentially to a situation by acknowledging a *doxastic* commitment, in *perception*, can also be done for reliable dispositions to respond noninferentially to the acknowledgment of a *practical* commitment by producing a performance, in *action*. To be agents, the 'A' creatures must learn to respond to the acknowledgment of some kinds of commitments by making-true some claims—just as to be perceivers, they must learn to respond to some facts by taking-true some claims, that is, by acknowledging some kinds of commitments. Interlocutors can observe each other's nonlinguistic intentional performances just as they can their perceptual reports. Suppose then that locutions are introduced making it possible to *asccribe* performances to individuals as intentional. These are *agentive* locutions, which may be represented for present purposes as having the form "S does  $q$ ." For S to do  $q$  in this sense is for its being true that  $q$  to be brought about by the exercise of a reliable disposition to respond differentially to S's acknowledgment of a practical commitment. As in the case of perception, these locutions are first used with chest-and-back names and then extended so as to give a use to "A does  $q$ ," in the case where the one making that claim can observe the doing and trace the movements it involves to a wrist on which 'A' is inscribed. Thus at this stage the use of 'A' is still purely observational and liable to misidentification. At this stage, scorekeepers learn to treat S's claim "A does  $q$ " as in order whenever the scorekeeper would endorse the ascription "S does  $q$ "—those are the circumstances in which S is, according to the scorekeeper, entitled to the commitment expressed by "A does  $q$ ." "A does  $p$ " commitment-entails "Someone does  $p$ " and is commitment-entailed by "Everyone does  $p$ ," in just the same way that "S does  $p$ " does.

For the next stage, then, suppose that limited scorekeeping equivalence of attributions of doings and of self-ascriptions of doings is extended in practice by interlocutors learning reliably to respond to practical commitments to do  $q$ , not only by doing  $q$ , but also by undertaking doxastic commitment to the

self-ascription "A does *q*." Again, this is just adding to an already-established differential responsive disposition an alternate response that the creatures are already capable of producing in other circumstances. Here again, a substantial difference in the use of 'A' is being envisaged: one need not see a wrist label in order to use 'A' this way. The result is that although it is possible for *S* noninferentially to claim "A does *q*" without in fact doing *q*—for instance when the agentive ascription in response to acknowledgment of the practical commitment is not prevented, but doing *q* as a response is prevented—in the cases where reliability prevails and both responses are elicited by that acknowledgment, *S* will have *nonobservational* knowledge of those doings and will express such knowledge by the 'A' ascription. Furthermore, uses of "A does *q*" that are elicited in this way are immune from errors of misidentification, as those based on observation of a labeled wrist were not.

At this point the initially impoverished and unselfconscious speakers of the 'A' language have a rudimentary but recognizable first-person pronoun. They have, in favored cases, immediate, nonobservational knowledge of their own doings and perceivings. They are not incorrigible about them, for they cannot guarantee the truth of what they take themselves to perceive or the success of what they take themselves to do, but when they are right, they *know* they are right: scorekeepers within the community will attribute to them a hybrid deontic status that corresponds to justified true belief. The claims about their perceivings and doings that, if true, express this sort of immediate, nonobservational knowledge are immune from errors of misidentification. Challenges to *S* of the form "Were you right to say 'A perceives . . .' or 'A does . . . ? Perhaps it was really *T*!'" will not be in order, except as queries about whether the warrant for the claim really was observational—that is, elicited by the inspection of wrist labels rather than by the exercise of nonobservational responsive dispositions. The *immediacy* of these acknowledgment-expressing uses of 'A'—as either noninferentially elicited concomitantly with the acknowledgment of doxastic commitments ('perceives') or noninferential responsive consequences of acknowledgments of practical commitments ('does') elicited concomitantly with actions—ensures that the question of whether one should not look more closely and see whether it is someone else's commitments that are at issue cannot arise.

Finally, this fable promises to make intelligible not only Anscombe's but Perry's observations about the use of 'T', though only the briefest sketch can be offered here of how this story would go. What one expresses using 'T' plays a special motivational role. 'Does' as described here in effect attributes the intentions-in-action into which prior intentions mature. It would be possible to extend this discussion by introducing next ascriptions of *prior intentions*, which would lead to self-ascriptions of the form "A shall do *q*." When acknowledgments of these ascriptions come to be among the noninferential responses to acknowledgments of practical commitments regarding future doings, they express nonobservational knowledge of one's intentions, rather

than just of one's actions. (Discussion of the motivational role of 'now' that Perry points to would require going further into the expressive role it plays in connection with these ascriptional locutions in making explicit the practical reasoning implicit in the maturation of prior intentions into intentions-in-action.) The final stage is reached when training secures it that when one says 'I shall' or uses 'I' in substitution inferences that lead to that conclusion, then one is disposed to respond to (the acknowledgments that are made explicit in) such concomitant expressions by producing a performance. There will be no other expressions, *S* or *T*, which *de jure* have this significance. For each of them, it is only if something that means "I am (=) *S*" is endorsed that the motivational significance is guaranteed to be in play, just as in Perry's examples. The fable shows just how "*A* perceives a bear" or "*A* is making a mess" can have a significance in practical reasoning across various possible situations that is not matched by any of the chest-and-back labels, or indeed any other designators that do not play the expressive role with respect to perception and action that has been built into the use of '*A*' by the end of the story.<sup>55</sup>

### 3. Anaphoric Quasi-Indexicals and the Ascription of Essentially Indexical Beliefs

In this way, then, it is possible to understand the basic outlines of the expressive role played by the first-person pronoun 'I'. There is no necessity that every idiom include an expression playing that role; there need be nothing incoherent in descriptions of communities of judging and perceiving agents, attributing and undertaking propositionally contentful commitments, giving and asking for reasons, who do not yet have available the expressive resources 'I' provides. Yet when this bit of vocabulary is available, it makes explicit something important about such a set of discursive practices. 'I' is a logical locution. While those who lack it can be *conscious* in the sense of *sapient*—can be explicitly aware of things by making judgments about them—they are in an important sense not yet *self-conscious*. One of the normative social statuses instituted by any scorekeeping practices that qualify as discursive is that of being an individual *self*: a subject of perception and action, one who both can *be* committed and can *take* others to be committed, a deontic scorekeeper on whom score is kept. Selves correspond to coresponsibility classes or bundles of deontic states and attitudes—an indispensable individuating aspect of the structure of scorekeeping practices that institutes and articulates discursive commitments.

The notion of *one* performer who is responsible for *two* different claims is implicit in the practical acknowledgment of relations of inferential consequence among claims. One interlocutor is not responsible for the inferential consequences of commitments undertaken by another; such consequence relations govern only the commitments of a single interlocutor. The notion

of incompatible propositional contents similarly presupposes the assignment of responsibility for commitment to those contents to reidentifiable interlocutors. There is nothing wrong with one interlocutor being entitled to commitment to a content incompatible with a content to which another is committed; what is forbidden is that one and the same interlocutor should count as entitled to a commitment incompatible with another claim to which that same interlocutor is committed. Asserting, challenging, deferring, justifying—all these significances that performances can have according to basic discursive scorekeeping practices depend on sorting commitments and performances into the concomitancy classes corresponding to individuals.

That two commitments are to be assigned to the same individual is accordingly a fundamental social status instituted by those scorekeeping practices. It is not arbitrary how such commitments are sorted into equivalence classes, for the reliability of the responsive dispositions that make possible perception and action, and hence the empirical and practical aspects of conceptual contents, are tied to the careers of ongoing organisms. The sorting depends on the bodily relations between the retina that is irradiated and the mouth that can be trained to make reliable reports as responses to what was seen, between what that mouth acknowledges and the limbs that can be trained reliably to move in response to that acknowledgment, and so on. The empirical and practical abilities presupposed by our conceptual contents essentially involve locating beliefs—connecting for instance the public coordinate addresses appealed to by commitments to physical existence and the egocentric space induced by the use of demonstratives<sup>56</sup>—that depend in turn on the fact that our bodies trace out continuous paths in space.

'I' plays the substitutional role of a singular term; it is a pronoun and so should be counted, when correctly used, as giving us access to an object. Selves are precisely the kind of thing it gives us access to. But so would chest-and-back names, provided they played the proper role in scorekeeping practices. The difference lies in the sort of privileged access that 'I' gives us to those objects. The sort of self-consciousness that the use of first-person pronouns makes possible is not available through the use of any other sort of expression. Thus what can be expressed only by the use of 'I' is an example of a special kind of conceptual content—one that is *essentially indexical*. Beliefs with this sort of content embody a special sort of acquaintance with their objects; they are *de re* in an epistemically strong sense. Doxastic and practical commitments that are explicitly expressed (in the sense of overtly acknowledged by the production of propositionally contentful speech acts) by the use of other indexical expressions such as 'now' and 'here', including demonstrative ones such as 'this' and 'that', can also be strongly *de re* in this sense. The sort of cognitive access to their objects that they incorporate is expressible only by the use of indexicals; the role they play in the conceptual economy of the believer cannot be duplicated by beliefs expressible entirely in nonindexical terms.

Beliefs that are about their objects in the epistemically strong sense characteristic of this kind of *de re* beliefs are essentially indexical; being able to use an indexical expression to pick out a particular object depends on being in the right circumstances or context (in the case of 'T', it depends on *being* the right object). Thus one must be in the right circumstances or context in order to have strong *de re* beliefs (or intentions) about an object. Scorekeepers cannot pick each other out by using 'T', and it is only rarely that in assessing and attributing the commitments of another a scorekeeper will be in a position to pick out demonstratively all the objects with which the other was demonstratively acquainted. What constraints do these facts impose on the *communication* of the strong *de re* beliefs expressed by the use of indexicals? Does the fact that demonstrative references are only *accidentally* shareable across interlocutors, and that first-person references *never* are so shareable, mean that essentially demonstrative thoughts are only accidentally shareable, and first-person thoughts never are? Can the audience keeping score on another's commitments understand the essentially indexical commitments that the interlocutor undertakes, in cases where the audience is not (whether *de facto* or *de jure*) in a position to pick out the same objects indexically? Are strong *de re* beliefs strictly or largely incomunicable?

Concluding that they are is adopting what Perry calls the "doctrine of limited accessibility." He takes the collateral costs of such a theoretical commitment to be sufficiently high to motivate resisting this doctrine and suggests a strategy for doing so. His way out is to distinguish two components of belief: a belief *state* exhibited by the believer and a belief *content* exhibited by that state. These correspond to Kaplan's notions of *character* and *content*. Two different people who each believe something they could express by asserting the sentence "I am threatened by a bear" share a belief state, though the contents of their beliefs (who it is they take to be so threatened) differ. Michele, who believes "I am threatened by a bear," and Nicole, who believes "Michele is threatened by a bear," are in different belief states, which share a content. Thus belief states are shareable, and belief contents are shareable; what is not shareable is their conjunction: only Michele can have a belief with that particular content by being in the belief state she could express by using the sentence "I am threatened by a bear." The question then is what constraints the fact that sharing this conjunction of state and content may be impossible places on the theoretical understanding of *communication* of the state one is in and the content that it has. Clearly in some cases one understands the commitments of another only if one understands both the belief *state* the other is in and the *content* of that belief. But both of these are shareable. What needs to be worked out is an account of the sense in which being able to share these individually suffices to make intelligible (hence communicable) their conjunction, which is not shareable.

McDowell is led to a different attitude by these same considerations

concerning essentially indexical beliefs: he endorses the doctrine of limited accessibility toward strong *de re* beliefs but denies that doing so incurs the cost of making it unintelligible how they can be understood by and communicated to others. No one else can have the thought that Michele would express by saying, "I am threatened by a bear," but that does not mean that no-one else can understand what thought she would be expressing by that claim. He says: "Frege's troubles about 'I' cannot be blamed simply on the idea of special and primitive senses; they result, rather, from the assumption—which is what denies the special and primitive senses any role in communication—that communication must involve a *sharing* of thoughts between communicator and audience. That assumption is quite natural, and Frege seems to take it for granted. But there is no obvious reason why he could not have held, instead, that in linguistic interchange of the appropriate kind, mutual understanding—which is what successful communication achieves—requires not shared thoughts but different thoughts which, however, stand and are mutually known to stand in a suitable relation of correspondence."<sup>57</sup> So McDowell proposes that different *de re* senses can be sorted in two ways: one corresponding to Kaplan's characters (according to which you and I express senses of the same sort by saying 'T'), and one corresponding to my use of 'T' and your use of 'you'. Successful communication requires only that the audience be able to associate an appropriately corresponding *de re* sense with the speaker's utterance. It is on this basis that he retreats from his earlier endorsement of *samesaying* as the relation between reported and reporting tokenings in (*de dicto*) ascriptions of such beliefs and their expressions. That the senses expressed by 'T' in my mouth and 'you' in yours, or by 'today' uttered today and 'yesterday' uttered tomorrow, correspond in the right way is conceived as a requirement stronger than mere coreference but weaker than identity of sense or content.

Clearly the facts about essentially indexical or epistemically strong *de re* beliefs can be rendered in either way. Employing the machinery of the previous section makes it possible to describe more precisely the phenomenon that each of these approaches renders in general terms. For the sort of understanding that is the cognitive or semantic uptake that must be appealed to in defining the notion of successful communication can be expressed in scorekeeping terms by the requirement that the scorekeeping audience interpreting the performances of the speaker or agent be able to *attribute* the very same commitments that the performer *undertakes* or *acknowledges* by those performances. If one wishes to construe such understanding on the model of shared contents, then it is this pragmatic scorekeeping condition that determines what counts as grasping the *same* conceptual content; if one wishes to construe it instead on the model of grasping (in the cases at issue strongly *de re*) senses of the same sort ("corresponding senses"), then it is this pragmatic scorekeeping condition that determines what counts as those contents being of the *same* sort. So from the point of view of deontic scorekeeping,

the phenomenon to focus on is that of *attributing* essentially indexical commitments. This is an implicit practical capacity; as always, the best route to theoretical understanding of something *implicit* in a practice is by consideration of the use of the expressions by which it is made *explicit*. Since attributions are made propositionally explicit in the form of *ascriptions*, this methodological precept dictates a look at *ascriptions* of essentially indexical commitments.

Transposed into these terms, the problem becomes that of understanding how an ascriber can specify the content of an ascribed commitment, even though no tokening in the mouth of the ascriber could have just the same significance as a tokening of, say, "I am threatened by a bear," in the mouth of the one to whom the commitment it expresses is to be ascribed. So put, the problem has a straightforward solution. Castañeda is concerned that the ordinary use of a *de dicto* ascription such as "Michele believes that she is threatened by a bear" does not distinguish between its use to report the belief Michele could express as "I am threatened by a bear" and other beliefs she might express by making such claims as "Michele is threatened by a bear" or "The woman under the pine tree is threatened by a bear."<sup>58</sup> He suggests that the genuinely first-person belief might be reported by some such idiom as "Michele believes that *she herself* is threatened by a bear." Anscombe points out that Greek, among other languages, contains special *indirect reflexive* constructions, which are used precisely to ascribe the use of reflexives to others.<sup>59</sup> Castañeda suggests regimenting the ascription of essentially indexical propositional attitudes by the use of what he calls "quasi-indicators," which are stipulated to perform this expressive function. In his notation, what is expressed in the informal example above by 'she herself' is expressed by the use of the quasi-indicator '*she\**', so that the ascription "Michele believes that *she\** is threatened by a bear" is correctly used only to ascribe the first-person belief Michele could express by using "I am threatened by a bear."

An analogous difficulty arises for the reporting of other essentially indexical beliefs, notably those the believer would express using *demonstratives*. As was pointed out in the discussion of the expressive role of 'that' clauses in specifying the contents of *de dicto* ascriptions, one of the characteristics that distinguishes indirect discourse from direct discourse is that indexical and demonstrative expressions appearing in them are used by the ascriber, and so evaluated according to the context of the ascriber's speech act, not that of the speech act ascribed. Thus ascriptions such as "Danielle believes that *this* is an interesting painting" do not settle whether the belief ascribed is one Danielle is in a position to express by claiming "*This* is an interesting painting," rather than by using some nondemonstrative singular term, perhaps by claiming "Edward Hopper's *Excursion into Philosophy* is an interesting painting"—a belief one could have without knowing which painting one was talking about, in the strong sense of 'knowing which' that goes with

demonstrative acquaintance. Castañeda's quasi-indicators can mark this distinction as well: "Danielle believes that *this\** is an interesting painting" ascribes the use of a demonstrative without using one.<sup>60</sup>

How do quasi-indicators perform this expressive function? Castañeda notes that tokenings functioning as quasi-indicators are not used demonstratively and that they have antecedents, but he does *not* claim that they are anaphoric dependents. The only reason he gives is a bad one: they are not replaceable by their antecedents. But it was pointed out already in Chapter 5 that anaphoric dependents are not in general replaceable by their antecedents; syntactically lazy anaphora is only one species. The analysis in Chapter 7 of asymmetric repetition structures of tokenings in terms of inheritance of governing substitutional commitments showed how anaphora can be understood more generally. In this context, the indirect indexical constructions Castañeda regiments as "quasi-indicators" can be understood to be functioning in a straightforwardly anaphoric fashion. They differ from ordinary anaphoric dependents in the extra information they carry about their antecedents—namely, the antecedent of a tokening of *this\** is a demonstrative of the sort the ascriber would express using 'this'.<sup>61</sup>

There is no special difficulty in understanding how anaphoric proforms could convey such added information; pronouns that do so are more common in natural language than purely anaphoric forms. Thus the use of 'he' carries the added information that its antecedent picks out an animate organism, and a male one. In many languages, of course, the gender information conveyed by a pronoun concerns the antecedent itself, rather than what that antecedent refers to. Anaphorically indirect indexical expressions are to be understood as working the way these pronouns do, except that the additional information they convey concerns not gender and number but the status of the antecedent as an indexical. In this way they can be compared to anaphorically indirect definite descriptions such as "the one he referred to as 'that airhead,'" which also convey information about the expression on which they are anaphorically dependent.

To add the expressive power of these expressions to the regimented ascriptions of the previous section, let  $\lceil a^{(t)} \rceil$  stand for an expression that is anaphorically dependent on a tokening of type  $\langle t \rangle$ , such as 'he<sup>(1)</sup>' (which expresses what Castañeda's "he himself" does).<sup>62</sup> This usage differs from Castañeda's in that where he simply indicates that there is an indexical or demonstrative antecedent by using a '\*\*', here the type of that antecedent is explicitly specified in the superscript. 'Quasi-indicator' is Castañeda's technical term, and he does not use this convention or understand the functioning of the expressions to which it applies in anaphoric terms. Thus it seems well to adopt a different term. They might be called 'anaphorically indirect indexicals', but this would suggest that they are a kind of indexical, whereas they are really a kind of anaphoric dependent. So the term 'quasi-indexicals' will be used here, as a reminder of the origin of this category in Castañeda's quasi-indicators, with the difference being the anaphoric analysis of their expressive function.

The addition of these expressive resources makes it possible to ascribe essentially indexical beliefs as such. "Danielle believes that  $\text{it}^{(\text{this})}$  is an interesting painting" is a *de dicto* ascription of such an epistemically strong *de re* belief. Since what a strong *de re* belief is a belief *of* or *about* is essential to its having the content that it does (see the discussion below of the "object-dependence" of this sort of belief), a full specification of its content requires also a *de re* specification.

Danielle believes that  $\text{it}^{(\text{this})}$  is an interesting painting, and the painting she believes that of or about is Edward Hopper's *Excursion into Philosophy*.

This is a *de dicto* ascription conjoined with an identity one side of which is an ascriptionally indirect *de re* definite description. It could be expressed in a single mixed *de dicto/de re* ascription by something like:

Danielle believes of Edward Hopper's *Excursion into Philosophy* that  $\text{it} = \text{it}^{(\text{this})}$  is an interesting painting.

The basic effect of this mixed ascription can be achieved without the use of quasi-indexicals by employing the full "... of . . . as . . . that . . ." structure mentioned at the end of the previous section. Thus the ascription

Danielle believes of Edward Hopper's *Excursion into Philosophy*, as 'this', that it is an interesting painting<sup>63</sup>

specifies the content of the ascribed belief in full. Understanding the anaphoric function of both quasi-indexicals and indirect definite descriptions formed using 'refers' makes it possible to appreciate that this is because that ascription has the sense of

Danielle believes of Edward Hopper's *Excursion into Philosophy*, which she refers to by using a tokening of type ⟨this⟩, *that* it is an interesting painting.

What have been considered so far are *de dicto* ascriptions of strong *de re* beliefs. The use of quasi-indexicals also permits a straightforward analysis of *de re* ascriptions of beliefs that are strongly *de re* in the sense of being essentially indexical or demonstrative.

$S$  believes of<sub>strong</sub>  $t$  that  $\Phi(\text{it})$

is an ascription expressing the perspectively hybrid deontic attitude that consists of the *attribution* that would be expressed by the *de dicto* ascription

$S$  believes that  $\Phi(\text{that}^{(\text{this})})$ ,

and the *undertaking* of the substitutional commitment that would be expressed by the identity

$\text{that}^{(\text{this})} = t$ .<sup>64</sup>

This understanding just adapts the reading of weak *de re* ascriptions by combining it with a restriction on the terms the ascriber takes it the believer could use to express the belief. Such an account combines:

- The denotational reduction of weak *de re* to *de dicto* ascription;
- Kaplan's and Sosa's idea that stronger sorts of rapport with objects correspond to a “frankly inequalitarian attitude toward different ways of specifying objects”;
- Perry's point that there are essentially indexical or demonstrative beliefs, and Burge's suggestion that it is just these that are strongly *de re*;
- Davidson's and McDowell's paratactic account of *de dicto* ascriptions in terms of relations between the reporting tokening, in the ascriber's mouth, and the (possible) reported tokening, in the mouth of another, which is or expresses what is ascribed.

The final theoretical ingredient, which serves to combine all the rest, is then

- An anaphorically rendered notion of quasi-indexicals, whose expressive role in ascriptions is modeled on Castañeda's quasi-indicators and indirect reflexives.

With these expressive resources it is possible even for an interlocutor who cannot have that 'T thought and that demonstrative thought to ascribe *exactly* the beliefs that Michele would express using the first-person pronoun "I am threatened by a bear"<sup>65</sup> and that Danielle would express using a demonstrative "This is an interesting painting." What is required is only that the reporting tokening occurring in the ascription determine in a systematic fashion what token is being reported—either the utterance being ascribed or the tokening that would express the state being ascribed. For that it is not necessary that the content-specifying token in the ascription be one that *would*, if uttered by the ascriber as a freestanding assertion, undertake the very same commitment as is being ascribed. That samesaying relation between the reported and the reporting tokening holds in the case of *de dicto* ascriptions of ordinary beliefs, but not of essentially indexical ones. This is the point McDowell makes in the passage cited above. The discussion of quasi-indexicals has shown how to make precise his notion of suitably related corresponding thoughts or *de re* senses. The relation in question is stronger than mere coreference in just the way that anaphoric dependency is stronger than coreference: a tokening used quasi-indexically and the token used indexically that is its anaphoric antecedent are members of one single token-repeatability structure.

The quasi-indexicals used in the *ascription* of essentially indexical beliefs are *logical vocabulary*. They make explicit the implicit practical understanding of another's indexical beliefs that is required for their *attribution*, and therefore for their *communication*. They show what is required to attrib-

ute *exactly* the belief that another undertakes, even in the cases where the essentially indexical character of the belief makes it impossible, in principle or in practice, for speaker and audience, believer and scorekeeper, to *share* that belief. In the least demanding extensional or representational sense, communication requires securing only coreference. Further grasp of what is expressed by the terms the speaker did (or the believer would) use is not needed. In this sense it is sufficient for an interpreter to understand another's remark, and so for communication to count as successful, if that scorekeeper adopts a deontic attitude of attributing a commitment that would be undertaken explicitly by asserting a (weak) *de re* ascription. Understanding in this sense is gathering the information conveyed by an utterance—knowing what is being talked about and how it is being classified.

The use of quasi-indexicals in ascriptions of strong *de re* beliefs preserves more than reference, however. The ascriber's access to the object the belief is strongly about is anchored, by anaphoric links, in the believer's demonstrative acquaintance with the object. The sort of attribution that is in this way made explicit is attribution not just of a belief *about* that object but of belief *strongly* about that object. This is a more demanding, broadly intensional sort of communication and understanding of another's claim. It is an extension—via anaphoric commitments to defer regarding what substitutional commitments govern the use of an expression—of the sort of sameness-of-claim content defined by deference of assertional responsibility to vindicate entitlement to a claim. A member of the audience of a speech act containing a demonstrative reference can pick it up anaphorically (perhaps, but not necessarily quasi-indexically) and so understand it, without thereby counting as able to entertain the very same (demonstrative) thought the speaker expressed.

#### 4. Object-Dependent Singular Thought

The propositional contents of demonstrative beliefs and claims are paradigmatic of a wider class, sometimes called 'singular thoughts' in virtue of the peculiarly intimate cognitive relations to their objects that they embody. Russell thought that the best way to accommodate this sort of strong *de re* belief in semantic theory is to treat the propositional contents they express as having the objects themselves as constituents.<sup>66</sup> Less ontologically alarming contemporary descendants of these Russellian propositions are proposed by those who treat demonstratives and indexicals as devices of *direct reference*, as picking out their objects without employing any conceptual resources or mediating senses.<sup>67</sup> Each of these is motivated by the thought that the behavior of demonstratives and indexicals requires an anti-Fregean semantic theory. McDowell, following Evans, has argued forcefully to the contrary—that the underlying phenomenon is the *object-dependence* of the thoughts or contents expressed by the use of these locu-

tions, and that a thoroughly Fregean account of object-dependent thoughts and singular senses is possible.<sup>68</sup> There are some thoughts one can entertain regardless of whether the singular terms occurring in the sentence tokenings that express them succeed in picking out objects. Thus the claim that the woman who wrote *Sordello* was hopelessly sentimental and syntactically unsound expresses a genuine thought—even though the singular term fails to refer (since Robert Browning authored the work that one would by implication be criticizing by endorsing that claim). Object-dependent thoughts, by contrast, are those that can be entertained only if the singular terms occurring in the sentence tokenings that express them do succeed in picking out objects. Prime among them, the suggestion is, are those propositional contents whose expression requires the use of demonstratives or indexicals.

That essentially demonstrative thoughts are object-dependent is a claim about their accessibility—their accessibility not to different individuals but to the same individual in different possible circumstances. If there is indeed a cup in my hand, the thought I could express by asserting “This cup (in my hand) was on my desk yesterday” is one that might be true or false, depending on the previous career of that cup. But if there is no such cup, if I am hallucinating or dazzled by reflections into mistaking a stone in my hand for a cup, then that thought is not available for me to think.<sup>69</sup> That thought is strongly about a particular cup, and if it is not available to be demonstrated, I am not in a position to think a thought that is strongly about that cup. This need not mean that I am not thinking at all; I may be thinking (falsely) that there is a cup in my hand and that it was on my desk yesterday. That is a thought, but it is not the same one that I would express by the use of the demonstrative in the different case in which I am not mistaken about there being a cup in my hand. Calling singular thoughts and beliefs—those that are strongly about this very object in the way that is paradigmatically expressed by the use of demonstratives and indexicals—‘object-dependent’ is intended to be a neutral specification of the phenomenon addressed theoretically by Russell’s claim that the object itself is a *constituent* in the proposition entertained and by the claim that such thoughts and beliefs refer *directly* to their objects, without the use of concepts or associated senses.

The phenomenon of object-dependent propositional contents is widely thought to be of particular significance for semantic theory. The reason for this is that it makes manifest a fundamental feature of the way in which thought and talk (insofar as they have an empirical and practical dimension) relate to what they are about—the objects that must be consulted in order to assess the *truth* of what is believed and claimed. The expressive role played in semantic assessment by the theoretical vocabulary whose paradigm is ‘true’ has been explained here in terms of the adoption of perspectively hybrid deontic attitudes. Treating a claim as *true* is *attributing* one doxastic commitment while *undertaking* another which shares or anaphorically inherits its propositional content. Similarly, treating an action as *successful* is

attributing a practical commitment while undertaking a doxastic commitment with a corresponding propositional content.

The object-dependence of strong *de re* propositional contents can also be understood in terms of perspectivally hybrid deontic attitudes. Taking someone to have a strong *de re* propositional attitude is *attributing a doxastic commitment* the expression of which would involve the demonstrative or indexical use of a singular term (or more broadly, of an anaphoric dependent traceable back to such an antecedent)—an attitude that can be *ascribed de dicto* using a quasi-indexical. But it also requires that the one attributing that attitude *undertake an existential commitment* regarding the singular-term tokening in question. The object-dependence of a propositional content consists in the additional existential commitment *undertaken* by anyone who *attributes* a status or attitude exhibiting such a content.<sup>70</sup> (Recall that existential commitments were explained in Chapter 7 in substitutional terms, with respect to a special class of designators that are canonical for the sort of existence that pertains to the objects at issue—in this case, accessible spatiotemporal ones.)

For this reason, *de re* ascriptions of epistemically strong *de re* beliefs involve existential commitments. Recall that (weak) *de re* ascriptions as here conceived need not involve such commitments. If someone who has never heard the name ‘Pegasus’ believes that Bellerophon’s horse has wings, I can specify the content of that belief in the weak *de re* way by saying that he believes of Pegasus that he has wings, without undertaking any existential commitment to the existence of such a horse. But if I say that he believes of<sub>strong</sub> Pegasus that he has wings, I am committed to his being able to pick out Pegasus by using a demonstrative, and hence to the spatiotemporal accessibility of that horse in the common environment he and I share, which is what a commitment to the physical existence of the horse comes to on the analysis presented in Chapter 7. That *de re* ascription of an epistemically strong *de re* belief commits me to there being some indexical expression *t* such that

- (a) he believes that  $a^{(t)}$  has wings, and
- (b)  $a^{(t)} = \text{Pegasus}$ .

Undertaking commitment to that identity involving the quasi-indexical  $a^{(t)}$  commits me to its antecedent tokening /*t*/ picking out Pegasus, and this is the source of the existential commitment I undertake in making the *de re* ascription of a belief strongly of Pegasus. The result is that *de re* ascriptions of essentially indexical beliefs provide a way of understanding what Quine was after with his notion of *relational* belief. For his conception of such beliefs embodied both the idea that they involve a special sort of epistemic access to or rapport with objects, and that *de re* ascriptions of them involve existential commitments.

Expressions making explicit the stronger sense of ‘of’ can then be under-

stood in terms of their role in this kind of ascription. If S attributes the strong *de re* doxastic commitment that would be ascribed *de dicto* by asserting "*T* believes that she<sup>(that woman)</sup> is a spy" and undertakes the doxastic commitment that would be acknowledged by asserting "The one *T* referred to as 'that woman' is (=) Rosa Kleb," "She<sup>(that woman)</sup> is Rosa Kleb," or in suitable circumstances just "She is Rosa Kleb," then *T* is also committed to the *de re* ascription of that commitment: "*T* believes *of<sub>strong</sub>* Rosa Kleb that she is a spy." The subscript distinguishes the 'of' as expressing a strong *de re* belief by marking the requirement that the underlying *de dicto* ascription involves the use of a quasi-indexical (or a proper name functioning analogously to a quasi-indexical), and hence ascribes a strong *de re* belief.

A scorekeeper who attributes a demonstrative tokening but who does not undertake an existential commitment regarding it—and so does not take it to pick out an actual object—does not take the thought or belief in whose expression it occurs to be strongly of or about any object. What is expressed is taken to be only a mock-thought, one that could be mistaken for a thought with demonstrative content but that in fact has none (though it may be associated with genuinely contentful nondemonstrative thoughts).<sup>71</sup> In this case the scorekeeper must not undertake substitutional commitments with respect to the original demonstrative tokening—and so not with respect to its anaphoric dependents, including quasi-indexicals. It is for this reason that *de re* ascriptions of essentially indexical beliefs involve existential commitments and so count as attributions of successful object-dependent beliefs.

Besides the sort of error involved in taking it that someone (perhaps oneself) has a demonstrative thought when there is no such thought to be had, it is also possible to have two genuine demonstrative thoughts that are strongly about one and the same object, without realizing that they are. In the classic case, one could say of Venus in the morning, "This planet is the last one visible as the sun rises," and of Venus in the evening, "This planet is the first one visible as the sun sets," without realizing that the two claims were about one planet. Conversely, it is possible to have two genuine demonstrative thoughts that are strongly about different objects while believing that they are about one single one. I may at one moment think truly of the cup in my hand (as "this cup") that it is the one that was on my desk last night and, after a moment of confusion, think the same thing of the indistinguishable duplicate that replaces it, without realizing that the first thought is true and the second false, since they are about different cups.

In Fregean terms, what one grasps when one has practical mastery of the use of an expression—understands the thoughts expressed by that use—is its *sense*. The possibility of this sort of individuative mistake regarding the use of demonstratives shows that the notion of *de re* or object-involving senses or conceptual contents is incompatible with the doctrine that such senses are epistemically *transparent*. According to the doctrine of their transparency, grasping a sense entails individuative omniscience regarding it: one

cannot grasp two different senses without realizing they are different, nor the same one twice without realizing that. On this line, the fact that one may fail to realize that the objects picked out by two different senses are different or that they are the same—may endorse false identities and deny true ones—just shows that one may be ignorant about the world, not that by grasping the senses one does not thereby know everything about them. According to this ultimately Cartesian picture of conceptually articulated senses, they lie entirely within the mind, open to its survey in a way that rules out the possibility of error or ignorance. Only in collaboration with the world lying wholly without the mind do they determine referents, however, and it is the epistemic opacity of that second semantic component that introduces the possibility of error and ignorance—which accordingly pertains not to the representational *purport* of those senses but only to its specific *success* or failure.

It may be doubted whether Frege was ever committed to this picture.<sup>72</sup> In any case, it does not work well for the sort of strong aboutness exhibited by essentially demonstrative or indexical claims and the beliefs they express. Here the model of sense and reference that has traditionally been associated with Frege must be modified. If strong *de re* or object-involving senses are admitted, then the transparency principle regarding the individuation of senses must be relinquished; if the transparency principle is retained, then object-involving thoughts must be understood as picking out their objects without the intervention of senses. The latter option is the one pursued by theories that understand strong *de re* claims and beliefs as *directly referential*—namely, that the object picked out by singular terms such as demonstratives, indexicals, and indeed proper names is the *only* contribution their occurrence makes to the claims expressed by the sentence tokenings in which they appear.

Two related phenomena raise particular challenges for such a view. First is the version of Frege's puzzle about identities that arises for the special case of directly referential expressions. How is one to explain the difference in inferential significance between acquiring commitment to a claim of the form  $t = t'$  and acquiring commitment to a claim of the form  $t = t$ ? The former, as Frege points out, can be “rich in consequences,”<sup>73</sup> furthering one's knowledge by substitutionally entitling one to new claims; the latter is not something that could ever count as a discovery—it can license no new conclusions. Yet if the two terms are understood as directly referential, it would seem that claims of the two forms say just the same thing: they are strongly about an object, and they say of it just that it is identical to itself. The second phenomenon concerns how this difference is reflected in explicit *ascriptions* of belief or commitment. It seems that one can coherently ascribe the belief that  $t$  and  $t'$  are not identical even when those singular-term tokenings are expressions that secure an epistemically strong rapport with one and the same object and so according to this approach should be under-

stood as directly referential. But how is one to explain the content of the belief that is being ascribed in that case? The classical Fregean account of such ascriptions turns precisely on the *senses* of the expressions used to specify the content of the ascribed belief—and it is the essence of the direct reference approach to eschew appeal to any such senses.

To understand these phenomena one must look at the way singular reference can be passed on *anaphorically*. Anaphoric chains contribute to both of the theoretical tasks for which Frege postulated senses: they are ways in which objects can be given to us, and they determine the reference of the expressions occurring in them (whose senses they articulate). Anaphoric chains of tokenings—explained (as in Chapter 7) in terms of inheritance of substitutional commitments—provide a model for object-involving, *de re* senses. The access such anaphoric senses provide to the objects they pick out is *conceptual*. For the conceptual is understood in this work in terms of *inferential* articulation, and anaphora is a fundamental mechanism whereby unrepeatable tokenings are linked into repeatable classes subject to *substitutional*, and hence indirectly to *inferential*, commitments (see further discussion below in 9.1). The cognitive accessibility of empirical objects depends in no small part on the possibility of grasping object-involving anaphoric senses—even though such senses are not themselves epistemically transparent to those who grasp them.

Singular thought, which is strongly about particular objects,<sup>74</sup> extends beyond essentially demonstrative or indexical thought. The discussion of the use of quasi-indexicals in ascriptions shows how singular reference, originally secured by the use of demonstratives or indexicals, can be preserved and extended anaphorically. This should come as no surprise, in light of the argument in Chapter 7 to the effect that unrepeatable tokenings, paradigmatically those used demonstratively and indexically, can function as singular terms at all, and so as picking *out* objects, only in virtue of the possibility of picking *up* their references anaphorically, to construct token repeatables that can figure in substitutional commitments. If my claim “*This cup* was on my desk last night” succeeds in being strongly of or about a particular cup, then so does the anaphorically dependent claim (whether in my mouth or in that of another) “And *it* was also there the night before.” The singular reference can likewise be extended by the use of explicitly anaphoric definite descriptions, such as “the cup he referred to on that occasion as ‘this cup’,” which give the sense (anaphoric, not demonstrative) of ‘that cup’ in sentences such as “That cup is *not* the one that was on your desk last night” when uttered by another, later, out of sight of the cup in question.

Indeed, without such mechanisms one often could not so much as state identities linking essentially demonstrative or indexical thoughts. One cannot, for instance, be in a position to express the strong *de re* version of the discovery that Frege was interested in by saying “This planet [Venus, in the morning] is (=) this planet [Venus, in the evening].” For at any time when one

could perform the one demonstration, one could not perform the other. One must rather say something such as "This planet [Venus in the morning] is (=) that planet, that is, the one I referred to this morning as 'this planet'." Even in cases where one is situated so fortunately as to be able to say "this is (=) this" or "this is that," (where 'that' is used demonstratively, rather than anaphorically), such an utterance has the significance of asserting an *identity* only insofar as it can be used to license *substitutions*. (The appearance of 'is' or '=' can mark that aspiration but cannot by itself make the remark able to perform that function.) That it can be so used depends on the possibility of picking up those demonstrative references anaphorically, using expressions such as 'it', 'that planet', 'the cup N.N. referred to yesterday as "this"', and so on. So much as stating the Frege puzzle about the difference between the claim that would be expressed by a sentence of the form  $t = t'$  and one of the form  $t = t$  requires the repeatability of  $t$  in order to formulate the trivial identity (or the overtly anaphoric reflexive construction in  $t =$  itself). It is only as initiators of anaphoric chains that demonstrative and indexical tokenings provide ways of talking or thinking about objects at all, and hence potentially as *strongly* of or about them. These chains provide the *point* of using demonstratives or indexicals, and they articulate the *significance* of doing so.

### 5. *The 'Puzzle about Belief'*

Proper names can be used to express singular thoughts. The object-involvingness of these uses of proper names has been an important topic addressed by theories of direct reference. It was suggested already in Chapter 7 (Section IV) that the considerations that motivate causal-historical theories of proper names can be understood equally as motivating an *anaphoric* theory of proper-name usage. According to such an approach, the tokening-repeatables corresponding to proper names should be understood not as equivalence classes of lexically cotypical tokenings but as anaphoric chains. Such chains are anchored in a tokening that plays the role of anaphoric initiator, which corresponds to the introduction of the name. In favored cases, that initiating token might be a singular term used demonstratively or indexically: "I hereby christen *this ship* 'Beagle'." The role of the introducing token as anaphoric initiator, and hence as reference-fixer, means that even where it has the syntactic form of a definite description, it should be understood as used demonstratively—that is, as 'dthat'ed, and so rigidified. This is not a special feature of anaphoric chains that are proper names but a general feature of anaphora. Consider the following discourse:

#The president of the organization is a brilliant woman. She would have been a brilliant woman even if she had never become president of the organization.#

The initiator 'the president of the organization' serves to pick out an individual; anaphoric dependents on it, such as the two tokenings of 'she', continue to refer to that same individual even in counterfactual situations in which one could not pick out that individual with that description or title. Indeed, as this example suggests, this feature of the use of anaphoric dependents is essential to the possibility of expressing such counterfactual circumstances.

This approach assimilates the way proper names express singular thoughts to the mechanism of overtly anaphoric preservation of object-involvingness just considered. On the anaphoric account of proper-name usage, one who uses a tokening of a proper name is committed thereby (in the eyes of those keeping deontic score) to the pragmatic significance of that tokening being determined by the same substitutional commitments that govern its anaphoric antecedents—whatever those are. The one using the proper name need not know what the antecedents are (never mind what the original initiator of the chain is) in order to undertake that anaphoric commitment. Nor, provided the one using the name is generally competent in the relevant practices, need that individual *intend* to use the tokening to pick out the same object as the one uttering some particular antecedent did. The anaphoric commitment to inherit substitutional commitments can be undertaken and attributed in the absence of any such intention, just as one can undertake an assertional commitment as an inferential consequence of acknowledging another one, without intending to or realizing that one has done so. Each scorekeeper who treats a tokening as a proper name, and hence as an anaphoric dependent, must as part of that interpretation of the utterance assign it to some anaphoric chain (or at least, take it that there is some such chain), for it is that chain that determines what commitments one undertakes by the utterance, and hence its pragmatic significance. Both intentions and conventions may be relevant to such scorekeeping and may be appealed to by other scorekeepers in assessing the correctness of an interpretation. But the essential thing to understand theoretically is what it is for a scorekeeper to treat one tokening as anaphorically dependent on another, and that is not a matter of intention or convention but of commitment.<sup>75</sup>

If anaphoric chains of tokenings are the semantically relevant structure that governs the behavior of proper names, then it is easy to understand how different tokenings of the type ⟨Aristotle⟩, for instance, can refer to different people. This happens when they owe allegiance to different anaphoric chains, anchored in different antecedents—in the same way that different tokens of the type ⟨it⟩ can refer to different objects. Individual speakers are not omniscient about the commitments they undertake by their use of various expressions; they need not know much about the anaphoric chains they participate in and appeal to in order nonetheless to have their commitments determined (according to scorekeepers) by those chains. As a result, they may not realize that two singular propositions they are entertaining are strongly about the

same individual, or they may mistakenly think that one individual is at issue when two are. These are the phenomena Kripke addresses in his essay about the use of object-involving expressions in ascriptions of propositional attitude, "A Puzzle about Belief."<sup>76</sup> Seeing through the difficulties Kripke raises requires bringing together the treatment of expressions of strong *de re* belief with the previous discussion of *de dicto* ascriptions.

The puzzle concerns an apparent incoherence in ascriptions of belief that arises when someone uses two different proper names without realizing they refer to the same object, or mistakenly takes different occurrences of one name to refer to different objects. The example that is usually discussed concerns Pierre, who, though generally bilingual in French and English, does not realize that 'Londres' and 'London' are names of the same city. Keeping separate substitutional accounts under the two headings, he endorses both "Londres est joli" and "London is not pretty." The question is whether or not he should be said to believe that London is pretty.

The puzzle arises from two seemingly harmless principles governing the specification of the content in indirect discourse ascriptions of belief, which Kripke calls the *disquotational* principle and the principle of *translation*. The translation principle governs the relation between reported tokening (which would express the belief ascribed) and the reporting tokening inside the scope of the 'that' (which specifies the content of the belief ascribed). It is just the principle considered in Section IV above in the discussion of Davidson's "On Saying That." In fact, however, only the disquotational principle is required. As Kripke says after discussing the bilingual case: "Even if we confine ourselves to a single language, say English, and to phonetically identical tokens of a single name, we can still generate the puzzle . . . Only the disquotational principle is necessary for our inference; no translation is required."<sup>77</sup> The case he considers is one where someone learns of the pianist Paderewski, and so it is true to say, based on the avowals he is willing to make:

Peter believes that Paderewski had musical talent.

But then he learns of the nationalist leader Paderewski and concludes from general beliefs what can be reported, based on his avowals, as

Peter believes that Paderewski had no musical talent.

From the vantage point of this example, it is clear that the example of Pierre involves a translation principle governing the ascription of beliefs only because it amounts to a case like that of Peter, but in which one of the lexical expressions under which he has segregated his beliefs (in spite of their actually being about one individual) is translated into a different language. Thus the translation principle really adds nothing essential to the case.

The puzzle or paradox comes in two forms. In the weaker form, the sincere avowals of Peter lead us to attribute to him inconsistent beliefs, namely the belief that Paderewski did have musical talent and the belief that Paderewski

did not have musical talent. Now one might think that there is nothing particularly paradoxical about having inconsistent beliefs. Certainly in the *de re* sense, one can believe of Benjamin Franklin, as the inventor of bifocals, that he did not invent the lightning rod, and also belief of Benjamin Franklin, as the inventor of the lightning rod, that he did invent the lightning rod. But what is at issue is *de dicto* ascriptions of belief. The relevant parallel would accordingly be ascribing both the belief that Benjamin Franklin did invent the lightning rod and that Benjamin Franklin did not invent the lightning rod. Kripke makes contradictions of the sort exemplified by the Paderewski case particularly telling by invoking in addition a principle that may be called the *transparency of inconsistency*:

We may suppose that [the believer], in spite of the unfortunate situation in which he now finds himself, is a leading philosopher and logician. He would *never* let contradictory beliefs pass. And surely anyone, leading logician or no, is in principle in a position to notice and correct contradictory beliefs if he has them.<sup>78</sup>

A straightforward application of the principles of disquotation and translation yields the result that Pierre holds inconsistent beliefs, that logic alone should teach him that one of his beliefs is false.<sup>79</sup>

The principle of the transparency of inconsistency is what underwrites the inference from holding inconsistent beliefs to the claim that logic alone should show that one must be discarded. Again, with respect to another analogous case: "Our Hebrew speaker both believes, and disbelieves, that Germany is pretty. Yet no amount of pure logic or semantic introspection suffices for him to discover his error."<sup>80</sup> In the stronger form of the paradox, the inconsistency is charged not against the one whose beliefs are being reported but against the ascribers. For those ascribers are charged with both taking it that Peter does believe that Paderewski has musical talent and taking it that Peter does *not* believe that Paderewski has musical talent.

It is worth looking closely at the disquotational principle that causes all the trouble. Here is the full passage that introduces and specifies it:

Let us make explicit the *disquotational principle* presupposed here, connecting sincere assent and belief. It can be stated as follows, where 'p' is to be replaced, inside and outside all quotation marks, by any appropriate standard English sentence: "*If a normal English speaker, on reflection, sincerely assents to 'p', then he believes that p.*" The sentence replacing 'p' is to lack indexical or pronominal devices or ambiguities that would ruin the intuitive sense of the principle (e.g. if he assents to "you are wonderful," he need not believe that you—the reader—are wonderful) . . .

A similar principle holds for sincere affirmation or assertion in place of assent.<sup>81</sup>

Kripke says of this principle: "Taken in its obvious intent . . . the principle appears to be a self-evident truth."<sup>82</sup>

There are two principles with quite different functions being endorsed here. One of them connects overt linguistic *avowals* of belief (assertions, or affirmations) with reports, attributions, or ascriptions of belief. It says, in effect, that the best evidence one could have, or very strong evidence, or *prima facie* evidence that is hard to override, for ascribing a belief with a certain content is a sincere avowal or affirmation with that content. 'Self-evident' is not a word whose use is encouraged or endorsed by the point of view pursued in this work (in view of its standard circumstances and consequences of application), but it can nonetheless be agreed that anyone who denied all principles along these lines would thereby simply have changed the subject and shown that he or she was not talking about our concept of belief at all. At any rate no philosophical ground is to be gained by denying that there is any connection of this sort between avowal and ascription of belief.

The second principle concerns the relation between the expression used to specify the content of the belief avowed, on the one hand, and the expression used to report or ascribe that belief, in indirect discourse, on the other. The relation asserted here is the one that gives the principle Kripke appeals to its name: the *disquotational* principle. Under this heading the claim is that the very same words used to avow the belief are to be used to report it. The relation between direct discourse quotation of what was said and indirect discourse reporting of what was claimed is that of lexical cotypicality. This principle has a very different status from the first. It should not count as "self-evident" in anyone's book. Indeed, as Kripke immediately acknowledges, it is evidently false for a number of different sorts of locution. The whole problem of indirect discourse consists in specifying the relation between the tokening actually uttered, or that the believer could utter to express the belief in question, on the one hand, and the tokening another can admissibly use in reporting what was asserted or believed, on the other. Specifying this relation involves a myriad of subtleties, which the 'disquotational' strategy simply ignores. These are just the subtleties that were discussed in Section IV in connection with McDowell's emendation of Davidson's paratactic theory of the expressive role of content-specifying 'that' clauses in *de dicto* ascriptions of propositional attitude. The upshot of that discussion is precisely that translation plus disquotation (= type repetition) will not do in general as an account of the relation between the reported and the reporting tokening in indirect discourse ascriptions.

In the passage above, Kripke explicitly formulates the disquotational principle, on which the paradox or puzzle rests, so as to acknowledge the existence of locutions to which it does not apply. He mentions both the occurrence of indexicals (presumably including demonstratives), and of pronominal or anaphoric constructions. Why should we not conclude from the

puzzle cases offered just that proper names, at least under some circumstances, are also locutions for which the disquotational principle is inadequate? Such a conclusion seems particularly apt in view of the fact that demonstratives, indexicals, and tokenings that are anaphorically dependent on them are the paradigm of object-involving singular-term usage, and that Kripke sets up the problem about proper names by appeal to the very features of their use that have made direct reference theories attractive (both for them and for demonstratives, indexicals, and so on). Kripke nowhere addresses the possibility that the disquotational principle does not apply to proper names in the same way and for the same reasons that it does not apply to demonstratives, indexicals, and their anaphoric dependents.

The concessive qualification that follows the explicit statement of the principle does not quite say that the principle is to be taken as holding except in those cases where it does not. Three classes of exceptions are acknowledged. Consider first "ambiguities that would ruin the intuitive sense of the principle" (that is, presumably, render it false). What sort of ambiguities are these? Kripke considers a case which he does not take to present the difficulties characteristic of his "puzzle." Suppose that Arthur has in his idiom two names that share the lexical type *(Cicero)*. Some of his tokenings of that type are intended to refer to a Roman orator, and others are intended to refer to a famous spy from the Second World War. Then Arthur would affirm both tokenings of the type *(Cicero was a spy)*, when the latter is under discussion, and tokenings of the type *(Cicero was not a spy)*, when the former is under discussion (the denouncer of Catiline has been accused of many sins, but so far as Arthur knows, never of being a spy). In such a case Kripke does not take it that Arthur has inconsistent beliefs, believing both that Cicero is and is not a spy, nor that we have inconsistent beliefs in both taking him to believe that Cicero is a spy and not to believe that Cicero is a spy. The forms of words parallel the Paderewski case, but the dual use of the name 'Cicero' prevents inconsistency, and therefore paradox. This seems to be the sort of ambiguity that is intended to be excluded by the third phrase in the qualification of the disquotational principle.

How does this case differ from that of Paderewski? Peter believes that he is using different tokens of the type 'Paderewski' systematically as different names of different individuals, in a way exactly parallel to Arthur's use of 'Cicero'. As it turns out, this is not the case. But to claim that the one situation is unproblematic and the other is paradoxical evidently requires more of a difference than simply that Peter's uses are in fact coreferential, while Arthur's are not. After all, 'the inventor of bifocals' and 'the inventor of the lightning rod' are coreferential, like the various uses of 'Paderewski', and there is no suggestion that paradox lurks in the beliefs of someone who attributes a property to Franklin by the use of the one expression that he denies to him by the use of the other. Kripke's appeal to the transparency

principle in stating his puzzle commits him, however, to very strong criteria of adequacy on the difference he discerns between these cases. The principle of the transparency of inconsistency, cited above, makes explicit a commitment to Peter's being able to tell his case from Arthur's by the use of "pure logic and semantic introspection alone"—for only this can set up Kripke's dilemma. At the very least, such resources ought to suffice to tell to which cases the disquotational principle applies ("self evidently"!), as in Peter's case, and to which it does not, as in Arthur's. Yet clearly Peter and Arthur are in no position to tell their cases apart. Kripke offers no argument whatever for not assimilating Peter's case (and therefore Pierre's) to Arthur's ambiguity case.<sup>83</sup> Neither identity of lexical type nor identity of referent will do the job. Presumably the difference has something to do with the fact that 'Paderewski' is used as *one* name by the community that Peter and also we, who are reporting on his beliefs, belong to, while 'Cicero' is used by us as two names. But these are not facts that are accessible to Peter and Arthur, least of all by "logic and semantic introspection." Thus it is difficult to see how such facts can do the work Kripke needs to have done in formulating the paradox. Again, why is not the proper conclusion from the puzzle cases just that sometimes proper names are so used that the disquotation principle does not apply to them because of dual uses not only of the 'Cicero' type but also of the 'Paderewski' type? An answer is not forthcoming.

#### 6. *Tactile Fregeanism: Proper Names as Anaphoric Dependents*

The remarks so far have simply been critical of Kripke's argument. But it should be clear that his puzzle can lead to positive conclusions, not so much about belief and its ascription as about proper names. Indeed, it is best thought of as a puzzle about proper names, rather than about beliefs—insofar as it deserves to be called a 'puzzle' at all. A good place to begin is with the observation that one cannot tell simply from the lexical type of an expression whether it is used in such a way that the disquotation principle applies to it. For instance, putting aside the category of proper names, which is currently at issue, only some occurrences of definite descriptions (the least 'Millian' of expression types) are suitable for the application of the disquotational principle. In particular, only 'attributive' uses of definite descriptions are suitable—'referential' uses are not. Thus it is possible to use the definite description 'the man in the corner drinking champagne' to refer to someone who is not in fact drinking champagne. Such 'referential' or speaker's referential uses can for many purposes be assimilated to demonstrative uses. In any case, such cases are not in general happily reported in indirect discourse by attributing "the belief that the man in the corner drinking champagne . . ."

Kripke offered three categories of exceptions to the disquotational princi-

ple: ambiguous, indexical, and pronominal uses. It has been suggested here that there might be reasons to assimilate the uses of proper names that are puzzling if they are assumed to fall under the disquotational principle either to the case of ambiguities or of indexical uses. What about the possibility of assimilating them to the pronominal cases? The disquotational principle does not properly describe the relation between the tokening used to avow a belief and the tokening used to report or ascribe that belief in case the former contains an anaphoric dependent, because the ascriber's repeating another token of the same type need not involve anaphorically picking up the same antecedent as the original. Thus if Jones says: "I met a man from the Biology Department yesterday" and if Smith responds with the avowal: "He must be the new statistician," it need not be correct for Brown later to report this belief with the ascription: "Smith believes that he must be the new statistician." Only if special arrangements are made to see to it that the antecedent of the pronoun tokening that appears in the content-specifying portion of the ascribing expression is either the reported tokening or its antecedent can a type-identical term be used in the indirect discourse report.

Is there any reason to think that tokenings of the type 'Paderewski' behave oddly in reports of Peter's beliefs because they are functioning in a way analogous to anaphorically dependent tokenings? Kripke describes his own approach to proper names this way: "Ordinarily the real determinant of the reference of names of a former historical figure is a chain of communication, in which the reference of the name is passed from link to link."<sup>84</sup> This is the "causal-historical" approach to proper-name usage. As has already been suggested, it can usefully be developed in the form of an *anaphoric* understanding of proper names. In the favored case, there is a 'baptism' of an object by using an indexical or definite description to pick out the object the name is to be attached to. This reference-fixing specification becomes the antecedent, standing at the beginning of a "chain of communication, which on the present picture determines the reference" of further tokenings of the name in the chain that is anchored by that initial antecedent.<sup>85</sup> These further tokenings inherit their reference from earlier elements in the chain. This is just anaphora. Anaphoric chains of tokenings create repeatables that can play the same role in substitution inference that is played by classes of cotypical tokenings, for instance, semantically definite descriptions, used attributively. Just so for chains of name tokenings. In each case, the commitments one is undertaking by using a dependent expression late in a chain are to be determined by tracing the chain back. The "legitimacy of such a chain" of name tokenings depends on how the reference is passed,<sup>86</sup> just as with an anaphoric chain. Kripke says that for proper names, differences in the beliefs of users do not change the reference of those name tokenings, so long as the user "determines that he will use the name with the referent current in the community."<sup>87</sup> This sounds like a primitive (and not ultimately satisfactory)

account of what it is to use a pronoun with a certain antecedent (as though one needed to have the *concept* of reference in order to use an expression in an anaphoric and so in fact reference-inheriting way!). In either the canonical anaphoric-dependent case or the case of proper-name tokenings, to take a tokening as continuing a chain is to take it as inheriting its substitution-inferential role from those antecedent tokenings.

Sometimes distinct anaphoric chains of proper-name tokenings are anchored in antecedents picking out different objects, as in the 'Cicero'/'Cicero' case. Sometimes distinct chains are anchored in a single object, as in the 'Paderewski'/'Paderewski' case. Both of these structures can also occur with ordinary anaphoric dependents, such as 'it'. Just as one cannot tell "by pure logic and semantic introspection" whether two chains that one is continuing are anchored in one object or in two for ordinary anaphoric dependents, so one cannot for the anaphoric chains that govern the use of proper names. Just as for canonically pronominal expression types such as 'he' or 'it', cotypicality is no guarantee of coreference—one must consult the anaphoric chain to which a tokening belongs in order to determine its reference. This is just the situation in which the disquotational principle is not applicable. The cases Kripke presents do not generate a puzzle; they just show that proper names can be used in such a way that the disquotational principle does not apply to them. Kripke's own approach to proper-name reference in terms of chains of tokenings suggests exactly why. The result is that these cases present good reason for treating proper names on an anaphoric model.

With the exception of the anaphoric initiator that is their common antecedent, all the tokenings of a single proper name are of the same lexical type. (As cases like Kripke's 'Cicero' case show, however, not all tokenings of the same lexical type need belong to one chain and so be tokenings of one proper name.) This explains why there is no need for a special category of *quasi-names* performing an expressive role in *de dicto* ascriptions of strong *de re* beliefs that corresponds to that performed by quasi-indexicals.<sup>88</sup> Quasi-names would be used to attribute a belief that would be properly expressed by the believer by the assertion of a sentence containing a proper name and would function as anaphoric dependents of such attributed tokenings. But if the ordinary uses of proper names are already anaphoric, with antecedents stipulated to be of the same lexical type, then proper names can function as ascriptional quasi-names without further alteration. The fact that for this reason one can often specify the content of an ascribed belief by using a tokening of the same type as would be used by the believer to express it is responsible for the impression that the disquotational principle ought to apply to proper names, as it does to *attributive* (that is, nondemonstrative, nonanaphoric) uses of definite descriptions. The underlying anaphoric structure of proper-name usage—and so the way in which the disquotational principle incorrectly generalizes from some well-behaved cases—is high-

lighted in cases where different anaphoric chains govern tokenings drawn from the same lexical type.

The invisibility of this alternative anaphoric analysis and the impression that the ‘puzzles primarily has to do with belief, rather than with proper names, both arise out of Kripke’s commitment to a ‘Millian’ theory of the semantics of proper names. The opposition Kripke sets up between ‘Millian’ and ‘Fregean’ theories of proper name usage is misleading in a number of ways. It is evidently not exhaustive. It is not even clear that Kripke’s own theory is best described as Millian—that is, directly referential in the sense that it is not legitimate to appeal to anything other than the referent (in particular, to an anaphoric antecedent or chain) in explaining the contribution the occurrence of a name makes to the use of a sentence in which it occurs. Nor does it seem sensible to treat other sorts of expressions as working this way. Demonstratives such as ‘this’ are not really “directly referential” because they require implicit sortals to pick out their referents. Pronouns such as ‘he’ are not, because they convey gender information, are differently used depending on what anaphoric chain they are part of (over and above what their ultimate referent is), and involve an all-but-explicit personal sortal. On the other side of the opposition, Kripke’s view is that “the opposing Fregean view holds that to each proper name, a speaker of the language associates some property (or conjunction of properties) which determines its referent as the unique thing fulfilling the associated property (or properties).”<sup>89</sup> It is hard to see how this view can qualify as Fregean. Properties are in the realm of reference, not in the realm of sense—they are not immediately graspable, for Frege. Furthermore, how is this account supposed to extend to the senses of predicates, for Frege? Presumably they are not also specified by a set of properties that picks them out. In any case, the possibility of *de re* senses as Evans and McDowell describe them is overlooked by this way of setting up the issue.

The conceptual contents that are expressed by proper names, which determine their substitutional and so inferential roles, correspond to anaphorically structured constellations of tokenings. These constellations are like Fregean senses in that they determine the referents of the name tokenings whose significance they govern. They are also like Fregean senses in that they provide cognitive access to the particular objects to which they refer—via intersubstitutability equivalence classes of token-repeatables, some of which include demonstrative and indexical tokenings. Such chains anchor our thought and talk in particular objects that it is about—the objects that must be consulted in order to assess the truth of our claims and the beliefs they express. They determine (according to the interpretation of a scorekeeper) not only *that* we are talking and thinking about particular objects by correctly using proper names but also *what* we are thereby talking and thinking about. The connection they establish between our thought and talk and its objects

is so tight that the propositional contents we express by their means are object-involving: the singular or strong *de re* thoughts we express by using proper names when everything goes well (according to a scorekeeper) are not ones we can so much as entertain in cases that are just like those in every way possible except that the anaphoric chains do not (according to the scorekeeper) provide access to a unique object.

Conceptual contents expressed by proper names—understood in terms of constellations of tokenings that are articulated by anaphoric commitments regarding the inheritance of substitution-inferential commitments—are unlike Fregean senses as traditionally conceived in that they are not epistemically transparent in their individuation. We can be confused about which anaphoric chain a certain tokening is beholden to, and hence about whether two tokenings belong to the same or to different chains. In that sense we do not always know what we are saying or thinking; even where the propositional contents are object-involving, confusion is possible regarding which objects our strong *de re* beliefs are about. Conceptual (because ultimately inferentially articulated) contents of this sort are best thought of on a *tactile*, rather than a *visual*, model. The Cartesian visual model is the one that gives rise to demands for transparency, to the idea that error and ignorance should be impossible regarding what is in the mind. But it is just this model of inside and outside that makes it unintelligible in the end how what is in the mind should essentially involve representational purport regarding what is outside the mind—how thought can seem to be, and when all goes well in fact be, *about* environing things.

Frege's own favorite metaphor for our cognitive relation to senses is that of *grasping* rather than *seeing*. One can grasp an anaphoric chain as one grasps a stick; direct contact is achieved only with one end of it, and there may be much about what is beyond that direct contact of which one is unaware. But direct contact with one end gives genuine if indirect contact with what is attached to the other end. Indeed, the more rigid the stick (or chain), the better are the antics of what is attached to its far end communicated to the end one grips, and the more control one can exert, albeit at a distance, over it. The Cartesian model of conceptual contents restricts them to the part of the stick touching one's hand, at the cost of mystery about how our cognitive reach can exceed that immediate grasp. A *tactile* Fregean semantic theory, of the sort epitomized by understanding proper names as constellations of singular-term tokenings articulated by anaphoric commitments, effaces this impermeable boundary between the transparency of the mind and the opacity of its objects. The model of thought it presents incorporates, as two sides of one coin, both the possibility of ignorance of and error about our own concepts and the possibility of genuine aboutness of those concepts and genuine knowledge of the objects with which those concepts put us in touch.

## VI. THE SOCIAL-PERSPECTIVAL CHARACTER OF CONCEPTUAL CONTENTS AND THE OBJECTIVITY OF CONCEPTUAL NORMS

### 1. Weak and Strong Aboutness

The strategy pursued in this chapter is to explain the representational dimension of discursive practice by offering an account in deontic scorekeeping terms of what is expressed by the locutions of ordinary language that are used to make it explicit. These are the expressions used to say what claims and beliefs are *of*, *about*, or *represent*, in the sense of what objects must be consulted in order to assess their *truth*. Locutions such as ‘*of*’, ‘*about*’, and ‘*represents*’ play the expressive role of representational locutions in virtue of their use in *de re* specifications of the contents of ascriptions of propositional attitude. The basic way of specifying the propositional content of an ascribed commitment is the *de dicto* style. This style of content specification corresponds to a particular way of indicating a possible sentence tokening (*dictum*) on the part of the one to whom the commitment is ascribed that that individual would, according to the ascriber, acknowledge as expressing the commitment ascribed. The ascriber does this by embedding in the ascription—in regimented ascriptions, as a clause inside the scope of a ‘*that*’—a sentence tokening that stands in a relation to the indicated tokening that is potentially complicated (where it must be anaphorically articulated) but that in the simplest cases amounts to mere repetition from the perspective provided by the ascriber’s different language or indexical circumstances, preserving the claim expressed.

The *de re* style of specifying the contents of ascriptions derives from this basic *de dicto* style by allowing substitutions that are licensed by the commitments *undertaken*, rather than those *attributed*, by the ascriber. Thus the very same commitment that *S* can ascribe *de dicto* by asserting “*T* believes that Benjamin Franklin could speak French” can be ascribed *de re* by asserting “*T* believes of the inventor of bifocals that he could speak French,” provided only that *S* endorses the substitution-licensing identity “Benjamin Franklin is (=) the inventor of bifocals”—regardless of whether *S* also takes it that *T* endorses that identity. The *de re* form accordingly specifies what individual, according to the ascriber, it is whose properties must be investigated in order to determine whether the ascribed belief is *true*. *S* will endorse *T*’s claim—take it to be true—if and only if *S* endorses the claim that the inventor of bifocals could speak French. *De dicto* ascriptions specify the content of the attributed commitment from the point of view provided by what the one to whom that commitment is attributed would, according to the attributor, acknowledge. *De re* ascriptions specify the content of that same commitment from the point of view provided by what the one attributing the commitment would acknowledge. These are two different sorts of ascription, two ways of specifying the content of a single commitment, not ascriptions of two different sorts of belief or commitment.

The previous section of this chapter acknowledges that besides two styles of content specification that can be employed in ascribing any belief whatsoever, one of which can sensibly be called *de re*, there are also two different sorts of belief, one of which can sensibly be called *de re*. Besides weak *de re* ascriptions of belief there are strong *de re* beliefs—which like the others can be ascribed either *de dicto* or *de re*. Distinguishing strong *de re* beliefs is the result of adopting a “frankly inegalitarian attitude” toward different ways of specifying an object—in particular, by acknowledging the special sort of rapport secured by the capacity to pick out an object by certain sorts of uses of demonstratives, indexicals, and proper names. The task of the last section was to account in scorekeeping terms for the use of the sentences that express and ascribe such strong *de re* beliefs. This story had three parts. The first examined the sort of strong *de re* rapport established by using certain kinds of singular terms to pick out objects. The prime example considered was that of first-person indexicals. Some attention was also paid to the anaphoric extension of demonstrative and indexical acquaintance, widening the category of singular or strong *de re* thought to include what is expressed in favored cases by the use of proper names. The second part of the story about strong *de re* belief concerned *de dicto* ascriptions of such beliefs. The key to understanding these was found in an anaphoric construal of quasi-indexicals—expressions used to ascribe the use of indexicals and demonstratives. (Understanding proper names as anaphoric chains of cotypical tokenings makes it clear how this one category can play both the expressive role of undertaking a commitment and that of ascribing it, so that no separate category of quasi-names is required.) The final move is then to derive the structure of *de re* ascriptions of strong *de re* beliefs from that of these *de dicto* ones.

Although the special, strong *de re* acquaintance with the individual is secured by the availability to the believer of a special sort of singular term to pick it out, what is secured is acquaintance with that individual. In *de re* ascriptions of strong *de re* beliefs, the individual in question can be specified in any way available to the ascriber; these are not restricted to the use of singular terms that themselves express strong *de re* rapport with the object of the belief. So if according to *S*, but unknown to *T*, Rosa Kleb is the only actual KGB colonel mentioned in Ian Fleming's *From Russia with Love*, then according to *S*, *T* believes of<sub>strong</sub> the only actual KGB colonel mentioned in Ian Fleming's *From Russia with Love* that she is a spy. As with *de re* ascriptions of ordinary beliefs, information that would be conveyed by a *de dicto* ascription about *how* the believer represents an object is suppressed, in favor of information about *what* object is represented; in the case of strong *de re* beliefs, *de re* ascriptions specify *that* the believer has a belief that is strongly of or about an object, and *which* object it is, but not (as a *de dicto* specification of the content of the same belief would) *how*—by the use of what sort of expression—that rapport is established.

Thus the second, stronger sense of ‘of’ is built on, and must be understood in terms of, the first and more fundamental sense. The representational dimension of discourse, its aboutness or intentional directedness at objects and states of affairs, is accordingly to be explained in the first instance as what is expressed by (weak) *de re* specifications of the contents of ascribed commitments—whether or not those commitments themselves qualify as strong *de re*. That *de re* style of content specification in turn reflects the *social* difference in doxastic perspective between the one *attributing* an ascribed commitment and the one *undertaking* it, the one keeping score and the one whose score is kept, the interpreting interlocutor and the interpreted interlocutor. *De re* ascriptions are the explicit expression of perspectively hybrid deontic attitudes, the attribution by the ascriber of one discursive commitment (doxastic or practical, and so propositionally contentful) and the undertaking by that ascriber of another (substitutional) commitment. Commitments of these kinds are conceptually contentful in virtue of their inferential significance; talk about the *representational* dimension of the conceptual content of intentional states should be understood in terms of the *social* dimension of their inferential articulation.<sup>90</sup>

## *2. The Essentially Perspectival Character of Conceptual Contents*

Conceptual content is understood in this work as what can be made explicit in discursive practice. Discursive practice has as its defining core *claiming*. Claims are a kind of commitment that can be understood in terms of the functional role things of this kind play in social scorekeeping practices—practices and practical attitudes that accordingly can be thought of as instituting this sort of deontic status. Commitments of this kind are in turn appealed to both to pick out specifically *propositional* contents and to pick out the specifically *sentential* locutions used to express them. Because it must incorporate performances with the pragmatic significance of claimings, discursive practice is *linguistic* practice. Linguistic practice is what makes it possible for something to be made *explicit* (and logical practice is what makes it possible for linguistic practice itself to be made explicit). Making anything explicit, *saying* it, requires using one linguistic expression rather than another. Specifying the content of a discursive commitment by using a particular sentence involves undertaking an *expressive* commitment concerning the inferential role of the locution employed.

Expressive commitments were mentioned already in Chapter 2, where it was pointed out that the use of any particular linguistic expression involves an *inferential* commitment—a commitment to the propriety of the inference from its circumstances of application to its consequences of application.<sup>91</sup> As with any sort of deontic status, expressive commitments should be understood in terms of the attitudes of *undertaking* and *attributing* them. These

differ not only in *who* endorses the propriety of the inference involved but also (typically, though not invariably) in *what* inference is taken to be involved. For what one takes to follow from what (committively or permissively) depends on what collateral premises one is committed or entitled to. This principle applies no less to the inferences incorporated in conceptual contents than to those that relate conceptual contents among themselves. To see the inferential relations internal to concepts as pragmatically and methodologically on a par with those that are external to concepts is the essence of the sort of holism about meaning that Quine introduces in "Two Dogmas of Empiricism," just as a result of considering the role of auxiliary hypotheses in determining the significance of particular commitments. Neither category can be traded in theoretically for (eliminated in favor of) the other in toto—any more than the implicit and the explicit can. But there is a deep practical equivalence between inferential commitments that are implicitly endorsed by employing particular bits of vocabulary and those that are overtly endorsed in the scorekeeping practice of taking or treating certain discursive commitments as consequentially related to, and so as providing reasons for, others.

We discursive creatures—rational, logical, concept-using ones—are construed here in *expressive* terms; we are the ones who can make it explicit. One manifestation of that overarching feature of the approach is the methodological principle that what is implicit is to be made theoretically, as opposed to practically, intelligible precisely by exercising our defining attribute—by making it explicit. When the inferences implicit in the use of a word are made explicit in the form of conditionals, the fact that the proprieties governing them are relative to a background of collateral commitments is manifest. Which conditionals one endorses depends in part on what other claims one endorses. When the circumstances and consequences of application potentially associated with a particular word are made propositionally explicit, which conditionals connecting them an interlocutor endorses varies with the other commitments, doxastic and inferential, that the interlocutor undertakes. This relativity of explicit inferential endorsements to the deontic repertoires of various scorekeepers reflects the underlying relativity of the inferential endorsements implicit in the concepts expressed by particular words, according to various scorekeepers. A word—'dog', 'stupid', 'Republican'—has a different significance in my mouth than it does in yours, because and insofar as what follows from its being applicable, its consequences of application, differ for me, in virtue of my different collateral beliefs (and similarly for circumstances of application—consider 'murder', 'pornographic', 'lyrical').

One way of characterizing this situation, which at least since Quine has been sufficiently popular among those endorsing holism that it is often thought to be one of its necessary consequences, is to say that different interlocutors—most importantly those undertaking commitments and those

keeping score by attributing them—use different concepts, attach different meanings to their words. There are not enough words to go around, however; we all use basically the same words to express the contents of our commitments. As was pointed out in Chapter 7, this way of putting things accordingly seems to threaten the intelligibility of mutual understanding and so of successful communication. The discussion of indexicals and demonstratives in the previous section provides a different model for construing communication and the conceptual contents that are communicated, in the face of variations in the significance of utterances in different mouths. The fact that the word 'I' can never have the significance in my mouth that it does in yours (not just practically or for the most part, but in principle) in no way precludes my understanding what you express by using it.

Communication is still possible, but it essentially involves intralinguistic *interpretation*—the capacity to accommodate differences in discursive perspective, to navigate across them. Anaphora can secure not only coreference but token repeatability across the different repertoires of commitments that correspond to different interlocutors. This capacity becomes explicit in the *ascription* by one individual of propositionally contentful commitments on the part of another individual. There quasi-indexicals can secure contact with and express the understanding of otherwise unrepeatable indexical and demonstrative tokenings by the one to whom a commitment is attributed. Anaphorically indirect definite descriptions make these connections explicit. Not only can one attribute an expressive commitment without undertaking it by making explicit what is implicit in the use of scare quotes—"the one John referred to as 'that murderous lowlife'"—but (at the usual risk of perhaps puzzling prolixity) by their means one can make explicit what is implicit in the use of quasi-indexicals—"the one John referred to as 'T,'" "the book he referred to as 'this'."

Even where the anaphoric mechanisms that make them possible are not on the surface, the difference between the expressive commitment acknowledged by one who uses a word to make a claim explicit as an overt assertion and the expressive commitment that would be undertaken by the use of the same word on the part of one who attributes that claim is explicitly taken account of and managed in ascriptions. For the difference between the expressive commitments undertaken by the ascriber and those attributed by the ascriber is overtly marked by the difference between expressions occurring *de re* and those occurring *de dicto* in the content-specifying portion of the ascription. Thus *T* may have beliefs that could be expressed in *T*'s mouth by the assertion of the sentences

That murderous lowlife should be locked up

or

The seventh god has risen.

*S* may not be in a position to undertake the expressive commitments involved in using the expressions 'that murderous lowlife' or 'the seventh god' in the way needed to express these claims and so may not be able to ascribe them *de dicto* by asserting

*T* claims that that murderous lowlife should be locked up

or

*T* claims that the seventh god has risen.

Nonetheless, *S* can pick up the expressions anaphorically and express the claims *de dicto* using scare quotes or indirect definite descriptions, which explicitly disavow the expressive commitments those expressions incorporate (attributing, rather than undertaking, responsibility for their use in picking out who or what is talked about):

*T* claims that '<sup>s</sup>that murderous lowlife<sup>s</sup>' (the one *T* refers to as 'that murderous lowlife') should be locked up

or

*T* claims that '<sup>s</sup>the seventh god<sup>s</sup>' (the one *T* refers to as 'the seventh god') has risen.

It is the same anaphoric capacity to express such identities as

He ('<sup>s</sup>that murderous lowlife<sup>s</sup>', the one *T* refers to as 'that murderous lowlife') is (=) the mayor

or

It ('<sup>s</sup>the seventh god<sup>s</sup>', the one *T* refers to as 'the seventh god') is (=) the sun

that make it possible for *S* to specify the content of the very same belief *de re*, as

*T* claims of the mayor ('<sup>s</sup>that murderous lowlife<sup>s</sup>', the one *T* refers to as 'that murderous lowlife', as 'that murderous lowlife') that he should be locked up

or

*T* claims of the sun ('<sup>s</sup>the seventh god<sup>s</sup>', the one *T* refers to as 'the seventh god', as 'the seventh god') that it has risen.

Where the parenthetical expressions are omitted, the underlying anaphoric links that make such *de re* content-specifications possible are not apparent on the surface of the ascription.

In ascribing a discursive commitment, the ascriber must specify its content and to do so must undertake various expressive commitments. Since the

same words in the ascriber's mouth often do not express the same claim that they would in the mouth of the one to whom a claim is ascribed, the content specifications must take account of the difference in discursive perspective between the ascriber and the target of the ascription. Claims must often be differently expressed from the two points of view. Nonetheless, what is attributed explicitly in a *de re* ascription can be the very same claim that would be acknowledged, using different words, in an assertion by the one to whom it is ascribed. The difference in the inferential significance of words in one mouth and words in another, due to differences in collateral commitments, should not be understood as meaning that one interlocutor cannot strictly be said to understand what another says; it should only be taken to mean that the content they both grasp (if all goes well in the communication of it) must be differently specified from different points of view. Conceptual contents are *essentially expressively perspectival*; they can be specified explicitly only from some point of view, against the background of some repertoire of discursive commitments, and how it is correct to specify them varies from one discursive point of view to another. Mutual understanding and communication depend on interlocutors' being able to keep two sets of books, to move back and forth between the point of view of the speaker and the audience, while keeping straight on which doxastic, substitutional, and expressive commitments are undertaken and which are attributed by the various parties. Conceptual contents, paradigmatically propositional ones, can genuinely be shared, but their perspectival nature means that doing so is mastering the coordinated system of scorekeeping perspectives, not passing something nonperspectival from hand to hand (or mouth to mouth).

The expressive resources that have been considered here—the distinction between vocabulary used *de dicto* and vocabulary used *de re* in ascriptions, anaphorically and ascriptionally indirect definite descriptions, quasi-indexicals, scare quotes, and so on—provide mechanisms for making explicit the systematic differences in scorekeeping significance that constitute essentially socially perspectival conceptual contents. The representational dimension of such contents—what is expressed in *de re* ascriptions using terms such as 'of', 'about', and 'represents'—is one manifestation of their socially perspectival character. Thus wherever propositional or other conceptual contents are attributed, there is some implicit appeal to an *I-thou* social practice in which one scorekeeper interprets the performances of another. The representational dimension of propositional content is conferred on thought and talk by the social dimension of the practice of giving and asking for reasons, in virtue of which inferentially articulated contents are essentially perspectival. As the point was put at the opening of this chapter, the conceptual contents employed in *monological* reasoning, in which all the premises and conclusions are potential commitments of one individual, are parasitic on and intelligible only in terms of the sort of content conferred by *dialogical* reasoning, in which the issue of what follows from what essentially involves

assessments from the different social perspectives of scorekeeping interlocutors with different background commitments.

We sapients are discursive scorekeepers. We keep track of our own and each other's propositionally contentful deontic statuses. Doing that requires being able to move back and forth across the different perspectives occupied by those who undertake commitments and those who attribute them. Reidentifying conceptual contents through shifts in doxastic and practical point of view requires *interpretation* in Wittgenstein's sense—substituting one expression of a claim (he says 'rule') for another. The ordinary discursive understanding of propositional contents involved in all explicit knowing—that something is the case must be understood in terms of this implicit, practical-interpretive scorekeeping as knowing-how. Wittgenstein's pragmatism about norms—his insistence that norms made explicit in principles are intelligible only against a background of norms implicit in practices—is one of the master ideas orienting this entire enterprise. As was argued in Section II, however, one of the results of following that pragmatist idea out in detail by working out discursive scorekeeping practices sufficient to confer propositional contents that can codify norms in the form of claims and principles is that Wittgenstein is wrong to take that pragmatist methodological principle to be incompatible with understanding discursive practice as involving interpretation (in his sense) at every level, including the most basic. The capacity to interpret remarks, to substitute different expressions of a claim, rule, or principle—the propositional form in which things are made explicit—is a basic component of the fundamental practical capacity to grasp and communicate essentially perspectival conceptual contents.

Conceptual content is essentially something expressible, but it is expressible only from the point of view of some background repertoire of deontic statuses. What sentence should be used to express a particular semantic content varies from individual to individual. Assessments of the pragmatic significance of using a linguistic expression on a particular occasion—for instance what would entitle its utterer to that speech act, and what it commits its utterer to—differ depending on what claims are available to serve as auxiliary hypotheses. Conceptual content consists in the systematic relations among the various pragmatic significances. Thus grasping the semantic content expressed by the assertional utterance of a sentence requires being able to determine both what follows from the claim, given the further commitments the scorekeeper attributes to the assertor, and what follows from the claim, given the further commitments the scorekeeper undertakes. Where the expressive power of ascriptional locutions is available, this implicit capacity to keep and correlate multiple sets of books is made explicit in the availability of both *de dicto* and *de re* specifications of the contents of attributed discursive commitments. These features—the relativity to social perspective of the pragmatic significance of using a particular sentence or other linguistic expression, hence of which semantic content that expression

makes explicit, hence of which expression can be used to specify a given content—compose the socially perspectival character of semantic contents.

That perspectival character of the contents conferred on linguistic expressions by their role in deontic scorekeeping practices is the detailed consequence of the general methodological policy of subordinating semantic theory to pragmatic theory. This is the methodological (as opposed to the normative) sense of ‘pragmatism’—that the point of treating sentences and other bits of vocabulary as semantically contentful is to explain their use, the practices in the context of which they play the role of expressing explicitly what is implicit in discursive deontic statuses. (Which is not to say that those statuses are intelligible apart from consideration of such expression—as always, talk of what is implicit involves tacit reference to the process by which it can be made explicit.) The perspectival character of the expression of conceptual contents is a structural manifestation of the peculiarly intimate connection between semantics and pragmatics mandated by methodological pragmatism. In such a context, particular linguistic phenomena can no longer reliably be distinguished as ‘pragmatic’ or ‘semantic’.

A prime example is that of anaphora. This chapter and the previous one have described the crucial expressive role played by anaphoric connections in linking inherently unrepeatable performances (paradigmatically those involving demonstratives, indexicals, and even proper names) into repeatable structures that can express conceptual contents by being governed by (having their significance determined by) indirectly inferential substitutional commitments. It is anaphoric chains that tie different perspectives together and make it possible for scorekeepers to correlate them, structuring different pragmatic significances according to unified, though perspectival, semantic contents. Is anaphora a pragmatic or a semantic phenomenon? What about the distinction between *de re* and *de dicto* uses of linguistic expressions to specify the contents of ascribed commitments? Neither the latter, which expresses differences between scorekeeping perspectives, nor the former, which secures connections among them, is happily assimilated to either category. From the vantage point provided by this approach, philosophers have on the one hand maintained a tolerably clean line between semantics and pragmatics only by largely ignoring anaphoric phenomena (while lavishing attention on demonstrative and indexical ones that presuppose them) and have on the other hand been precluded by commitment to such a discrimination from a proper appreciation of the expressive function and significance of weak *de re* ascriptional locutions.

### *3. The Objectivity of the Norms Governing the Application of Concepts Is Part of Their Social-Perspectival Form*

Methodological pragmatism expresses the recognition that semantics is not ultimately intelligible apart from pragmatics; semantic con-

tents can be studied in abstraction from the pragmatic scorekeeping significance of altering deontic statuses and performing speech acts that have those contents, only provisionally and temporarily, subject to a promise of subsequent reunification. For commitments and the performances that express them are conceptually contentful in virtue of the role they play in the game of giving and asking for reasons. Their contents are perspectival because the inferential articulation of deontic scorekeeping practice has an irreducible social dimension. The social aspect of inferential articulation is apparent in the variations of substitutional commitments from one interlocutor to another, which correspond to different identity beliefs, and in the anaphoric commitments that link speech acts by different interlocutors. In fact, the perspectival relation between semantic contents and pragmatic significances reflects a prior perspectival relation between deontic *statuses* and deontic *attitudes*. For deontic statuses are *instituted* by the same essentially social scorekeeping practice of adopting discursive deontic attitudes that *confers* conceptual content on them. The basic phenomenon is the social-perspectival character of discursive *norms*—in deontic terms, the way in which *being committed* is to be understood in relation to being *taken* to be committed, by scorekeepers occupying different deontic points of view.

The discussion of the social-perspectival character of the representational dimension of semantic contents, expressed explicitly in the distinction between *de re* and *de dicto* ascriptions, can shed light on important aspects of discursive scorekeeping social practices—aspects that have necessarily been left somewhat obscure in the discussion of pragmatics up to this point. Thus it is appropriate to look once more at the relations envisaged between deontic statuses, which are the counters in terms of which discursive score is kept, and the deontic attitudes, whose adoption and alteration constitute the activity of scorekeeping that has been described as instituting those statuses. In particular, the expressive and explanatory raw materials are finally in place to address the fundamental issue of the *objectivity* of conceptual norms. It was pointed out already in Chapter 1 that it is a critical criterion of adequacy on any account of concepts that it make sense of a distinction between how they *are* applied in fact, by anyone or everyone, and how they *ought* to be applied—how it would be *correct* to apply them.

Failure to satisfy this criterion of adequacy is one of the primary objections that was offered against what were characterized as *I-we* construals of the social practices in which conceptual norms are implicit. Such construals fund a distinction between what particular individuals *treat* as or *take* to be a correct application of a concept, on the one hand, and what *is* a correct application, on the other, by contrasting individual takings with communal ones. This is the standard way of understanding objectivity as intersubjectivity. The cost of adopting this way of understanding the significance of the social dimension of discursive practice is, unacceptably, to lose the capacity to make sense of the distinction between correct and incorrect claims or

applications of concepts on the part of the whole community. This sort of view was criticized on the grounds that it depends on an illegitimate assimilation of linguistic communities to the individuals who participate in them—treating communities as producing performances and assessing them, undertaking commitments and attributing them. Besides its commitment to this mythological conception of communities, this view was found to be objectionable in that it avoids *regulism* (finding norms only in the explicit form of rules) only by falling into the complementary error of *regularism*, by rendering norms implicit in practice as mere regularities—albeit regularities of communal rather than individual practice.<sup>92</sup> Even if these further complaints are bracketed, however, the fact remains that an essential part of the representational dimension of our concepts—the way they purport to apply to an objective world—is that they answer for the ultimate correctness of their application not to what you or I or all of us *take* to be the case but to what actually *is* the case. Part of what it is for our concepts to be *about* an objective world is that there is an *objective* sense of correctness that governs their application—a sense of appropriateness that answers to the objects to which they are applied and to the world of facts comprising those objects. Even communally sanctioned takings or regularities of takings of what is correct concept application that are universal within a community can still be *mistakings*: even if all of us agree and always will agree that the mass of the universe is small enough that it will go on expanding forever, the possibility remains that we are all wrong, that there is sufficient matter undetected by us to make it collapse gravitationally.

One of the central challenges of an account of conceptual norms as implicit in social practice is accordingly to make sense of the emergence of such an *objective* notion of correctness or appropriateness of claims and applications of concepts. To begin with, the possibility of making such a notion intelligible in the present context appears to be threatened by the ontological disparity between concepts and the objects they apply to. For concepts are essentially perspectival, yet if they are to be *objectively* true or false of objects, those objects must be understood as nonperspectival in a strong sense. It makes no sense to specify or express a propositional or other conceptual content except from some point of view—which is subjective, not in a Cartesian sense, but in the practical sense that it is the point of view of some scorekeeping subject. How then is one to understand such propositional contents as answering for their truth, such conceptual contents as answering for their correct application, to some *one way* things are? Traditional philosophy says that beliefs are many, but the truth is one; the same point arises here in the contrast between scorekeeping perspectives, which are many, and the world, which is one. It is part of the representational purport of conceptual contents that thought and talk give us a perspectival grip on a nonperspectival world. To say that objects and the world of facts that comprises them are what they are regardless of what anyone takes them

to be requires a notion of *objective correctness* of claiming and concept application that is *not* perspective-relative in the way that what is *taken* to follow from what is.

In fact, however, understanding what is expressed by assessments of the objective correctness of applications of concepts requires appeal neither to nonperspectival facts (= true propositional contents) nor to community-wide commitments to propositional contents. Rather, the distinction between claims or applications of concepts that are objectively correct and those that are merely taken to be correct is a structural feature of each scorekeeping perspective. Indeed, the required notion of objective correctness is just what is expressed by *de re* specifications of the conceptual contents of ascribed commitments. For ascriptions in the *de re* style specify the objects that determine the truth or falsity—that is, the objective correctness—of the ascribed claim. Suppose the Constable has said to the Inspector that he himself believes that the desperate fugitive, a stranger who is rumored to be passing through the village, is the man he saw briefly the evening before, scurrying through a darkened courtyard. Suppose further that according to the Inspector, the man the Constable saw scurrying though the darkened courtyard is the Croaker, a harmless village character whom no one, least of all the Constable (who knows him well), would think could be the desperate stranger. Then the Inspector can identify the objective representational content of the Constable's claim by an ascription *de re*: "The Constable claims of the Croaker (a man who could not possibly be the fugitive) that he is the fugitive." Of course he does not take it that the Constable claims *that* the Croaker (a man who could not possibly be the fugitive) is the fugitive. The Constable claims only *that* the man he himself saw scurrying through a darkened courtyard is the fugitive. For the Inspector, the contrast between the *de re* and the *de dicto* content specifications is the contrast between saying what the Constable has in fact, willy-nilly, undertaken commitment to—what object his claim is *about*, in the sense that matters for assessments of truth—on the one hand, and what the Constable *takes* himself to be committed to, acknowledges, on the other hand.

Given the way the world actually is (according to the Inspector)—in this case, given who actually was scurrying through the courtyard—the Constable has without realizing it committed himself to a claim that is true if and only if the Croaker is the fugitive. That is the objective content of his commitment. The *de dicto* ascription specifies the content of that same commitment *subjectively*, that is, from the point of view of the one who acknowledges that commitment. In keeping the two correlated sets of books on the Constable that are made explicit as ascriptions in *de re* and *de dicto* form, the Inspector implicitly distinguishes between the deontic *status* undertaken by the Constable and the deontic *attitude* adopted by the Constable. This is the distinction between what the Constable is *committed* to by what he acknowledges and what he *acknowledges*. The difference arises

because one is committed to the inferential consequences of what one acknowledges, but one may nevertheless not acknowledge those consequences. What follows from the claim expressed by a certain sentence, an integral part of its content, can be assessed (by a scorekeeper) either objectively—using facts, true claims, as auxiliary hypotheses—or subjectively, using as auxiliary hypotheses only claims that the one who endorses the original claim acknowledges commitment to. It is this difference that is expressed by the difference between ascriptions *de re* and ascriptions *de dicto*.

As the terms are used here, to have *undertaken* a commitment is just to *be committed*; it is for a commitment to be in force. One can count as having adopted this attitude, and so as occupying this status, even if there is no overt performance that is the undertaking in virtue of which one is committed. Some commitments (and entitlements) we just come with, by default. In order to understand how, in spite of this coincidence, status and attitude can nonetheless be categories distinguished from the point of view of every scorekeeper, one must appreciate the perspectival character of deontic statuses. What is causally efficacious (for instance in action and perception) is the attitude of *acknowledging* a commitment. Commitments may be undertaken by default, by acknowledgment, or consequentially, as a result of what one does acknowledge. Indeed the content of the commitment one undertakes by a speech act of acknowledgment (assertion) depends on these consequences. Acknowledging a commitment can be identified with attributing it to oneself.<sup>93</sup> Thus undertaking a commitment is to be distinguished from attributing it to oneself, which is only one species of that attitude. Attribution can be seen to be the fundamental deontic attitude. Undertaking commitments can be understood in terms of attributing them if the social articulation of scorekeeping attributions is kept in mind: an interlocutor can count as having undertaken a commitment (as being committed) whenever others are entitled—perhaps in virtue of that interlocutor's performances—to attribute that commitment.

The attributing of commitments (and the kind of self-attribution that is acknowledging them) is an *immediate* deontic attitude, while undertaking commitments includes not only the immediate deontic attitude of acknowledging them but also *consequential* deontic attitudes. Deontic statuses are just such consequentially expanded deontic attitudes. But the consequences of a particular acknowledgment are assessed differently from different perspectives—that is, by different attributors. From the vantage point of any particular scorekeeper, what one is *really* committed to by an acknowledgment (paradigmatically the assertive utterance of a sentence), what *really* follows from the claim (and hence its objective content), is to be assessed by conjoining it with truths—that is, statements of fact. But what plays this role for a scorekeeper is the set of sentences by the assertion of which the scorekeeper is prepared to *acknowledge*, and so undertake, doxastic commitment. Thus immediate deontic attitudes determine consequential ones, and so

deontic statuses, from each scorekeeping perspective. From the point of view of each scorekeeper, there is for every other interlocutor a distinction between what commitments that individual *acknowledges* and what that individual is really committed to—between (immediate) deontic attitudes and deontic status (or consequentially expanded attitudes). But how this line is drawn in particular varies from scorekeeper to scorekeeper. This perspectival structure is what has been implicitly appealed to throughout by talk of the *institution* of deontic statuses by deontic attitudes.

In this way, every scorekeeping perspective maintains a distinction in practice between normative status and (immediate) normative attitude—between what is objectively correct and what is merely *taken* to be correct, between what an interlocutor is actually committed to and what that interlocutor is merely *taken* to be committed to. Yet what from the point of view of a scorekeeper is objectively correct—what from that perspective another interlocutor is actually committed to by a certain acknowledgment—can be understood by us who are interpreting the scorekeeping activity entirely in terms of the immediate *attitudes*, the acknowledgments and attributions, of the scorekeeper. What appears to the scorekeeper as the distinction between what is objectively correct and what is merely taken to be or treated as correct appears to us as the distinction between what is acknowledged by the scorekeeper attributing a commitment and what is acknowledged by the one to whom it is attributed. The difference between objective normative status and subjective normative attitude is construed as a social-perspectival distinction between normative attitudes. In this way the maintenance, from every perspective, of a distinction between status and attitude is reconciled with the methodological phenomenism that insists that all that really needs to be considered is attitudes—that the normative statuses in terms of which deontic score is kept are creatures instituted by the (immediate) normative attitudes whose adoption and alteration is the activity of scorekeeping.

On this account, objectivity is a structural aspect of the social-perspectival *form* of conceptual contents. The permanent possibility of a distinction between how things *are* and how they are *taken* to be by some interlocutor is built into the social-inferential articulation of concepts. The distinction is in the first instance available to each scorekeeper regarding the commitments of *others*, of those to whom the scorekeeper *attributes* commitments, for it just reflects the difference in perspective provided by the different inferential significance of claims in the context of auxiliary premises provided by commitments acknowledged (and so undertaken) by the scorekeeper and commitments that are attributed to others by that scorekeeper. Although grounded in essentially social, other-regarding scorekeeping, however, the possibility of a distinction between how things actually are and how they are merely taken to be by some interlocutor remains a structural feature, even, as will be seen below, in the case of attributions to oneself.

#### 4. *I-Thou Symmetry of Subjective Discursive Attitudes and Objective Discursive Statuses*

It was pointed out above that the account of (weak) *de re* ascriptions as expressing the difference in social perspective between *attributing* a commitment (in the paradigmatic case, a doxastic one) and *undertaking*, by acknowledging, a commitment (in the paradigmatic case, a substitutional one) just generalizes the previous account of assessments of truth to include nonpropositional conceptual contents. The objective representational content of someone's belief or remark can be specified in a *de re* ascription:

Max believes of quinine that it prevents jungle fever (Max represents quinine as preventing jungle fever).

Such a characterization specifies the objects that must be consulted in order to assess the objective truth or falsity of the belief—what it is true of, if it is true at all. This connection emerges particularly clearly if what is exported to *de re* position in the ascription is not just a singular term, predicate, or sortal but (a nominalized version of) an entire sentence:

The senator believes of the first sentence of the Communist Manifesto that it is true.

In these cases, that the expressions occurring in the scope of the *de re* 'of' express what matters for determining the *truth* of the ascribed claim is made explicit by the fact that the ascription-structural anaphoric dependent left as a trace of the exportation within the scope of the *de dicto* 'that' is a prosentence, which accordingly is formed using 'true'.

Our practical grasp of the objective dimension of conceptual norms—normative assessments of the objective truth of claims and objective correctness of applications of concepts—consists in the capacity to coordinate in our scorekeeping the significance a remark has from the perspective of the one to whom the commitment it expresses is attributed and its significance from the perspective of the one attributing it. This requires recognizing the different specifications of the same claim that correspond to extracting its inferential consequences and antecedents in the context of other commitments that are acknowledged as true by the scorekeeper, on the one hand, and extracting them in the context of other commitments acknowledged by the target of that scorekeeping, on the other. This is just the difference between employing as auxiliary hypotheses claims that are true (according to the scorekeeper) and employing as auxiliary hypotheses claims that are merely held true (according to the scorekeeper) by the interlocutor whose commitments are being assessed. Thus every scorekeeping perspective incorporates a distinction between what is (objectively) true and what is merely (subjectively) held true.

This account accordingly provides one detailed way of understanding what

lies behind some of Davidson's most important claims, quoted with approval back in Chapter 3:

If this account of radical interpretation is right, at least in broad outline, then we should acknowledge that the concepts of objective truth and of error necessarily emerge in the context of interpretation. The distinction between a sentence being held true and being in fact true is essential to the existence of an interpersonal system of communication . . . The concept of belief thus stands ready to take up the slack between objective truth and the held true, and we come to understand it just in this connection . . . Someone cannot have a belief unless he understands the possibility of being mistaken, and this requires grasping the contrast between truth and error—true belief and false belief. But this contrast, I have argued, can emerge only in the context of interpretation, which alone forces us to the idea of an objective, public truth.<sup>94</sup>

The notion of objective truth conditions makes explicit what is implicit in our grasp of the possibility of mistaken belief and so of the distinction between what is *merely* held (true) and what is *correctly* held (true). It emerges only in the context of interpretation—that is, discursive scorekeeping—because that is the practical activity in which the commitments acknowledged (held true) by one interlocutor are compared and contrasted with those acknowledged (held true) by another, the scorekeeper who attributes the first set.<sup>95</sup>

The objectivity of conceptual norms—the dimension of correctness in the application of conceptual contents assessed according to how the things they are applied to actually are rather than to how they are *taken* to be—is here given an essentially *social* construal. Objectivity appears as a feature of the structure of discursive intersubjectivity. But traditionally intersubjectivity has been understood in the *I-we* way, which focuses on the contrast between the commitments of *one* individual and the commitments of the *community* (collectively), or those shared by *all* individuals (distributively). In the broadly Davidsonian account offered here, by contrast, intersubjectivity is understood in the perspectival *I-thou* fashion, which focuses on the relation between the commitments *undertaken* by a scorekeeper interpreting others and the commitments *attributed* by that scorekeeper to those others. From the point of view of this latter sort of understanding of intersubjectivity, *I-we* accounts mistakenly postulate the existence of a *privileged* perspective—that of the 'we', or community. The objective correctness of claims (their truth) and of the application of concepts is identified with what is endorsed by that privileged point of view. The identification of objectivity with intersubjectivity so understood is defective in that it cannot find room for the possibility of error regarding that privileged perspective; what the community *takes* to be correct *is* correct. The community, it may be said, is *globally* privileged.

Understanding the sort of privilege that is at issue in these claims can, like

any other normative status, be approached by distinguishing the *circumstances* under which it is appropriately acquired, on the one hand, and the *consequences* of possessing it, on the other. The consequences of a perspective being privileged in the relevant sense—what that privilege consists in—is that one *cannot* hypothetically adopt a third-person point of view with respect to it and evaluate it from the outside by contrasting what it merely *takes* to be true with what actually *is* true. Thus a scheme that identifies truth with what is taken true by all the members of a community, or by the experts in a community, or what will always be taken true by them, or what would be taken true by them under some ideal conditions for inquiry, thereby identifies circumstances of application for a perspective to be privileged in just this sense. Truth is *defined* as how things are according to a discursive perspective that meets certain conditions; it is settled in advance that any perspective from which a distinction appears between how things seem from such a privileged point of view and how things in fact are is itself without any authority at all. According to an account that incorporates a global privilege of this sort, the distinction between what is true and what is taken-true does not vary from one perspective to another.

By contrast, according to the *I-thou* construal of intersubjectivity, each perspective is at most *locally* privileged in that it incorporates a structural distinction between objectively correct applications of concepts and applications that are merely subjectively taken to be correct. But none of these perspectives is privileged in advance over any other. At first glance this egalitarian attitude may seem just to put off the question of what is really correct—so that one must choose between relinquishing the notion of objectivity by acquiescing in a regress of attributors assessing attributions and securing it by acknowledging an infallibly authoritative taker (for instance, the community). The alternative is to reconstrue objectivity as consisting in a kind of perspectival *form*, rather than in a nonperspectival or cross-perspectival *content*. What is shared by all discursive perspectives is *that* there is a difference between what is objectively correct in the way of concept application and what is merely taken to be so, not *what* it is—the structure, not the content.

The crucial feature of the perspectival structure of objectivity is the *symmetry* of state and attitude between ascriber and the one to whom a commitment is ascribed. The author of a *de re* ascription treats it as specifying the objective representational content of the attributed commitment (as specifying the status in question, and what it is really about), and its *de dicto* correlates as specifying the subjective attitude that the target of the attribution has toward that state—what the one whose status it is *takes* it to be about. From the point of view of the one to whom the status is ascribed, however, things are the other way around. The attributee (a woman, let us say) treats her own specifications of the content of the commitment she has undertaken as authoritative—she takes it that she knows what she means,

what she has committed herself to, and what she is talking about. The *de re* ascription by the other (a man, let us say), insofar as its terms diverge from those appearing in the corresponding *de dicto* ascriptions (which answer directly to what the attributee acknowledges) expresses from her point of view simply the attitude of the ascriber—his version of her actual status. His divergent *de re* characterization of the content of the commitment she undertook just indicates a difference between the commitment he *attributes* to her, and what she actually undertook. The difference between status and attitude, between commitment and its attribution, between what is correct and what is merely taken to be correct, between objective content and subjective view of it, is as essential a feature of the attributee's point of view as it is of the attributor's—though the author and the target of ascriptions line up these distinctions with that between the author's *de re* and *de dicto* attributions in a complementary way. In each case, from each point of view, there are distinct attitudes of taking or treating as correct and taking or treating as taken or treated as correct. This symmetric pair of perspective types, that of attributor and attributee, each maintaining this fundamental normative distinction, is the fundamental social structure in terms of which communities and communal practice are to be understood. This symmetric *I-thou* social distinction is presupposed by the *I-we* social distinction appealed to by the other kind of construal of intersubjectivity.

Its symmetry ensures that no one perspective is privileged in advance over any other. Sorting out who should be counted as correct, whose claims and applications of concepts should be treated as authoritative, is a messy retail business of assessing the comparative authority of competing evidential and inferential claims. Such authority as precipitates out of this process derives from what various interlocutors say rather than from who says it; no perspective is authoritative as such. There is only the actual practice of sorting out who has the better reason in particular cases. The social metaphysics of claim-making settles what it means for a claim to be true by settling what one is doing in *taking* it to be true. It does not settle which claims *are* true—that is, are *correctly* taken to be true. That issue is adjudicated differently from different points of view, and although these are not all of equal worth, there is no bird's-eye view above the fray of competing claims from which those that deserve to prevail can be identified, nor from which even necessary and sufficient conditions for such deserts can be formulated. The status of any such principles as probative is always itself at issue in the same way as the status of any particular factual claim.

### 5. Objectivity Proofs

The objectivity of conceptual norms, it has been claimed, is a reflection of the perspectival distinction between undertaking and attributing inferentially articulated commitments. Attributions are made explicit by

ascriptions; the perspectival structure that objectivity consists in can be expressed explicitly in propositionally contentful commitments. Using pieces of logical vocabulary whose expressive roles have been specified systematically in these pages, the challenge to *I-we* construals of conceptual norms that was put in Chapter 1 is to show that it is the case neither that

(i)  $\{p\}[p \rightarrow \{S\} \{S \text{ claims that } p\}]^{96}$

nor, perhaps more threateningly, that

(ii)  $\{p\}[\{S\} \{S \text{ claims that } p\} \rightarrow p]$ .

These entailments together would codify the equivalence of  $p$ , the obtaining of the state of affairs represented by  $p$ , its being *true* that  $p$  (given the prosentential rendering of the expressive role of ‘true’), to universal or communal *commitment* to  $p$ . One of the motivating criteria of adequacy of the present account is that it underwrite no such collapse of an objective state of affairs into a communal attitude. It should be obvious that neither of these conditionals holds on the *I-thou* deontic scorekeeping account, given the story that has been told about the expressive role of propositional-attitude-ascribing locutions as explicating commitments and conditionals as explicating inferences. But to say that is not to say that there is no point in offering *proofs*, which can help engender illumination as much as conviction.

The quantified conditional (i) might be called the *No Communal Ignorance Condition*. It is surely not plausible, and so not a desirable consequence of any theory. It is worth working through an argument in a bit of detail here, however, since the principles appealed to will be of use in more interesting demonstrations below. The conditionals that are of interest here are those that express incompatibility entailments (introduced in Chapter 2). For these express a *modal* force,  $p \rightarrow q$  saying in effect that it is not *possible* for  $p$  to obtain without  $q$  obtaining. And this is the sense desired for the objectivity-denying principles (i) and (ii) above. The official definition of this sort of conditional is that  $p \rightarrow q$  just in case everything incompatible with  $q$  is incompatible with  $p$ —so that ‘Wulf is a dog’ incompatibility-entails ‘Wulf is a mammal’, because everything incompatible with the latter is incompatible with the former.<sup>97</sup> To show that the conditional in (i) does not hold, then, it is enough to find some claim  $q$  that is incompatible with its consequent, and not with its antecedent.

To see that (i) is not a consequence of the present account, it suffices to notice that nothing in that account requires that one who *undertakes* a commitment to  $p$ , and is furthermore *entitled* to that commitment, therefore must *attribute* that commitment to *anyone* else (never mind to *everyone* else). Since propositional contents are incompatible just in case commitment to one precludes entitlement to the other,<sup>98</sup> this means that  $p$  will not in general be incompatible with  $\neg\{S \text{ claims that } p\}$ —except for a few very special cases.<sup>99</sup> So it is possible to pick some particular  $p^\alpha, S^\alpha$  such that it is not the

case that  $p^\alpha/\sim(S^\alpha$  claims that  $p^\alpha)$  (using ' $p/q$ ' to express  $p$ 's incompatibility with  $q$ ). This is to say that  $\sim(S^\alpha$  claims that  $p^\alpha)$  is not incompatible with the antecedent of the conditional in (i). But it is incompatible with the consequent:  $(S)(S$  claims that  $p^\alpha)/\sim(S^\alpha$  claims that  $p^\alpha)$ . For the right-hand side just denies a substitution instance of the left-hand side. Thus for any (nonascriptional)  $p^\alpha$ , this recipe shows how to produce a  $q^\alpha$  that is incompatible with the consequent of (i) but not with the antecedent, which shows that the entailment does not obtain for any (nonascriptional)  $p$ .

The quantified conditional (ii) might be called the *No Communal Error Condition*. It is also not plausible, and so not a desirable consequence of any theory. It is, however, an unavoidable consequence of *I-we* theories of conceptual norms. (As pointed out in Chapter 1, this is one basis on which McDowell rightly criticizes Wright's and Kripke's Wittgenstein.) To show that this conditional is not a consequence of the deontic scorekeeping account, it is necessary to find a claim that will defeat the incompatibility entailment it makes explicit. Such a defeating claim is one that is incompatible with the consequent of the conditional, but not with its antecedent. Thus it suffices to find a  $p^\alpha$  and a  $q^\alpha$  such that  $q^\alpha/p^\alpha$  and *not*  $q^\alpha/(S)(S$  claims that  $p^\alpha)$ .

Intuitively, what is wanted is just something that is incompatible with a claim but not with everyone's believing the claim. Given that actual conceptual roles have been officially defined only for logical locutions, a bit of subtlety is required to produce a pair of claims that *provably* stand in the desired incompatibility relations. Here is a recipe for producing such a pair. Let  $p^\alpha$  be:

$$(p)(\sim!\exists Dx \text{ claims that } p),$$

which claims that the unique  $D$  does not have any beliefs; and let  $q^\alpha$  be:

$$\exists Dx \text{ claims that } p^\alpha,$$

that is:

$$\exists Dx \text{ claims that } (p)(\sim!\exists Dx \text{ claims that } p),$$

which claims that the unique  $D$  believes that the unique  $D$  does not have any beliefs (so the unique  $D$  believes this of the unique  $D$ —but not, of course, *as* the unique  $D$ ). For instance, ' $\exists Dx$ ' might be 'the object weighing 100 lbs. or more that made the scrape marks on the flagstones in the garden'. Then it might be that John endorses  $p^\alpha$  because he believes that the scrape marks were made by a rock rolling down the hillside into the garden. Since rocks do not have beliefs, he believes that what made the marks does not have beliefs. But I, who take it that John unknowingly made those marks with his hobnailed boots, may endorse  $q^\alpha$ .

Notice, first, that  $q^\alpha/p^\alpha$ , since  $q^\alpha$  attributes a belief to  $\exists Dx$  and  $p^\alpha$  denies that  $\exists Dx$  has any beliefs. But second, it is *not* the case that  $q^\alpha/(S)(S$  claims

that  $p^\alpha$ ). For  $q^\alpha$  is just an instance of that quantified claim and so is entailed by it: if  $(S) \{S \text{ claims that } p^\alpha\}$ , then  $\exists x D x$  claims that  $p^\alpha$ .<sup>100</sup> Thus  $q^\alpha$  is incompatible with the consequent of the conditional (ii), but not with the antecedent, and so the conditional does not hold. Not all the counterexamples to (ii) have this form, but the pattern indicated shows how to construct a whole family of them (notice that it is not even essential that the singular terms involved be definite descriptions).

The explicit expression of the structural distinction of perspective between undertaking and attributing commitments is a general acknowledgement by each interlocutor of the possibility for *any S* and *p* that *S* believes that *p* but it is not true that *p*. Rehearsing rigorously the coherence of such an attitude in particular cases does not show, however, that this principle extends even to the self-attributions expressed by self-ascriptions. It does not yet show the coherence of the attitude that would be endorsed by saying:

It is possible that (*I* believe that *p* and it is not true that *p*).

It is the coherence of such an acknowledgment that one would expect to be hard to fund on an *I-thou* social account of conceptual norms.

The equation of what one is saying in claiming that *p* with a claim about universal or communal belief threatens *I-we* social accounts of conceptual norms; it is no surprise to find that it is not a threat to the *I-thou* variety. What would be more impressive is to see a similar sort of wedge driven between the use of '*p*' and '*I* believe [or claim] that *p*' on the deontic score-keeping account. That there is an apparent difficulty in avoiding a strong equivalence between these two very different claims is evident from the fact that in any language with the expressive resources both of propositional-attitude-ascribing locutions and of strong first-person pronouns, each interlocutor will be committed to what is expressed by '*I* claim [am committed to the claim] that *p*' when and only when that interlocutor is committed to what is expressed by '*p*'. An inability to distinguish these claims would show that there is a problem with the *local*, perspectival privileging of each score-keeper's repertoire of commitments that corresponds, in the *I-thou* rendering of discursive social practice, to the failure to account for the objectivity of conceptual norms that afflicts the *I-we* rendering of social-discursive practice in virtue of its *global* (or quantificational) privileging of the communal repertoire.

Nonetheless, utterances of '*I* claim that *p*' and '*p*', though having the same assertibility conditions, express different propositional contents; as the point is usually put, they have different truth conditions. One wants to say that their equivalence is a pragmatic rather than a semantic matter—that is, that it derives from features of *claimings*, rather than of what is *claimed*. The Geach-Frege test for distinguishing these was endorsed and exploited in Chapter 5. It depends on the fact that pragmatic force or significance (para-

digmatically the assertional variety) attaches to a whole speech act, not merely to the utterance of an embedded component of an asserted sentence. Thus one can strip off the pragmatic consequences of claiming by embedding a content, paradigmatically as the antecedent of a conditional. The practical cash-value of the different truth conditions associated with ' $p$ ' and ' $I$  claim (or believe) that  $p$ ' is that they behave differently as the antecedents of conditionals.<sup>101</sup> In particular,  $p \rightarrow p$  holds generally, while ' $I$  claim that  $p$ '  $\rightarrow$   $p$  does not. Thus the real challenge to an *I-thou* social account of conceptual norms is posed, not by the conditionals (i) and (ii), but by:

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

and, more threateningly,

(iv)  $\{p\} [\{\text{I claim that } p\} \rightarrow p]$ .

Showing that the conditional in (iii) (the *No First-Person Ignorance Condition*) does not hold in the deontic scorekeeping account requires finding a claim that is incompatible with what is expressed by ' $I$  claim that  $p$ ', but not with  $p$  itself. Consider ' $I$  do not claim that  $p$ ' (or anything that entails it, such as 'I make no claims' or 'I do not exist'). This is evidently incompatible with the ascription that appears in the consequent, since it is just the denial of that claim. Is it incompatible with the antecedent? That I cannot be entitled to both  $p$  and ' $I$  do not claim that  $p$ ' is just the tenor of Moore's paradox. But this does not mean that the contents of these claims are incompatible. For if they were incompatible, then *anyone* committed to the one content would be precluded from entitlement to the other. But if I endorse the explicit ascriptions

$S$  claims that I do not claim that  $p$ , and  $S$  claims that  $p$ ,

I have not attributed incompatible claims to  $S$ .<sup>102</sup> That is, I could take  $S$  to be *entitled* to both these commitments. Thus they are not incompatible. For one involves what commitments  $S$  *attributes* to me, and the other involves what commitments  $S$  *undertakes*, and these do not collide.<sup>103</sup>

Ascriptional locutions make explicit the possibility of taking up hypothetically a sort of third-person scorekeeping attitude toward my own present commitments and entitlements (much as I must do for my *past* commitments and entitlements in any case). Here such ascriptions show that what precludes entitlement both to the claim that  $p$  and to my denial of a self-ascription is a pragmatic matter concerning attitudes, not a semantic matter concerning the contents to which they are addressed. My denial that I claim that  $p$  collides with what I am *doing* (claiming that  $p$ ), not with what I am *saying* (that  $p$ ). To distinguish these, I must look at someone *else's* attitudes toward the *same* contents. The social dimension of their inferential articulation is essential to the semantic nonequivalence of ' $p$ ' and ' $I$  claim (hold) that  $p$ '.

The intuition underlying the distinction between what is expressed by  $\lceil p \rceil$  and  $\lceil I \text{ believe (or claim) that } p \rceil$  on the deontic scorekeeping account is straightforward. Although whenever one is *committed* to one of these contents, one is committed to the other, the same does *not* go for entitlements. The easiest way to see this is to think about someone with incompatible commitments  $p^\alpha$  and  $q^\alpha$ . One of the distinctive virtues of understanding intentional states in the first instance in terms of normative statuses rather than in causal-functional terms is precisely that sense can be made of the undertaking of commitments with incompatible contents. The subject of such commitments will not be *entitled* to any of them, but this in no way precludes nonetheless being committed. If  $p^\alpha/q^\alpha$ , then if I am committed to both, I cannot be entitled to either. But I can still be entitled to the ascriptional commitment expressed by  $\lceil I \text{ claim that } p^\alpha \rceil$ , for this just correctly ascribes one of the commitments I in fact have. Thus though *commitment* to what is expressed by  $\lceil p \rceil$  and by  $\lceil I \text{ claim that } p \rceil$  go hand in hand, *entitlements* to those commitments do not.

This idea provides a general recipe for producing counterexamples to the conditional in (iv) (the *No First-Person Error Condition*). Such defeasors are claims that are incompatible with the consequent of that conditional, but not with its antecedent. But *any* claim  $q$  such that  $q/p$  can serve this role. For it will not be the case that  $q/\lceil I \text{ claim that } p \rceil$ . I *can* be entitled to the ascriptional commitment expressed by  $\lceil I \text{ claim that } p \rceil$ , even if I am committed also to  $q$ , which is incompatible with  $p$ . Then I just have incompatible commitments (namely  $p$  and  $q$ ) and am correctly ascribing one of them to myself (perhaps my commitment to  $p$  arises as a consequence of another commitment I have overtly avowed by asserting it, and I have just realized this about my deontic status). Thus (iv) does not hold.

The result of these discussions of conditionals (i)–(iv) is that the perspectival notion of objectivity that arises out of the deontic scorekeeping account of inferentially articulated conceptual contents can be *shown* not to collapse into *I-we* quantificational intersubjectivity, and it can be *shown* not to collapse into a mere privileging of one's own perspective. Admittedly, this is to demonstrate only that a fairly weak necessary condition on a conception of objectivity has been satisfied—that the contents of ordinary claims, such as “Snow is white” and “The mass of the universe is great enough to produce eventual gravitational collapse,” are not equivalent to those of any claims about who is committed to what. But it is no small thing to be able to *prove* that a notion of objectivity that satisfies even these minimal constraints can be funded starting with a notion of *social* practice. Facts are true claims—‘claims’ in the sense of what is *claimed*, which does not depend on what people do, not of *claimings*, which do. The demonstrations just presented define a robust sense in which the facts as construed in this work are independent of what anyone or everyone is committed to. The *claim-making* practices described here are accordingly properly understood as making pos-

sible genuine *fact-stating* discourse, for they incorporate practices of assessing claims and inferences according to their *objective* correctness—a kind of correctness that answers to how things actually are, rather than to how they are *taken* to be, by anyone (including oneself) or everyone.

## 6. Conclusion

A central task of this work has been to show how such objective conceptual norms can be made intelligible in terms of social-deontic scorekeeping practices governing attitudes of taking or treating oneself and others as having inferentially articulated commitments and entitlements. The fine structure of the arguments presented above shows that essential appeal is made to both the *social* articulation of practical attitudes into attributing and undertaking (in addressing (iii)) and the *deontic* articulation of normative statuses into commitments and entitlements (in addressing (iv)) in the *pragmatics*, in showing the *objectivity* of the *semantic* (inferential) contents those practices confer. These are, of course, the two central structural elements in terms of which the deontic scorekeeping pragmatics has been elaborated.

The avowed aim of the model of discursive practice motivated and introduced in Part 1 and developed in detail in Part 2 has been to describe scorekeeping practices that are sufficient to confer various sorts of conceptual content: to begin with, fundamental nonlogical *propositional* contents, then the sorts of conceptual contents associated with predicates and singular terms (including pronouns, demonstratives, definite descriptions, and proper names), and finally the specifically *logical* expressive conceptual content of conditionals, negation, quantifiers, identity locutions, traditional semantic vocabulary, and ascription-forming operators, which help make explicit crucial implicit features of the inferential and social articulation of the scorekeeping practices that confer conceptual contents in the first place. A major criterion of adequacy of that account is that it explain how *objectivity* can arise as a structure within *intersubjectivity* construed in the *I-thou* fashion characteristic of the deontic scorekeeping approach to discursive practice. Satisfying this condition required showing how the *representational* dimension of conceptual contents—the sort of correctness of concept application that answers to how things are with the things represented and that contrasts in principle with the sort of correctness that answers only to how things are taken to be by some individual or whole community—can be understood in terms of the *inferential* articulation that defines those contents. The understanding that emerges from this discussion construes the objective representational dimension of conceptual contents as a reflection of the *social-deontic* articulation of the practices of giving and asking for reasons, which confer conceptual contents on the states, attitudes, performances, and expressions suitably caught up in them.

It turns out to be a consequence of the inferentialist way of conceiving conceptual content that it makes sense to specify the content of a state only from some point of view, relative to some set of collateral concomitant commitments, which can serve as auxiliary hypotheses in inferences involving it. For both the attributor and the attributee of any contentful state, there are two relevant sets of background commitments available in determining the practical significance (for what else one should go on to do or be committed to) of adopting a state with a specified content—that of the attributor and that of the attributee. Thus two socially related kinds of perspective are always in play, for each interlocutor. This fact, it has been claimed, secures and gives meaning to the possibility of a genuinely *normative* significance for the occurrence of contentful states. It is also what is expressed by *representational* idioms such as ascriptions *de re*. It is because of this fundamental social deontic structure, then, that propositional and other conceptually contentful states are always *representationally* contentful states. This analysis of the nature of the objective representational norms that govern the application of concepts makes it possible to see why *only* what plays a suitable role in essentially *social*, indeed *linguistic*, discursive deontic scorekeeping practices should count as *conceptually* contentful in the fundamental sense. The understanding of intentional or conceptual contentfulness that is finally arrived at vindicates the initial commitment to understanding discursive practice as social linguistic practice.

### **Appendix: The Construction and Recursive Interpretation of Iterated Ascriptions That Mix *De Dicto* and *De Re* Content Specifications**

The expressions that serve to specify the content of the commitments attributed by undertaking assertional commitment to basic ascriptions can play two different sorts of expressive role. In the regimented language employed here, this distinction of roles is marked by a distinction of two sorts of position in which content-specifying expressions can occur in ascriptions. Expressions may occur either in the scope of a 'that' operator (what is called, following the tradition, "*de dicto* occurrence") or in the scope of an 'of' operator ("*de re* occurrence"). The leading idea of the explanatory strategy developed in this chapter is that the significance of an expression's occurring in *de dicto* position is that the expressive commitment to the effect that the content of the attributed assertional commitment can properly be expressed by the use of that expression is *attributed* along with the ascribed assertional commitment, while the significance of an expression's occurring in *de re* position is that the expressive commitment to the effect that the content of the attributed assertional commitment can properly be expressed by the use of that expression is *undertaken*, along with the assertional commitment to

the whole ascription. The distinction between *what* is represented, the objective, *de re*, relational content of what is ascribed, on the one hand, and *how* it is represented, its subjective, *de dicto*, notional content, on the other, is based on this fundamental social-perspectival distinction of deontic attitudes (the distinction between undertaking a commitment and attributing one). One of the distinguishing characteristics of the present approach to propositional-attitude-ascribing locutions is understanding what is expressed by the difference between ascriptions *de dicto* and ascriptions *de re* as indicating the difference between two perspectives or attitudes an ascriber can adopt when specifying the content of an ascribed state, rather than as distinguishing two *kinds* of state (as Quine does), or two *components* in the content of any intentional state (as McGinn does).<sup>104</sup>

One of the criteria of adequacy of any account of this difference is its capacity to deal with *iterated* ascriptions—ascriptions of assertional commitments that themselves have ascriptional contents, expressions for which accordingly contain *embedded* ascriptional expressions. For interpretation of such compound expressions requires the recognition of many more than just two ways in which expressions can function in specifying the contents of ascribed commitments. It turns out to be straightforward to extend the social-perspectival account to handle the complexities of iteration. It is much less clear how the motivation behind distinguishing *de dicto* and *de re* ascriptions as ascriptions of different kinds of belief, or of different components of beliefs, fares once iterated ascriptions are taken onboard.

How, then, can all the iterated ascriptions be constructed in the regimented language of the scorekeeping model? To keep things under control, two simplifying assumptions are adopted—working at the level of types rather than tokenings, and only considering the case of singular-term exportation. These are straightforwardly dispensable in favor of more general formulations. Consider a basic nonascriptional content expressed using an *n*-adic predicate  $\Phi(x_1, x_2, \dots, x_n)$ . In ascribing commitment to a claim of this form, one might attribute the expressive commitments associated with the use of all of the terms used to specify the content, in the pure *de dicto* form

$S_2: S_1$  claims that  $\Phi(t_1, t_2, \dots, t_n)$ .

Or one might undertake all those expressive commitments and syntactically export all of the terms to *de re* position, in the pure *de re* form

$S_2: S_1$  claims of  $\langle t'_1, t'_2, \dots, t'_n \rangle$  that  $\Phi(it_1, it_2, \dots, it_n)$ ,

where each  $it_i$  is an ascription-structural anaphor dependent on  $t'_i$ . But such exportation need not be an all-or-none thing. Some terms may remain in *de dicto* position while others are removed to *de re* position, as in

Russell believed of Hölderlin's roommate that he was not a worthy successor to Kant,

where 'Kant' remains in *de dicto* position, but the other term has been exported to *de re* position, leaving the anaphoric trace 'he'. Thus arrayed between the pure *de dicto* and pure *de re* forms will be a variety of mixed ascriptions, in which some terms appear in *de dicto*, and some in *de re* position. So the general form of first level ascriptions is

$S_2: S_1 \text{ claims of } \langle t_1', t_2', \dots t_k' \rangle \text{ that } \Phi(t_1, t_2, \dots t_n),$

where  $k$  is less than or equal to  $n$ , and for all  $i$  less than or equal to  $k$ , there is a  $j$  less than or equal to  $n$  such that  $t_i$  is an ascription-structural anaphoric dependent of  $t_j'$ , symbolically: Depends ( $/t_i/, /t_j'/$ ). Since each term can appear in two positions, either exported or not, corresponding to an  $n$ -adic predication  $\Phi(x_1, x_2, \dots x_n)$  there will be  $2^n$  different first-level ascriptions.

Each of these ascriptions still contains  $n$  independent terms in its content-specifying regions. The *de dicto* positions will always contain  $n$  argument places. Each of them is filled either by an independent term or by an ascription-structural anaphoric pronoun. But corresponding to each anaphoric pronoun is exactly one term that is exported to *de re* position. Such exportation accordingly does not change the total number of terms occurring in the content specification. All the forms of first-level ascriptions of commitment to a nonascriptional claim involving  $n$  argument places can then be thought of as (ascriptionally complex)  $n + 1$  place predication. (The extra argument place is that which specifies the one to whom the commitment is ascribed; it occurs outside the content-specifying regions of the ascriptional expression.) These first-level ascriptions of the form  $\Psi(t_1, \dots t_n)$  can now themselves be treated as specifying the content of commitments that can be ascribed, by second-level ascriptions. A second-level ascription is one like

Russell claims that Hegel claims of Pluto that it does not exist,

in which the commitment ascribed is itself an ascriptional commitment. Clearly all of the terms that occur independently in the first-level ascriptional content  $\Psi(t_1, \dots t_n)$  (that is, all the terms except the ascription-structural anaphoric pronouns left as syntactic traces of terms exported to *de re* position) are available either to be left in what is *de dicto* position with respect to the outermost, second-level, ascriptional content-specification, or to be exported to the *de re* position of that outermost, second-level, ascriptional content-specification. Thus there can be second-level forms such as

$S_2 \text{ claims of } t_1 \text{ that } S_1 \text{ claims of it that } \Phi(it_1, t_2)$

and

$S_2 \text{ claims of } t_1 \text{ that } S_1 \text{ claims that } \Phi(it_1, t_2)$

that differ in that in the first what is exported to the second-level *de re* position occurred in *de re* position in the embedded first-level ascription as

well, whereas in the second what is exported to the second-level *de re* position occurred in *de dicto* position in the embedded first-level ascription.

Just as the first time around, exportation does not change the total number of independently occurring terms, and each of the  $n + 1$  independent terms available for possible exportation in the first-level ascriptions can either be exported or not. So for each of the  $2^n$  first-level ascriptions, there are  $2^{n+1}$  second-level ascription forms, or  $2^n \cdot 2^{n+1} = 2^{2n+1}$  second-level ascriptions (based on the nonascriptional  $n$ -adic expression  $\Phi(x_1, x_2, \dots, x_n)$ ) in all. For the general case of  $m^{\text{th}}$ -level iterated ascriptions, there will be  $2^{n+(n+1)+(n+2)+\dots+(n+m-1)}$  distinct ascription forms, which is  $2^{n \cdot m + (m^2 - m)/2}$ . If one treats *de dicto* and *de re* ascriptions as ascriptions of two different kinds of belief, then this is how many different kinds there are, not just two. And if one sees first-level ascriptions of the two sorts as specifying two different components of the content of the state that is attributed, then this is how many different components one is committed to discerning in the content of an  $m^{\text{th}}$ -level iterated ascription, not just two. The complication in this calculation results from the argument place for the target of the ascription—which behaves like a term in the scope of the *de re* operator ‘of’, except for not having ascription-structural anaphoric dependents. Putting those occurrences aside, an  $n$ -ary nonascriptional predication generates  $2^{n \cdot m}$  possible mixed ascription forms, where  $m$  is the number of iterated applications of the ascription-forming locution.

A criterion of adequacy of an account of the content of ascriptions is that it determine for each of these myriad iterated ascription-forms the pragmatic significance, in context, of undertaking commitment to an assertion with that form. The theory must offer a reading of each, specifying what an ascriber becomes committed to by asserting it. In the context of the sort of pragmatics or speech-act theory in play here, this means saying how the deontic score is changed by an ascriptional undertaking, which is to say what attitudes it expresses. For instance, looking only at a single iteration of ascribing operators and at a one-place predicate, two of the forms that must receive interpretations are:

- (i)  $S_2: S_1$  claims of  $t_1$  that  $S_0$  claims of  $it_1$  that  $\Phi(it_1)$ .
- (ii)  $S_2: S_1$  claims of  $t_1$  that  $S_0$  claims that  $\Phi(it_1)$ .

These examples are representative of the new sorts of structural anaphoric connections across ascriptional boundaries that become possible with iteration. How do the attitudes involved in these complex ascriptions unpack according to the rules for the regimentation suggested in Section IV?

Suppressing type-tokening niceties, one can begin reading (i) by stripping off the outermost *de re* occurrences:

- (i')  $S_2: \text{For some term } x_2, S_1 \text{ claims that } [S_0 \text{ claims of } x_2 \text{ that } \Phi(it_2)], \text{ and } t_1 = x_2$ .

The second-level *de re* ascription by  $S_2$  is interpreted in terms of a nonascriptional identity (that is, symmetric substitution-inferential) commitment undertaken by  $S_2$  and a second-level *de dicto* ascription by  $S_2$ . The undertaking of a *de dicto* ascriptional commitment (of whatever level) is itself readily interpreted in terms of the attributing it expresses. Bracketing, in the interests of simplicity, elaborations required to deal with indexicals, and foreign languages (details that the discussion of Section IV shows how to reintroduce as needed), substitution instances of the first clause of (i') are interpreted by

(i'')  $S_1: S_0$  claims of  $t_2$  that  $\Phi(it_2)$ ,

that is, by attributions of first-level ascriptions, in this case, *de re* ones. And now the same procedures that were applied to turn (i) into (i') can be applied to (i''), followed in turn by the procedures that turned (i') into (i''). Those procedures suffice to interpret  $n^{\text{th}}$ -level ascriptions in terms of attitudes toward  $(n - 1)^{\text{th}}$ -level ascriptions. Repeatedly turning the crank on this machinery provides a recursive procedure for assigning a reading to each of the arbitrarily complex iterated ascriptions in the hierarchy.

The procedure is the one followed with the example just considered. First, strip off the terms occurring in the scope of the outermost *de re* operator, resulting in an undertaken identity commitment and an attributed *de dicto* ascriptional commitment at the same ascriptional level. Then trade that *de dicto* ascriptional commitment for an attributing at the next lower ascriptional level. The construction of the hierarchy of regimented iterated ascriptions proceeded by arbitrary repetitions of two sorts of formation rule, one corresponding to *de dicto* ascriptions, another to *de re*. So these interpretive rules match the ones used in constructing the expressions, and it follows that for any complex regimented ascription, a finite number of repetitions of the two interpretive steps will render the complex attitude in terms of simpler, ultimately nonascriptional ones. Recall that the general form of an  $m^{\text{th}}$ -level ascription is:

$S_{m+1}: S_m$  claims of  $\langle t'_1, t'_2, \dots, t'_k \rangle$  that  $\Phi(t_1, t_2, \dots, t_n)$ ,

where  $k$  is less than or equal to  $n$  and for all  $i$  less than or equal to  $k$ , there is a  $j$  less than or equal to  $n$  such that  $t'_i$  is an ascription-structural anaphoric dependent of  $t_j$  (symbolically  $\text{Depends}([t_j], [t'_i])$ ), and  $\Phi(t_1, t_2, \dots, t_n)$  is itself an ascriptional sentence of level  $m - 1$ . It is clear from this that it suffices to reduce the ascriptional complexity of the ascriptions involved, first, to trade *de re* ascriptions for *de dicto* ones plus undertaken identity commitments and, second, to trade undertaking an  $m^{\text{th}}$ -level *de dicto* ascription for attributing an  $(m - 1)^{\text{th}}$ -level ascription.

In this way every complex ascription is shown to correspond to a set of deontic attitudes. A converse condition holds as well. That is, starting off with any set of nonascriptional attitudes on the part of various interlocutors, it is possible to express them explicitly, from any desired point of view, by

means of iterated ascriptions, in a recursively complete fashion. The attitudes with which one starts can be any combination of undertaken and attributed identity commitments (involving singular terms) and assertional commitments (involving sentences in which those terms occur). What is being claimed is the *expressive completeness* of the regimented ascriptional idiom, over a certain domain. For consider: given a grasp of the background entailments, any interlocutor can attribute any nonascriptional assertional commitment to any other *de dicto*, with the ascriber consulting only his or her other attributions to that individual. Furthermore, given one's own undertakings of commitment (particularly to identities), one can offer *de re* versions of those ascriptions, and so indicate what, according to the ascriber, the one to whom the commitments are ascribed is talking *about*. Thus all of the first-level attitudes—that is, perspectives on the states of interlocutors—can be expressed with assertional explicitness by the use of ascriptional locutions. The undertakings of assertional commitment to these first-level ascriptional claims, however, expand the community's stock of states beyond what was present before ascriptional locutions are introduced. So these states must be explicitly ascribable in their turn, if ascriptional locutions are in fact to make possible the explicit expression of all the deontic attitudes. Applying the ascriptional expressive machinery one more time permits this, yielding second-level *de dicto* ascriptions of all of the new states generated by the first application of the machinery, and then in turn *de re* ascriptions of all of them. Repeating these two procedures inductively then permits the expression by any interlocutor of the contents of ascriptional claims of arbitrary complexity, from either the point of view of the ascriber (*de re*, inferentially expanded by commitments *undertaken* by the ascriber) or the point of view of the ascribee (*de dicto*, inferentially expanded by commitments *attributed* by the ascriber).

# Conclusion

---

The meaning of words is to be determined by their use.

ISAAC NEWTON, *Principia*<sup>1</sup>

---

## I. TWO CONCEPTS OF CONCEPTS

### 1. *Three Kantian Dualisms*

The semantic core of the account of discursive practice presented here is the theory of conceptual content it incorporates. The distinctive features of that theory emerge most clearly when it is contrasted with more traditional ways of thinking about concepts. The most familiar conception, one that is pervasive in contemporary philosophical thought, traces its ancestry back to Kant. Its debt to Kant is most evident in its essentially *dualistic* character: the ways in which the conceptual is contrasted with the nonconceptual. It is in this regard that traditional views of concepts differ most strikingly from the nondualist alternative endorsed here.

Kant's account begins by elaborating two of his epoch-making insights: first, that *judgments* are the fundamental form of awareness, so that concept use must be understood in terms of the contribution it makes to judging; second, that cognition and action are distinguished from their analogs in nonrational beings by their liability to certain sorts of *normative* assessment (see Chapter 1). Kant combined these insights with a *classificatory* theory of concepts, in terms of which he synthesized the teachings of his rationalist and empiricist predecessors. It is this aspect of his account that has been most influential in subsequent thought—becoming so much a matter of

course as to be almost invisible as a presupposition. Although it is based on important dimensions of ordinary concept use, the classificatory conception generalizes inappropriately as a result of running together substantially different phenomena.

For Kant, concepts provide only one of the two elements required for judgment. Concepts without intuitions are empty, and intuitions without concepts are blind.<sup>2</sup> Kant's theory is essentially dualistic in that the notion of the conceptual element in judgments is that of one of a *pair* of contrasting aspects. That neither is intelligible apart from its collaboration with the other is one reflection of his healthy emphasis on the primacy of judgment. It remains unclear, however, how much remains of the picture of judgment as the joint product of two distinct faculties if those faculties can be understood only by abstraction, that is, in terms of their contribution to the activity of judging. Insofar as Kant's embrace of both intellectual and sensible faculties is construed as his saying "You're both right" to Leibniz and Locke, his insistence on their mutual presupposition is bound to look like the bit where he takes it back. On the other hand, insofar as sense can be made of the notion of distinct contributions to judgment made by concepts and the unconceptualized given, the nature of their collaboration seems bound to remain mysterious. What sort of 'fit' is envisaged between concepts and intuitions, in virtue of which it is *correct* (or just possible?) to apply some, but not other, concepts to the manifold of (preconscious) representations with which intuition in some sense presents the understanding? How does intuition *constrain* the application of concepts? Kant's appeal to the schematization of the concepts just moves the bump in the rug around. The capacity for judgment, for applying rules to particular instances, subsuming intuitions under concepts, is something that in the end<sup>3</sup> we must just accept *that* we have, without understanding just *what* we have. A distinction becomes a dualism when its components are distinguished in terms that makes their characteristic relations to one another ultimately unintelligible. (Descartes's dualism is, as always, the paradigm.)

Essential elements of Kant's dualistic conception of concepts are still with us today. They are the basis for the suspicion evinced by some (for instance Davidson) that talk of concepts inevitably commits us to a picture in which they play the role of *epistemological intermediaries*, which stand between us and the world we conceptualize and forever bring into question the very possibility of genuine cognitive access to what lies beyond them.<sup>4</sup> To see why such suspicions are justified, and to bring out the contrast between dualistic and nondualistic conceptions of the conceptual, it is helpful to disentangle three different sorts of contrast between the conceptual element in thought and some nonconceptual element in thought, all of which are in play in Kant's usage. Each of these contrasts represents a genuine distinction, but as these distinctions are elaborated and run together in Kant's classificatory model of concepts, each becomes a dimension of an unworkable dualism. For

Kant, concepts contrast with intuitions first as *form* to *matter*, which they structure or organize. Second, they contrast with intuitions as *general* to *particular*. Finally, they contrast with intuitions as products of *spontaneity* or intellectual activity, as opposed to products of *receptivity*.<sup>5</sup>

In the first, the conceptual is distinguished from the *material*, that which provides *content*, as opposed to the form (more specifically the *normative* form or rulishness), which is the contribution of concepts. In the second, the conceptual is distinguished from the particular, as what *classifies* to what is *classified*. In the third, the conceptual is distinguished from what is imposed on us from without, as what *we do* as opposed to what is *done to* or imposed on us. It is the beginning of wisdom in reading the first Critique to distinguish the roles played in various arguments by these different distinctions. It is central to Kant's account that the three contrasts (though different) all line up together. Once this is questioned, a host of alternatives to his arguments present themselves. The lines of thought developed in this work support the conclusions that:

1. there are genuine distinctions underlying the contrasts Kant points to, but
2. far from coinciding, they are each independent of and orthogonal to the others, and
3. none of them is properly understood as distinguishing the conceptual from some nonconceptual element in judgment.

What a judgment expresses or makes explicit, its content, is conceptual all the way down.

The first idea is that of concepts as organizing something else. This can take many forms. It can be claimed that what is organized is *experience*,<sup>6</sup> which is carved up by concepts, or alternatively lumped together by them. The material on which concepts work can be conceived of as perceptions or observations, sense data, or patterns of sensory stimulation. The concepts are supposed to be the source of *structure*, while something else provides the *content* or matter. Davidson has this picture in mind when he objects to the "scheme/content" dualism that he takes to be implicit in talk of alternative conceptual schemes.<sup>7</sup> Concepts contrast with the unconceptualized matter that they conceptualize, which thereby provides content to the judgments that result. The worry inevitably raised by this picture is that unless its activity is entirely unnecessary, in conceptualizing the unconceptualized the understanding that is deploying the concepts must somehow *alter* what it works on and is therefore liable to the possibility of systematically *falsifying* that matter in rendering it digestible to the intellect.<sup>8</sup> It should be admitted that it always remains pretty obscure what can be meant by either the form or the matter side of this opposition. (C. I. Lewis's heroic expository effort in *Mind and the World Order* is probably as clear a setting-out of this way of conceptualizing intuitions about concepts and intuitions as can be had.)

The second idea is that concepts are something *general*, something best expressed by the use of *predicates*. Along this dimension they contrast with nonconceptual *particularity*, as expressed by the use of (at least some kinds of) singular terms. The idea here is that predicates *classify* things, *say* something about them, as opposed to simply picking them out. This thought is the heir of Kant's treatment of intuitions as representations of particularity. The association of concepts with general terms rather than proper names is pervasive in the tradition. One important example is Frege,<sup>9</sup> for whom concepts are functions from (sequences of) singular terms to truth-values, and so essentially things that can be *true of* the objects picked out by singular terms, by contrast to those *objects*, which concepts can be *true of*.

Finally, the third idea is that the *conceptual* order contrasts with something like the *causal* order, which constrains it. This distinction is the heir of Kant's distinction between judgments as the joint products of the activity ('spontaneity') of the intellect and the receptivity of the senses. According to this line of thought, whatever is conceptually articulated shows the effects of the mind working on it, whereas the nonmental world that thought is largely about is not in itself conceptually articulated. Because of special features of Kant's view, he could not put this contrast in terms of concepts versus causes (since talk of causation is for him already talk that betrays traces of the activity of the concept-mongering intellect).<sup>10</sup> Nonetheless the tradition he inspired contrasts conceptually articulated expressions such as definite descriptions with those that are taken merely to register causal impingements—above all the uses of *demonstratives* that are so important in expressing the noninferential reports in virtue of which our concepts have empirical content. Kaplan's work is a prime example of contemporary versions of this distinction, as he worries about how to characterize the relation between the conceptual element in propositional contents, expressed by the use of predicates and descriptions, and the *nonconceptual, contextual, or causal* element, expressed by the use of indexicals.

In this contemporary form, Kant's distinction survives as the contrast between the *unrepeatable* character of indexical tokenings, reflecting their token-reflexive embeddedness in a causal context, and the *repeatable* concepts that are epitomized by definite descriptions.<sup>11</sup> Kant ran the second and third dualistic thoughts together—that is, the distinction between predicates and singular terms, on the one hand, and between repeatable and unrepeatable elements of thought, on the other, by systematically failing to distinguish between *representations of particularity* and *particular representations* (though elsewhere he is clear-headed about the distinction between representations of relations and relations among representations). In fact, however, singular terms, which represent particulars, are typically themselves as repeatable as predicates, while unrepeatable or token-reflexive indexical expressions can be of either grammatical category.

According to this broadly Kantian, dualistic, classificatory conception of

concepts, they function as epistemological intermediaries. They stand between the understanding mind and a world that is the source of their content or matter—a world composed of particulars that are grasped by means of general concepts and that imposes itself causally on a mind obliged somehow to conform itself to those causal impingements. As long as the conceptual is conceived in this way, Davidson is quite right to object to talk of conceptual schemes by means of which we render the world intelligible to and digestible by thought. But one of the lessons that ought to be drawn from the stories told here that this is not the only way to think about concepts. In particular, this broadly Kantian approach can be laid alongside another, inspired by Sellars, which avoids the dualistic understanding of each of the three distinctions that is characteristic of the Kantian one.

## *2. The Inferential Conception of Concepts Is Not Dualistic in Any of the These Ways*

The approach developed here thinks of concepts to begin with as inferential roles (see Chapter 2). It treats a reliably differentially elicited response as *conceptually* classifying the stimulus to which it is keyed just in case that response occupies a position in an *inferentially* articulated space of claims that can be offered as, and stand in need of, reasons. In order for it to count as a conceptually contentful performance, that response must be able to serve as a premise for inferences to the applicability of further concepts. The particular content of a given concept is accordingly thought of as the content of an inferential commitment: roughly the commitment to the propriety of the inference from any of the appropriate *circumstances of application* of that concept to any of the appropriate *consequences of application* of the concept.<sup>12</sup> In this way even the *empirical* content some concepts have in virtue of their connection with *noninferential* circumstances of application in *perception*, and the *practical* content some concepts have in virtue of their connection with noninferential consequences of application in *action*, can be seen to be inferentially articulated.

It is essential to this inferential approach to concepts that the inferences in question are what Sellars calls *material* inferences. This is to say that their correctness involves the particular *contents* of the concepts invoked by their premises and conclusions; it is not underwritten purely by the *form* of those premises and conclusions. A paradigm is the inference from “*A* is to the East of *B*” to “*B* is to the West of *A*,” whose correctness expresses part of the *content* of the concepts *East* and *West*.

The first point to notice, then, is that thinking of conceptual contents as articulated by the material inferences that determine their role in giving and asking for reasons involves no contrast between concepts as *form* and something else as *matter* or content. The inferential role, which is the conceptual role, is the content. If one likes, one can say that on this conception the form

of that content is inferential. But the concept itself is identified with the particular constellation of material-inferential transitions the concept is involved in. This is not a structuring of something *else* that contrasts with the concept. The inferences materially relate one concept to other concepts, not to something of another kind. Thus the first of the Kantian dualisms, contrasting the conceptual and the material, is simply not involved in the inferential conception of nonlogical concepts.

It is possible, however, to go on to erect a superstructure of formal proprieties of inference on this base of material proprieties of inference (see 2.4.2 above). This sort of inferential articulation is an essential part of the expressive role of specifically *logical* vocabulary, by means of which we make explicit to ourselves the contents of our nonlogical concepts. So a distinction between matter and form is discerned and exploited by the inferential approach, though not in a form recognizable as distinguishing a conceptual from a nonconceptual element in judgment. Indeed, the same Fregean procedure of noting invariants under substitution that gives rise to the notion of formal proprieties of inference—an inference being valid in virtue of its form with respect to some distinguished vocabulary-kind *K* (paradigmatically logical vocabulary) just in case it is a materially good inference and cannot be turned into one that is not good by substituting non-*K* for non-*K* vocabulary—is what makes it possible to distinguish various formal categories of subsentential expressions, such as singular terms and predicates.

Only claims can literally function as premises and conclusions in inference; so only what is expressed by *sentences* can directly have an inferential role and so be in the most basic sense inferentially articulated. This is the version of Kant's insight concerning the primacy of judgments in cognition that survives into the inferential conception of concepts. But subsentential expressions can nonetheless be conceptually articulated according to that conception—their occurrence in a sentence can have an *indirect* inferential significance. For substitution of one subsentential expression for another in a sentence can either result in preserving the goodness of inferences in which the sentence is involved or fail to preserve it. In this way subsentential expressions can be sorted into indirect inferential equivalence classes, by noting direct inferential invariances of the sentences that result from their substitution one for another. Thus the *inferential* approach to the conceptual articulation of sentences can be extended *substitutionally* to include the conceptual articulation of subsentential expressions.

When this is done, the subsentential categories of singular terms and predicates can be distinguished by the different patterns of *substitution inferences* in which they are involved. In particular, singular terms are distinguished by the de jure *symmetric* significance that their occurrence in a sentence has for substitution inferences involving it. For example, if the inference from "Benjamin Franklin spoke French" to "The popularizer of lightning rods spoke French" is a good one, then so is the converse inference.

By contrast, all predicates are involved in some *asymmetric* substitution inferences. For instance, the inference from "Benjamin Franklin could dance" to "Benjamin Franklin could move," is a good one, but the converse inference need not be. On the basis of such differences in substitution-inferential behavior, the difference between the sort of conceptual role played by singular terms and that played by predicates can be characterized (see Chapter 6).

This means that the second of the Kantian dualisms, though based on a genuine categorial distinction, also does not set off concepts as conceived by the material-inferential model. There is no restriction of the conceptual to the general, as expressed by predicates, in contrast to the particularity invoked by singular terms. Singular terms have an inferential role, represented by the set of terms intersubstitutable for them, just as predicates do. The difference between them is a formal difference of symmetric versus asymmetric substitution inference. It is not a difference that involves contrasting the conceptual as embodied in predicates, which express generalities, with something else, embodied in singular terms, which express particularity. Singular terms and predicates, the particular and the general aspects of claims, are equally (though not identically) inferentially articulated, and so equally conceptually contentful. *Particularity is as much a conceptual matter as generality*, on this conception. Thus the second dualism gets no grip on the inferential rendering of conceptual contentfulness, once that account has been extended to the subsentential level by invoking the notion of substitution inferences.

The third of the Kantian dualisms contrasts the conceptual, as the product of cognitive activity, with the nonconceptual impingement on cognitive receptivity in virtue of which that cognitive activity is constrained. Outside the strictures of Kant's own system, we can think of this as the conceptual/causal contrast, in which the application of concepts is constrained by the causal order, thought of as not itself conceptually articulated. The point of contact between the conceptual order and the causal order, according to this conception, takes place in *deixis*, where something is indicated without being characterized. In grasping this conception it is helpful to focus on the use of deictic expressions in noninferential reports, such as "This is red." For it is in such reports that the world most directly imposes itself on suitably trained concept-mongers, who find themselves passively acknowledging empirically contentful commitments.

Once again, it ought not to be denied that this sort of receptivity is essential to our empirical knowledge and that it ought to be distinguished from other, more spontaneous applications of concepts, for instance in purely inferential theorizing. Yet according to the inferentialist conception, unrepeatable deictic tokenings—for instance particular uses of 'this'—are fully conceptually articulated. Indeed, were they not, they could serve no cognitive purpose. To see how occurrences of deictic tokenings are assigned an inferential significance, and so taken to be conceptually contentful, is accordingly

to see that the third of the broadly Kantian dualisms—contrasting conceptual constraint with causal constraint on the application of concepts—fails to get a grip on the inferential conception of the conceptual. Just as the idea of *inference* needed to be supplemented by that of *substitution* in order to be brought to bear on subsentential expressions, so the idea of substitution-inferential significance needs to be supplemented by that of *anaphora* in order to be brought to bear on *unrepeatable* tokenings of subsentential expressions, rather than just on their repeatable types. To take one expression to be anaphorically dependent on another is to take it as *inheriting* its substitution-inferential role from the tokening that is its anaphoric antecedent.<sup>13</sup> If you say, “*That* is a porcupine” and I pick up that premise and conclude, “[so] *it* is a vertebrate,” the truth of the conclusion I have drawn is to be settled (according to an interpreter) by what substitutions are appropriate (according to the interpreter) for the demonstrative tokening that serves as the antecedent for my anaphorically dependent tokening. If (according to the interpreter) what the first speaker referred to by ‘*that*’<sup>14</sup> is the most cunning wooden replica of a porcupine in the room, then since this identity claim is to be understood as an intersubstitution license, I have unwittingly claimed of a cunning wooden replica of a porcupine that it is a vertebrate, and what I said is false.

Anaphora permits the formation of *chains* of tokenings, anchored by antecedents that can be deictic and therefore strictly unrepeatable. These *chains of unrepeatables are themselves repeatables and play the same role in substitution inferences that sets of cotypical tokenings play for repeatable expressions such as proper names and definite descriptions*. It is by means of anaphora, then, that substitution-inferential potential can be inherited by one expression from an unrepeatable tokening. In virtue of this mechanism, unrepeatable tokenings such as uses of demonstratives become available for service as premises in inference. In this way they acquire an *inferential* significance and so can be understood as expressing a *conceptual* content. This function of anaphora is essential to the existence of deictic expressions. For without the capacity to be picked up anaphorically, and so to have some inferential significance, deictic expressions would just be noises wrought from us by exposure to things—rather than genuinely linguistic expressions that can be used to *say* something. Thus anaphora is more basic than deixis, for there can be languages that have anaphoric mechanisms but no deictic ones, while there cannot in principle be languages with deictic mechanisms but no anaphoric ones (see Chapter 7).

In any case, with anaphora available to bring deictic expressions into substitution inferences, such expressions have indirect inferential roles, and so conceptual contents. There is no contrast between expressions like definite descriptions and those like demonstratives over the issue of whether or not they are inferentially articulated and so conceptually contentful. The structure of their contents is specifically different, for the latter are involved

in substitution inferences via anaphoric chains of unrepeatable tokenings potentially of a variety of types, while the former<sup>15</sup> are involved in substitution inferences via sets of repeatable, because cotypical, tokenings. But this difference plays a role analogous to that between symmetric and asymmetric substitution-inferential significances in distinguishing singular terms from predicates. In neither case is a contrast underwritten between the conceptual and something else, whether particular or causally responsive. Deictic tokenings play a role in the causal order, but they are not for that reason not also conceptually articulated.

Thus *none* of the Kantian dualisms—which contrast the conceptual as formal with the material, the conceptual as general with the particular, and the conceptual as spontaneous activity with the constraint of causes—applies to the inferential conception of concepts. That conception does not involve any commitment to a dualism of conceptual scheme and something else that it structures, classifies, or is about. So Davidson's proper fastidiousness about scheme/content dualisms and epistemological intermediaries ought not to motivate a rejection of appeals to concepts as here conceived. Concepts conceived as inferential roles of expressions do not serve as epistemological intermediaries, standing between us and what is conceptualized by them. This is not because there is no causal order consisting of particulars, interaction with which supplies the material for thought. It is rather because all of these elements are themselves conceived as thoroughly conceptual, not as contrasting with the conceptual.

The conception of concepts as inferentially articulated permits a picture of thought and of the world that thought is about as *equally*, and in the favored cases *identically*, conceptually articulated. *Facts* are just true claims.<sup>16</sup> Facts, like other claims, are conceptually articulated by their inferential and incompatibility relations to other claims. It is a feature of the conceptual articulation of claims, and hence of facts, that they are about particular objects.<sup>17</sup> (Indeed, the fact that we are accustomed to saying that facts, like claims, are *about* objects, rather than that they somehow *consist* of objects, is evidence for the correctness of identifying facts with true claims.) It is these facts and the propertied and related objects they involve that are cited as stimuli by interpreters who are specifying the reliable differential responsive dispositions in which the contents of empirical contents originate. These noninferential dispositions (the locus of our empirical receptivity) accordingly do not constitute the interface between what is conceptually articulated and what is not, but merely one of the necessary conditions for a conceptually articulated grasp of a conceptually articulated world—the world consisting of everything that is the case, all the facts, and the objects they are about.

In this way a story has been told about how the three nonconceptual poles of Kant's tripartite division of the conceptual and the nonconceptual contributions to the contents of judgments ought to be incorporated within the

conceptual realm. An approach to concepts that moves beyond exclusive focus on classification to include inferential connections among concepts as essential to their identity and individuation:

1. incorporates *content* by employing a notion of *material* proprieties of *inference*,
2. incorporates *particularity* by distinguishing between the symmetric role of singular terms in *substitution* inferences and the asymmetric role of predicates in substitution inferences, and
3. incorporates the deictic *unrepeatability* by which causal context affects conceptual content by explaining how *anaphoric* chains initiated by unrepeatable tokenings function as type-repeatables in substitution inferences.

The key theoretical concepts used to characterize the articulation of conceptual roles are material inference, substitution, and *anaphora*, so this can be called the ISA approach to semantics.

## II. NORMS AND PRACTICES

### 1. *The Normative and the Factual*

This inferential semantics is embedded in a normative pragmatics. Material proprieties of inference are understood as norms implicit in social practices that qualify as *discursive* inasmuch as they involve treating some performances as having the significance of assertions. Such inferentially articulated practices confer propositional contents on statuses, attitudes, and performances that are suitably caught up in them (since for an expression to have a certain conceptual content just is for its use to be governed by a corresponding set of norms). In this way the semantically primitive notion of material proprieties of inference is explained in the pragmatics—in the account of linguistic practice.

Such a pragmatic theory of the relations between meaning and use raises issues about the status of implicit practical norms. Does not talk of deontic statuses as instituted by social practices involve a residual dualism? When the orienting commitment to the normative character of discursive practice was first introduced and motivated, in Chapter 1, this insight of Kant's was presented in the context of a shift from a broadly Cartesian dualism of the mental and the physical to a broadly Kantian dualism of the normative and the factual. In these crude initial terms, Descartes's opposition of two kinds of descriptive properties (corresponding to ontological kinds of substances) was contrasted with a deeper opposition between descriptive and prescriptive attitudes—between attributing properties and attributing proprieties. Thus even if Kant's *semantic* dualisms have been overcome by the ISA approach, it would appear that the pragmatics in which that semantics is embedded

incorporates an overarching dualism that distinguishes the normative and the nonnormative. How should the relations between these categories be understood?

The deontic scorekeeping idiom acknowledges a distinction between normative and nonnormative claims, explained in terms of their different roles in practical reasoning, but that *distinction* does not underwrite a *dualism* of norm and fact. Indeed, looked at more carefully, neither does Kant's. (He is large; he contains multitudes.) The initial discussion of replacing one dualism with another (in Chapter 1, Section II) was only a temporary expository device, discarded in favor of a more nuanced treatment (in Section IV) once its purpose was served. For Kant, rules are the form of the normative as such. To call something 'necessary' is to say that it happens according to a rule, and everything that happens in nature, no less than everything done by humans, is subject to necessity in this sense. Concepts are rules, and concepts express natural necessity as well as moral necessity. So according to him there is strictly no nonnormative realm—no realm where concepts do not apply. Kant's fundamental innovation is best understood to consist in his employment of a normative metalanguage in specifying *both* what merely happens and what is done.

Of course he does distinguish between the realm of regularity and the realm of responsibility. This is the distinction between that to which concepts apply and those who apply concepts—between that which can acknowledge rules only implicitly by obedience (by having concepts be applicable to it) and those who can acknowledge them explicitly by the use of concepts (by applying concepts). It is only rules as explicitly acknowledged that can be both binding and *disobeyed*,<sup>18</sup> and it is the capacity for such acknowledgment—acting not just according to rules but according to *conceptions* of rules—that institutes distinctively normative statuses such as duty and responsibility. The applicable distinction is accordingly not between the normative and the nonnormative but between what can adopt explicitly normative *attitudes* and what cannot. Only we discursive (that is concept-mongering) creatures can *take* ourselves and others to be bound by the norms that are our concepts.<sup>19</sup>

This is the idea that is followed out in the deontic scorekeeping pragmatics presented here. The idiom in which the account of discursive commitment is expressed is normative throughout. Propositional contents are understood in terms of their explanatory role in specifying proprieties of claiming, judging, and inferring—in general, in terms of the role they play in the game of giving and asking for reasons. What it is for something to state or express a fact is explained in normative terms, and what it is for something to be stated or expressed is explained in turn by appeal to that practice. So what it is to be a fact—that is, true claim—is explained in normative terms. It is explained phenomenally, by appeal to the practice of fact-stating, which comprises the practical attitudes of taking a performance to be the

stating of a fact and purporting to state a fact by producing a performance. In this order of explanation, normative notions such as commitment and entitlement—which articulate implicit proprieties of practice—are more fundamental than the nonnormative properties they enable discursive practitioners to express explicitly.

However, only some of the vocabulary on which conceptual content is conferred by implicitly normative discursive practice plays the expressive role of making explicit specifically normative attitudes—for instance the attribution or acknowledgment of commitments. As explained in Chapter 4, the distinctive function of normative vocabulary is to express endorsement of patterns of practical reasoning—that is, in the first-person case, reasoning that leads from doxastic to practical commitments (presystematically: from beliefs to intentions). Social practices are implicitly normative in a way that mere behavioral regularities are not. Put phenomenally, that is to say that what a scorekeeper or interpreter has attributed counts as a *practice* in this sense (rather than a behavioral regularity or disposition) only if it is specified in explicitly normative terms—in terms of what, according to the practice, it is *correct*, or *proper* to do, what one *ought* to do, what one becomes *committed* or *entitled* to by a certain sort of performance, and so on. The account of practical reasoning explains in deontic scorekeeping terms how words have to be used in order to mean what such terms as ‘*correct*’, ‘*ought*’, and ‘*committed*’ do. By doing that, it makes sense of the distinction between normative statuses and attitudes, on the one hand, and nonnormative states and dispositions, on the other.

Explicitly normative vocabulary can be used to make claims (for example “Bank employees are obliged to wear neckties,” “One ought not to torture helpless strangers”). Those claims can be taken-true, can be put forward as, or purport to be, true. Since facts are just true claims (in the sense of what is *claimed*, not the *claiming* of it), this means that norm-explicitating vocabulary is in the fact-stating line of business. That is, corresponding to the distinction between normative and nonnormative vocabulary is a distinction between normative and nonnormative facts. (Indeed, this ontologically relaxed approach to facts finds nothing mysterious about negative, conditional, or modal facts, facts about the self-identity of objects, or in general facts expressed by any sort of declarative sentence at all.) In this way the normative is picked out as a subregion of the factual.

To revert to the previous point, however, this is a distinction made *within* the encompassing normative metalanguage in which the deontic scorekeeping roles characteristic of normative and nonnormative vocabulary are specified. The distinction between normative and nonnormative vocabulary, claims, and facts is itself drawn in normative terms. In this sense, the story is one in which it is norms all the way down—a Kantian story (on the pragmatic, rather than the semantic side).<sup>20</sup> Far from opposing one another, the realms of fact and norm mutually include one another: fact-stating talk

is explained in normative terms, and normative facts emerge as one kind of fact among others. The common deontic scorekeeping vocabulary in which both are specified and explained ensures that the distinction between normative and nonnormative facts neither evanesces nor threatens to assume the proportions of an ultimately unintelligible dualism.

## 2. *Where Do Norms Come From?*

The story told here is Kantian not only in that it is told in normative terms but also in the pride of place it gives to normative *attitudes* in explaining how we are both distinguished from and related to the non-us that surrounds us. On the one hand, such practical attitudes—taking or treating a performance as correct, attributing or acknowledging a commitment—have been appealed to in explaining our relations in perception and action to the causal order of nonnormative facts that we inhabit cognitively and practically. On the other hand, they have been appealed to in explaining where discursive norms come from—how sapience could have arisen out of the primordial nondiscursive ooze of mere sentience. For it has been claimed not just that we discursive beings are creatures of norms but also that norms are in some sense creatures of ours—specifically, that discursive deontic statuses are *instituted* by the practices that govern scorekeeping with deontic attitudes.

Norms (in the sense of normative statuses) are not objects in the causal order. Natural science, eschewing categories of social practice, will never run across *commitments* in its cataloguing of the furniture of the world; they are not by themselves causally efficacious—any more than strikes or outs are in baseball. Nonetheless, according to the account presented here, there are norms, and their existence is neither supernatural nor mysterious.<sup>21</sup> Normative statuses are domesticated by being understood in terms of normative attitudes, which *are* in the causal order. What is causally efficacious is our practically taking or treating ourselves and each other as having commitments (acknowledging and attributing commitments)—just as what is causally efficacious is umpires and players dealing with each other in a way that can be described as taking the score to include so many strikes and outs.<sup>22</sup>

It must then be asked how such an apparently reductive story about norms as instituted by social practices can be understood to be compatible with an insistence on the *irreducibly* normative character of the metalanguage in which norm-instituting social practices are specified. Here is the short answer: The work done by talk of deontic statuses cannot be done by talk of deontic attitudes *actually* adopted or relinquished, nor of *regularities* exhibited by such adopting and relinquishing, nor of *dispositions* to adopt and relinquish such attitudes. Talk of deontic statuses can in general be traded in only for talk of *proprieties* governing the adoption and alteration of deontic attitudes—proprieties implicit in social scorekeeping practices.

The crucial inferential articulation of discursive commitments consists in part in the fact that unacknowledged commitments can be (taken by other scorekeepers to be) undertaken consequentially, by acknowledging commitments to claims that (according to those scorekeepers) entail them. So according to the attributions (normative attitudes) of another, my commitments (normative statuses) outrun those I acknowledge (normative attitudes). In this way the social articulation of deontic scorekeeping attitudes is essential to the inferential (and hence discursive) articulation of the contents of the commitments they address. But this social articulation of scorekeeping practice is essentially normative in force. That I acknowledge commitment to *p* does not (according to the scorekeeper) mean that I *do* or *will* acknowledge commitment to its consequence *q*, only that I *ought* to—that I am, whether I realize it or not, *committed* to *q*.

It was shown at the end of the last chapter that the contents of ordinary empirical claims—objective proprieties governing the application of concepts—are not equivalent to the contents of any claims about who is committed to what. The implicit scorekeeping attitudes expressed by this difference in explicit contents accordingly distinguish what follows from *p* from what I or anyone *takes* to follow from *p*. What follows from *p* cannot be identified with how I or anyone *actually* keeps score; it is rather to be identified with a feature of *correct* scorekeeping (for it depends on what else is true, not on what anyone *takes* to be true). Conceptual contents on this inferential conception—and so what interlocutors are really committed to by using particular expressions (performing particular speech acts)—codify *proprieties* of scorekeeping. Any scorekeeper who attributes a conceptually contentful commitment may get these wrong, just as anyone who acknowledges or otherwise acquires such a commitment may get them wrong. Talk of inferentially articulated contents is a way of talking about implicit norms governing deontic scorekeeping practice;<sup>23</sup> this is the cash value of the claim that conceptual contents are *conferred* by such practice. But since commitments must be individuated at least as finely as their contents, if those contents are determined only by how it is *correct* to acquire and alter deontic attitudes, the commitments themselves must be understood as instituted also by proprieties of scorekeeping, rather than by actual scorekeeping. The scorekeeping account incorporates a phenomenalist approach to norms, but it is a *normative* phenomenism, explaining having a certain normative status in effect as being *properly* taken to have it.

At this point it can easily look as though the account of normative statuses as instituted by social practices is marching around in an unproductive circle (at best, unilluminating; at worst viciously circular and incoherent). For clearly the prior question arises once more: What is the relation between normative specifications of practices and nonnormative specifications of behavior? Actual scorekeeping, the adoption and alteration of practical normative attitudes (acknowledgments and attributions of deontic statuses),

consists of causally efficacious events and dispositions. If normative statuses could be understood as instituted by actual attitudes of acknowledging and attributing them, then the use of normative vocabulary specifying proprieties, commitments, and entitlements would straightforwardly supervene on the use of nonnormative vocabulary specifying performances and performative dispositions and regularities. If, however, as has been claimed, the institution of discursive deontic statuses should be understood rather in terms of the implicit practical *proprieties* governing such scorekeeping—not how the score is actually kept, but how according to the implicitly normative scorekeeping practices it *ought* to be kept, how scorekeepers are *obliged* or *committed* to adopt and alter their deontic attitudes rather than how they *actually* do—then the source and status of these norm-instituting proprieties of scorekeeping practice must be inquired into.

### 3. Interpretation

Proprieties are normative statuses—the status a performance has as *correct* or *incorrect* according to a rule or practice. This is so even when the practice whose proprieties are in question is itself a deontic scorekeeping practice. In that case what is being evaluated as proper or improper is the acquisition and alteration of deontic attitudes—that is the acknowledgment and attribution of further deontic statuses (commitments and entitlements). The (normative) phenomenalist strategy that has been pursued throughout is to understand normative statuses in terms of normative attitudes—in terms of (proprieties of) *taking* to be correct or incorrect. This strategy dictates two questions concerning proprieties of scorekeeping practice. First (apropos of phenomenalist about norms), what must one be doing in order to count as taking a community to be engaging in implicitly normative social practices—in particular in deontic-status-instituting, conceptual-content-conferring discursive scorekeeping practices? Second (apropos of its being a normative phenomenalist), what is it about the actual performances, dispositions, and regularities exhibited by an interacting group of sentient creatures that makes it correct or appropriate to adopt that attitude—to interpret their behavior by attributing those implicitly normative discursive practices?

The first question can be addressed by considering the different sorts of intentional stance that interpreters can adopt, according to the story told here. The central task of the pragmatic part of this project (the account of discursive practice) has been to introduce the model of *deontic scorekeeping*. Keeping discursive deontic score by attributing inferentially articulated deontic statuses—propositionally contentful commitments and entitlements to those commitments—is treating the one so interpreted as being in the game of giving and asking for reasons. Social practices are *linguistic* practices when interlocutors take up the *discursive scorekeeping stance* toward one another. Adopting this stance is (implicitly, or in practice) taking or treating others as

producers and consumers of propositionally contentful speech acts. Performances count as propositionally contentful in virtue of their relation to a core class of speech acts that have the pragmatic significance of *claims* or *assertions*.

Assigning this sort of significance to performances is treating them as making *explicit* the adoption of a normative status—that is, acknowledging (undertaking) a doxastic commitment by *saying* what one is committed to. Keeping discursive score on others is adopting deontic attitudes—that is, attributing discursive commitments by implicitly or in practice taking or treating another *as* committed. Such scorekeeping (and so linguistic practice generally) does not require that one be able *explicitly* to *attribute* deontic statuses—to say (assert) that someone is committed to the claim that *p*. The logical locutions whose expressive role is to make the adoption of such pragmatic attitudes explicit in the form of claimable contents—propositional-attitude-ascribing vocabulary such as the regimented “. . . is committed to the claim that . . .” or its vernacular correlate “. . . believes that . . .”—form an optional superstratum whose expressive role can be understood in terms of what is implicit in ground-level linguistic practice, but which is not required for, or presupposed by, such practice.

The production and consumption of speech acts of which participants in these fundamental discursive practices are capable accordingly differ as to whether the adoption of deontic attitudes (toward normative statuses) they involve is explicit or implicit. They can *explicitly acknowledge* (and so undertake) discursive commitments, in their assertional performances, but only *implicitly attribute* them, in their scorekeeping practice. Since acknowledging a commitment (the basic sort of undertaking or acquisition of that deontic status) is producing (or being disposed to produce) performances whose pragmatic significance is to make it appropriate for scorekeepers to attribute that commitment, to take someone to be a producer of speech acts is implicitly to take that practitioner to be also a consumer of them—a scorekeeper. Givers of reasons must be able to understand what it is to give a reason. As Davidson says: “One cannot be a thinker unless one is an interpreter of the speech of others.”<sup>24</sup>

Although performances cannot be accorded the significance of speech acts without implicitly treating the performer as a discursive scorekeeper, it is possible for those who are discursive scorekeepers to attribute a derivative sort of propositionally contentful discursive status and attitude to nonlinguistic creatures. This is adopting the *simple* intentional stance of interpreting something as a simple or practical intentional system. When this stance is adopted, the interpreter keeps a simplified sort of deontic score, by attributing propositionally contentful commitments, both doxastic and practical, which the subject is taken to acknowledge implicitly in its behavior. Its performances, dispositions, and behavioral regularities can be made intelligible by attributing sample pieces of practical reasoning, in the way Dennett

has described so well. The scorekeeping involved is simplified in that adopting the simple intentional stance does not involve attributing speech acts; it does not involve even implicitly treating the system in question as itself able to keep score (attribute, not just acknowledge deontic statuses); hence it does not involve treating it as a participant in the essentially social and linguistic game of giving and asking for reasons.

Discursive scorekeepers, participants in full-blooded linguistic practices, do two sorts of things that such simple, nonlinguistic intentional systems cannot: institute deontic statuses and confer conceptual contents. On the pragmatic side, both social flavors of deontic attitude—acknowledging and attributing—are needed to institute deontic statuses; reference to practical grasp of the possibility of *attributing* them is required to make sense of what is *acknowledged* as being inferentially articulated *commitments*. In the case of simple intentional systems, that essential pragmatic ingredient is supplied only by the interpreter, rather than attributed to the one being interpreted. On the semantic side, the social-perspectival dimension of inferential articulation is required to make sense of what states, attitudes, and performances exhibit as genuinely *propositional*, which includes having objectively *representational* conceptual content (see Chapter 8). In the case of simple intentional systems, that essential semantic ingredient is supplied only by the interpreter, rather than being attributed to the one interpreted. So the intentionality attributed by adopting this sort of stance is doubly derivative. On the side of pragmatics, the socially and inferentially articulated norms are derivative from the scorekeeping practices of the interpreter. As a result, on the side of semantics, the propositional and other conceptual contents employed to measure and systematize its behavior cannot be funded out of that behavior itself.<sup>25</sup>

By contrast, if one attributes genuinely linguistic practices to a community—takes its members to adopt the discursive scorekeeping stance to one another, and so to accord some performances the significance of speech acts, in particular assertional ones—one thereby takes them to exhibit *original* intentionality. The social practices one interprets them as engaging in are sufficient by themselves to institute inferentially articulated deontic statuses and so to confer genuinely conceptual contents. Describing the model of inferentially articulated deontic scorekeeping social practices is specifying in detail what one must *take* the members of a community to be doing in order for it to be *talking*—giving and asking for *reasons*, *making* their words and performances mean something by *their taking* them to mean something<sup>26</sup>—that one is thereby taking them to be doing. In short, the model specifies what structure an interpretation of the activities of a community must have in order for it to count as attributing original intentionality to that community—taking it as instituting socially and inferentially articulated deontic statuses and so conferring genuinely propositional conceptual content on them. This is adopting a further sort of stance.

So the difference between derivative simple and original discursive intentionality is presented in terms of the difference between two stances or forms of interpretation—in terms of the difference between the attitudes adopted in *attributing* them. The difference between these sorts of intentionality is not that one is construed in methodologically phenomenalist terms and the other is not. In keeping with the stance stance, this account is phenomenalist about both. The difference is that what one attributes in the case of genuinely discursive intentionality is (taken to be) autonomous in a way that what one attributes in the case of simple or practical intentionality is not.<sup>27</sup>

#### 4. Semantic Externalism and the Attribution of Original Intentionality

Interpreting a community as exhibiting original intentionality is taking its members to adopt the discursive scorekeeping stance toward each other. The content-conferring norms and proprieties that an interpreter who attributes discursive scorekeeping practices takes to be implicit in them have a number of important structural features. Central among them is the fact that the conceptual norms implicit in the practices attributed to a community outrun the nonnormatively specifiable behavioral discriminations members of that community are disposed to make. For this reason, conceptual norms can be understood as *objective*, and so as binding alike on all members of a discursive community, regardless of their particular attitudes. This feature of attributions of linguistic practices secures the sense in which concepts and the commitments they involve concerning appropriate circumstances and consequences of application can be understood to be *shared*, in spite of the many differences of attitude that correspond to the different scorekeeping perspectives of the discursive practitioners who keep track of each other's statuses. This normative surplus of practice (as attributed by an interpreter) over behavior (nonnormatively specified) is also what is appealed to in responding to the issue raised by the possibility of *gerrymandering* (introduced above in 1.3.5)—the problem of what privileges one of the many ways of projecting from actual applications of concepts (and regularities and dispositions regarding such performances) commitments regarding cases that have not arisen for practical adjudication.

The reason the conceptual contents conferred by the discursive scorekeeping practices a community is interpreted as engaging in can outrun the community's capacity to apply them correctly and to appreciate the correct consequences of their application is the empirical and practical *solidity* or concreteness of those practices. The assertible contents a discursive interpretation takes to be conferred by communal deontic scorekeeping practices are inferentially articulated, but they are not merely placeholders in abstract, purely formal, relational structures—hollow shells waiting to be filled up by supplying actual facts and objects that somehow 'fit' them. For the content-

conferring practices do not relate the deontic statuses that bear those contents only to other deontic statuses. Discursive practice comprises noninferential entries and exits as well, and these (according to the interpreter attributing those practices) relate contentful doxastic and practical commitments to the worldly states of affairs that properly elicit acknowledgments of those commitments and are properly elicited by such acknowledgments, respectively. Standard discursive practices—those that encompass both empirical and practical dimensions—are solid (even lumpy), in that they involve actual objects and states of affairs, as well as the deontic statuses in terms of which score is kept.<sup>28</sup>

In such practices, the actual causal provenance or consequences of a deontic attitude—and not just the proprieties that connect its adoption to the adoption of other deontic attitudes—can matter (according to the external intentional interpreter attributing the content-conferring practices) for the content of the status it is an attitude toward. So an interpretation of this sort takes it that what an interlocutor who performs a certain speech act is committed to thereby, according to the practices of the relevant community, can depend on how things are in the nonlinguistic world. The interpreter takes it that the solid, corporeal communal practices determine what is being talked about (whether or not any scorekeepers in the community realize it), for those practices incorporate it. And the interpreter also takes it that what is being talked about determines what it is correct to say and infer, including practically (whether or not any scorekeepers in the community attribute the right claims and consequences). Interpretations that attribute original intentionality are accordingly *semantically externalist* in Davidson's sense.<sup>29</sup> This is part of what was called above (8.5.6) 'tactile Fregeanism': our practice puts us in touch with facts and the concepts that articulate them—we grasp them. But what we grasp by our practice extends beyond the part we have immediate contact with (its handles, as it were); that is why what we grasp is not transparent to us, why we can be wrong even about its individuation. How the world really is determines what we have gotten a hold of; but even though for that reason we do not know all the details about it, we still genuinely grasp it.

In this way the proprieties governing the application of a community's concepts are in part determined (according to the interpreter) by the actual properties of and facts concerning the things the linguistic practitioners are perceiving, acting on, and so talking about—which are just features of *their* practice (according to the interpreter). How the things and properties they are talking about actually are determines the correctness of the commitments of all community members alike. They are all bound by the *same* conceptual norms, regardless of the differences in collateral commitments that make particular claims have different inferential significances for different scorekeepers. According to the practices the interpreter takes them to be engaging in, they *share* a common set of concepts, which determines how the attitudes of those who keep score on each other are answerable to the facts.

When concrete discursive practices (including perceptual reporting and intentional agency) are ascribed to a community, the states of affairs that properly noninferentially elicit the acknowledgment of doxastic commitments and those that are properly noninferentially elicited by the acknowledgment of practical commitments are specified in the interpreter's own language. For instance, in assessing the extent to which the claims made by various community members do express facts, and so are correct uses of their concepts, the interpreter compares the commitments he or she attributes to them to those the interpreter undertakes—and similarly for assessments of their reliability as perceivers and agents. Semantic externalism is perspectival externalism.

To treat those interpreted as linguistic practitioners who use particular concepts is to treat them as bound by proprieties that project beyond their actual behavior and dispositions. The interpreter uses the norms implicit in his or her own concepts in specifying how the conceptual norms that bind the community being interpreted extend beyond the practitioners' actual capacity to apply them correctly. All the resources of the interpreter's home language are available in distinguishing one such set of proprieties from another; taking the interpreted interlocutors to have bound themselves by even a slightly different set of proprieties would be offering a different interpretation, attributing a different set of practices. The general point is that while normative interpretation of a community as engaged in one set of practices rather than another is underdetermined by nonnormatively specified actual behavior, regularities of behavior, and behavioral dispositions, *relative to such an interpretation*, concepts nevertheless are objective, shared, and unambiguously projectable.

### 5. *Sharing Inferentially Individuated Concepts*

It has been acknowledged throughout this exposition that an inferential conception of concepts raises *prima facie* difficulties for understanding what is involved in communication between individuals with different repertoires of commitments. The inferential significances of utterances of the same sentence produced by different performers are different—even where anaphoric and indexical phenomena are not in play. For their different collateral commitments make available different auxiliary hypotheses; hence what consequential commitments the performer undertakes by producing those performances and what would entitle their utterer to them (according to the scorekeeper who attributes the collateral commitments) are different. So something special needs to be said about the sense in which interlocutors with different collateral commitments can nonetheless be said to be able to make the same claims and express the same inferentially articulated concepts. It is worth rehearsing briefly the features of the discursive scorekeeping model that are appealed to in providing such an account.

What is from many points of view the most natural way out of this difficulty is not the path taken here. The most straightforward approach would be to adopt an egalitarian attitude toward the different inferences a concept is involved in. A privileged class of inferences would be distinguished, which are taken to be constitutive of the concept, while the rest are accorded a secondary status as turning out to be correct ways of using the concept so constituted. There is an undeniable intuitive basis for such a distinction: The inferences from "This tractor is completely green" to "This tractor is not completely red" and from "This cloth is scarlet" to "This cloth is red," for instance, have a different status from the inferences from "This tractor is completely green" to "This tractor is made by John Deere" or from "The apple in the box is a ripe Winesap" to "The apple in the box is red." The correctness of the first inference plausibly is taken to be part of the concepts *green* and *red*, while the correctness of the second sort is equally plausibly taken to be just a matter of empirical facts about John Deere tractors and ripe Winesap apples—*inferences whose correctness involves the concepts red and green without in any way constituting them.*

Quine, of course, argues that one way of construing the sort of concept-(or meaning-) constitutive privilege that distinguishes the first class is defective because it does not correspond to the sort of difference in the use of the words (the practical status of the inferences) that the theory behind it entails.<sup>30</sup> There do not seem to be any inferential connections that are unrevisable in principle, immune to being undermined by suitable empirical evidence, and so *a priori* for those who grasp the concepts involved. But this is not to say that no pragmatic sense can be made of the intuitive difference in status between two sorts of inferences instanced above. Sellars,<sup>31</sup> for instance, does not take all the materially good inferences involving a concept to be essential to it.<sup>32</sup> He picks out the privileged concept-constitutive inferential connections as those that support *counterfactual* reasoning, and so count as having *nomological* force. This is a real practical difference; this way of drawing the line does not fall afoul of Quine's strictures, for it by no means follows that these conceptual matters are *a priori*—we need to investigate the world to find out what the laws are, as for any other facts. Since the laws involved are not *a priori*, unrevisable, or immune to factual evidence, this is not a version of analyticity. According to this view, not only claims but concepts can be correct or incorrect, depending upon whether the inferences they incorporate correspond to actual laws.

The difference between inferential connections among concepts that are counterfactually robust and those that are not is an important one, and this fact accounts for the felt difference between the two sorts of inferences mentioned above. Nonetheless, nothing is made of it here. This is partly because the notion of nomologicality and counterfactual reasoning, important though it is in other contexts, has not been reconstructed in discursive scorekeeping terms as part of this project (though the key notion required,

that of the *incompatibility* of claims, has been given a pragmatic interpretation). But neither this nor any other way of picking out a privileged subclass of concept-constitutive inferences has been appealed to in individuating concepts here, for two other reasons.

First, mastery of a special subset of distinguished inferences (for instance, the counterfactually robust ones) is not in general sufficient for grasp of a concept. For such grasp requires that one be hooked up to the *function* that takes as its argument repertoires of concomitant commitments available as auxiliary hypotheses and yields inferential significances as its values. Carrying on a conversation involves being able to move from perspective to perspective, appreciating the significance a remark would have for various interlocutors. (More is said about this below, in connection with the representational dimension of discourse.) The effect that various auxiliary hypotheses have on the inferential significance of a claim relative to a particular doxastic context cannot be determined just from the privileged inferences it is involved in (for instance, the counterfactually robust ones), unless it is assumed that the repertoire in question contains conditionals corresponding to all the other materially good inferences (for example from the ripeness of Winesap apples to their redness). Assuming that is contrary to the spirit of this enterprise: it depends on the formalist view of inference, which sees enthymematically suppressed conditionals behind every material propriety of inference. In particular, such a view would have the consequence that communities that do not yet have the expressive resources of logical vocabulary such as the conditional were precluded for that reason from counting as employing nonlogical concepts such as *red*.

The second reason that the inegalitarian attitude toward inferences is not taken in individuating concepts is that no matter how the privilege distinguishing some supposedly uniquely concept-constitutive inferences is construed (as counterfactual robustness or otherwise), endorsement even of these privileged inferences can still vary from perspective to perspective. There can be different views about what the laws of nature are, for instance, just as there can be differences about the colors of John Deere tractors and ripe Winesap apples. Failure to agree about such large-scale empirical matters does not preclude the interlocutors from nonetheless having a hold on the same concepts. This is the 'tactile Fregeanism' that explains why people can be counted as having radically false (nomologically precluded) views that are nonetheless genuinely *about*, say, arthritis.

Thus the response to the difficulty of reconciling the possibility of genuine communication with an account that individuates concepts by inferential roles comes in two parts. The first is the social-perspectival move. It allows inferential significances to vary with doxastic perspective, while conceptual content, which determines a *function* from perspective to significance, does not. But both the perspective-relative inferential significances of potential speech acts and the perspective-independent conceptual contents that deter-

mine them (in context) are thoroughly *normative* notions—consisting in proprieties of discursive scorekeeping.

The crucial second part of the response is accordingly the normative-interpretive move. It distinguishes the proprieties governing *correct* use in which the concepts grasped by individuals consist, on the one hand, from the dispositions to apply concepts, make inferences, and perform speech acts, in which an individual's grasping of a concept consists, on the other—and so distinguishes concepts from conceptions of them. Talk of grasp of concepts as consisting in mastery of inferential roles does not mean that in order to count as grasping a particular concept an individual must be disposed to make or otherwise endorse in practice all the right inferences involving it. To be in the game at all, one must make *enough* of the right moves—but how much is enough is quite flexible. One of the strategies that has guided this work is a commitment to the fruitfulness of shifting theoretical attention from the Cartesian concern with the grip we have on concepts—for Descartes, in the particular form of the centrality of the notion of *certainty*, that is infallibility about the content grasped, including its individuation (so long as we access it clearly and distinctly)—to the Kantian concern with the grip concepts have on us, that is the notion of *necessity* as the bindingness of the rules (including inferential ones) that determine how it is correct to apply those concepts.

Interpreting the members of a community as engaging in discursive practices is interpreting them as binding themselves by objective, shared concepts whose proprieties of use outrun their dispositions to apply them. There is no answer that could be given in advance as to how much one must be able to get right in order to be interpreted as hooked up to one concept or another. Massive individual differences in inferential dispositions among interlocutors are compatible with interpreting them all as nonetheless governed by (answerable to) the same set of conceptual proprieties. For it is compatible with interpreting them as talking about the same objects, answering to the same set of objective facts. In this way the perspectival account of propositional contents (and so conceptual contents generally) combines the intentional and extensional approaches to communication outlined above in 7.5.

## 6. Three Levels of Norms

The normative phenomenalist methodology applies a version of the stance stance to the problem of understanding normative statuses such as the proprieties implicit in discursive scorekeeping practices. It does so by focusing on when it is appropriate to adopt a certain sort of attitude—the stance of interpreting a community as engaged in inferentially articulated deontic scorekeeping practices that confer particular conceptual contents. It has been explained what it is for an interpreter to attribute to a community discursive practices that confer objective, shared, projectable conceptual con-

tents. The question that remains is, What is it that determines when it is appropriate or correct to adopt one rather than another of these interpretations, to attribute one rather than another of those sets of discursive practices? (Recall that the corresponding question that was asked without being answered above was rather what made it appropriate to adopt any such normative interpretation at all—to attribute practices rather than mere behavior.)

Once again, the issue of the origin of the warrant for employing a normative vocabulary seems just to have been put off. Norms have been appealed to at three different interpretive levels. First, talking and thinking, grasping and applying concepts, is described in terms of inferentially articulated norms; moves in the game of giving and asking for reasons are made intelligible in terms of alterations in what one is committed and entitled to at each stage. This is a normative reconstrual of the *discursive* in terms of *deontic statuses*. Second, what it is to take or treat interlocutors in practice as committed or entitled, as exhibiting deontic statuses, is explained in terms of scorekeeping practices. The norms implicit in these practices govern the alteration of *deontic attitudes*. At this stage in the account, deontic statuses are understood as instituted by proprieties of scorekeeping—of systematically altering deontic attitudes and thereby assigning pragmatic significances to performances, paradigmatically the fundamental speech act of assertion. Reference to deontic statuses is made only as the objects of deontic attitudes; the only thing one can do with a commitment is to attribute it or undertake it (perhaps, but not necessarily, by acknowledging it).

The third stage applies the methodological strategy of normative phenomenism one more time, doing for deontic attitudes what those attitudes did for deontic statuses. The focus is now on the practices of attributing deontic attitudes—interpreting a community as engaged in implicitly normative discursive practices, as keeping deontic score by attributing and acknowledging deontic statuses. The account of deontic scorekeeping on doxastic and practical commitments explains what one must interpret a community as doing in order for it to be talking that one is thereby taking them to be doing. More precisely, it specifies conditions on the structure of practices a theorist attributes to a community that are sufficient for community members, so interpreted, to be treating each other as exhibiting propositionally contentful doxastic and practical commitments. Thus the relation envisaged between original intentionality and the stance of the interpreter who attributes it is analogous, at a higher level, to that obtaining between deontic statuses and deontic attitudes—for in place of a direct explanation of what commitment and entitlement are, an account of what it is to take someone to have such a status was offered. The phenomenalist explanatory retreat from status to attitude is applied at two levels, within the interpretation and in the relation the interpretation stands in to what is interpreted.

Norms come into the story at three different places: the commitments and

entitlements community members are taken to be attributing to each other; the implicit practical proprieties of scorekeeping with attitudes, which institute those deontic statuses; and the issue of when it is appropriate or correct to interpret a community as exhibiting original intentionality, by attributing particular discursive practices of scorekeeping and attributing deontic statuses. It is normative stances all the way down.

Regularities of communal behavior and disposition specified in nonnormative terms cannot *dictate* the attribution of scorekeeping practices that institute a particular set of normative statuses and confer a particular set of propositional contents. In adopting such a stance, the interpreter takes the interlocutors being interpreted to be *committed* to keeping score according to specific patterns, associating pragmatic significances with discursive performances that correspond to the inferentially articulated contents of the doxastic and practical commitments they express. The interpreter thereby undertakes commitments to various sorts of assessments of propriety of performance of those interpreted. Such commitments on the part of the interpreter are compatible with an indefinitely large lack of fit between the norms attributed and the actual performance of those to whom they are attributed, including their performance in assessing each other. This means that the normatively specified practices attributed by a discursive interpreter are always underdetermined by nonnormatively specified actual performances and dispositions; various sets of practices could be attributed as interpretations of the same behavior. So whenever an interpreter takes a community to be engaging in scorekeeping practices whose implicit proprieties confer one set of propositional contents on the deontic statuses they institute, there will always be alternatives, other sets of contents that could be taken to determine the pragmatic significances that scorekeepers ought to associate with discursive performances. Because of this slippage between the normative and nonnormative specifications of what community members are doing, the interpreter has considerable leeway in how to interpret them.

It remains, then, to discuss the nature of the norms that govern the choice of an interpretation of a community as engaging in one set of implicitly normative, content-conferring discursive scorekeeping practices rather than another, or rather than describing their behavior exclusively in nonnormative terms. This issue is best approached by considering the relation between the discursive scorekeeping stance adopted by the members of a linguistic community (according to an interpretation), on the one hand, and the stance adopted by the interpreter who attributes implicitly normative linguistic practices governing such scorekeeping attitudes (and so original intentionality), on the other. On the face of it, one major difference between the two stances is that discursive scorekeepers take up attitudes toward other members of their own communities, while an interpreter who attributes original intentionality takes the members of some *other* community to be discursive

scorekeepers. This is a misleading appearance, however. The important difference between these two sorts of norm-attributing stance is of a different sort. Indeed, under the right circumstances, the difference dissolves entirely, and the two stances coalesce. This collapse of levels provides the key both to understanding the status of the concept-articulating norms implicit in our discursive practices and also to understanding ourselves as not merely *rational*, but *logical* normative creatures, as not merely *expressive*, but *self-explicating* ones.

### III. WE HAVE MET THE NORMS, AND THEY ARE OURS

#### 1. *Original Intentionality and the Explicit Discursive Scorekeeping Stance*

The relation between the attitudes of an interpreter who attributes to a community discursive practices (and hence original intentionality), on the one hand, and the proprieties of scorekeeping implicit in those practices, on the other, is modeled on the relation between the deontic attitudes of scorekeepers and the normative statuses they attribute. What the discursive scorekeeper does *implicitly* (taking or treating others, to whom speech acts and discursive commitments are attributed, as discursive scorekeepers), the attributor of original intentionality to a community does *explicitly* (ascribing discursive scorekeeping attitudes). The underlying difference between the two stances is, not the distinction between communally *external* and *internal* attitudes or interpretations, but the distinction between *explicit* and *implicit* ones. Only a creature who can make beliefs explicit—in the sense of claiming and keeping discursive score on claims—can adopt the simple intentional stance and treat another as having beliefs implicit in its intelligent behavior. Just so, only a creature who can make its attitudes toward the beliefs of others explicit—in the sense of being able to ascribe scorekeeping attributions—can adopt the explicitly discursive stance and treat others as making their beliefs explicit, and so as having original intentionality.

Discursive scorekeeping is what the members of a community must be *doing* in order for any of their performances to have the significance (for them) of *saying* something. To take them to be a community of discursive scorekeepers whose practices confer conceptual contents, an interpreter must be capable of saying what they are doing—making *explicit* the broadly inferential proprieties that are (taken to be) *implicit* in their scorekeeping practices. For those who can adopt only the basic scorekeeping stance can attribute commitments to others (even to nonlinguistic, simple intentional systems) and can also take performances to have the significance of assertions, that is of explicit acknowledgments of discursive commitments. They

thereby implicitly recognize others as scorekeepers, and hence as attributors of commitments.

But adopting the basic discursive scorekeeping stance does not require attributing specific attributions to others; it does not require keeping score on their *attribution*s, as well as their *acknowledgments* of discursive commitments. In contrast, interpreting the members of a community as engaging in discursive scorekeeping practices requires attributing to them the full range of deontic attitudes: attributing particular attributions as well as particular acknowledgments. And attributions can be attributed only by being ascribed, for it is only when made explicit in the form of propositional contents that they can be embedded in one another and so iterated. Only someone who can say something of the form “*S* is committed to the claim that *S*’s committed to the claim that *p*” can adopt the attitude that it makes explicit.<sup>33</sup>

To attribute a particular conceptual content to an expression is to say something about how it is correctly used; to attribute such content to a state or status is to say something about the circumstances under which it is appropriately acquired or relinquished and the appropriate consequences of doing so. Interpreting a community as exhibiting original intentionality is taking it that the broadly inferential proprieties that articulate the conceptual contents of their expressions, performances, and states are implicit in their deontic scorekeeping practices. So one capable of adopting that interpretive stance must be able to attribute not only scorekeeping attitudes but also those implicit inferential proprieties, which relate the adoption of one scorekeeping attitude to another. Altering a deontic scorekeeping attitude is a practical doing—the sort of thing a specification of which can play the role of the conclusion of a piece of practical reasoning. So proprieties of scorekeeping can be expressed as proprieties of practical reasoning. Again, only someone who can say something of the form “*S* is committed to the claim that if a scorekeeper *does* attribute to *A* commitment to *p*, then the scorekeeper *should* attribute to *A* commitment to *q*” can adopt the attitude that it makes explicit.

This is to say that interpreting a community as engaging in discursive scorekeeping practices, and so as exhibiting original intentionality, requires the full expressive resources of the logical locutions whose use has been reconstructed here in scorekeeping terms. Ascriptional locutions are needed so that both essential flavors of deontic attitude can be attributed, not just adopted: *attribution*s as well as acknowledgments of commitments. Sentential logical vocabulary, paradigmatically the conditional, makes it possible to attribute acknowledgment of specifically *inferential* commitments. Normative vocabulary is required so that endorsement of a pattern of practical reasoning can be attributed.<sup>34</sup> Subsentential logical vocabulary such as quantifiers and identity locutions enable the attribution of endorsements of *substitutional* commitments, and so on. The expressive power of these logical

locutions is necessary and sufficient to make possible the adoption of the *explicit* discursive scorekeeping stance.

## 2. Expressive Completeness and Interpretive Equilibrium

Of course it is not just a coincidence that foregoing chapters have explained how to introduce into the basic discursive scorekeeping model just the sorts of logical vocabulary needed to make explicit the various inferentially articulated proprieties implicit in that practice—the very proprieties in virtue of which the expressions, performances, and deontic statuses governed by them count as expressing or exhibiting nonlogical conceptual contents. One of the criteria of adequacy that has guided the project from the outset is that it be possible to elaborate the model of discursive practice to the point where it is characterized by just this sort of *expressive completeness*. This means that the model reconstructs the expressive resources needed to describe the model itself. By means of these logical resources, the theory of discursive practices becomes expressively available to those to whom it applies. What is required is just that the scorekeeping practices that confer conceptual contents on the fundamental sorts of explicating vocabulary used in stating the theory and specifying the content-conferring discursive scorekeeping practices in the first place be themselves specified within the terms of the theory. The hypothetical practitioners who play the idealized Sprachspiel of giving and asking for reasons herein described can then be understood as themselves capable of *saying* what they have been supposed to be *doing*: they can make explicit the implicit practical proprieties in virtue of which they can make anything explicit at all.

Once the expressive resources of a full range of semantically and pragmatically explicating logical vocabulary are in play, those who have mastered them can keep discursive score *explicitly*, by making *claims* about each other's doxastic, practical, and inferential commitments. They can theorize about each other's scorekeeping attitudes. The broadly inferential scorekeeping proprieties that otherwise remain implicit, in the shadows of the practical background, are brought out into the full revealing light of explicit, public, propositional awareness. Particular ascriptions of commitment and entitlement, endorsements of consequential relations among them, and acknowledgments of and failures to acknowledge deontic statuses become topics for public challenge, justification, and debate. Though all the deontic attitudes and practical inferential know-how involved in scorekeeping cannot be made explicit in the form of claims and principles at once, there is no part of that content-constitutive practice that is in principle immune from such codification—out of reach of the searchlight of explicitation. Having been all along implicitly normative beings, at this stage of expressive development we can become explicit to ourselves as normative beings—aware both of the sense in which we are creatures of the norms and of the sense in which they are

creatures of ours. Having been all along implicitly discursive beings, at this stage of expressive development we can become explicit to ourselves as discursive beings—aware both of the sense in which we are creatures of our concepts (the reasons we produce and consume) and of the sense in which they are creatures of ours.

The members of a linguistic community who adopt the explicit discursive scorekeeping stance to one another achieve thereby a kind of *interpretive equilibrium*. Each one interprets the others as engaging in just the same sort of interpretive activity, as adopting just the same sort of interpretive stance, as one does oneself. This symmetric taking of others to adopt just the same sorts of attitudes one is oneself adopting, characteristic of the discursive scorekeeping stance, contrasts markedly with the *asymmetric* relation obtaining between an interpreter who adopts the *simple* intentional stance and the nonlinguistic creature interpreted as a simple intentional system. In that case the interpreter does not take the system being interpreted to be able to do just what the interpreter is doing, namely *attributing* (as opposed to acknowledging) beliefs, intentions, and endorsement of patterns of practical reasoning. This is one of the reasons what is attributed by such interpreters deserves to be understood as a derivative sort of intentionality.

Linguistic practitioners who have not yet deployed logical vocabulary implicitly treat other interlocutors as adopting the same interpretive stance that they do—as being discursive scorekeepers. The relations between interpreter and interpreted in such basic nonlogical discursive practices are accordingly also symmetric; an interpretive equilibrium is achieved in that case as well. Their idiom is not semantically and pragmatically explicitly complete, however; they adopt attitudes they cannot make explicit as the contents of commitments that can be acknowledged by assertion. They do not attribute the sort of attitude they are adopting just by attributing propositionally contentful commitments. They can only *implicitly* treat one another as scorekeepers, by keeping score on each other.

They treat others as in the general line of business of attributing commitments (and so being scorekeepers) by treating some of their speech acts as having the force or pragmatic significance of acknowledgments of commitments. For it is a necessary condition of being able to acknowledge (and so undertake) discursive commitments in general that one can also attribute them. So the interpretive equilibrium exhibited by basic nonlogical discursive scorekeeping practices is implicit and expressively incomplete. There is still an asymmetry between the stance such scorekeepers are interpreted as adopting by one who attributes original intentionality to the community in whose practices they participate, on the one hand, and the interpretive stance adopted by the interpreter who attributes such content-conferring practices, on the other.

That gap disappears—a complete and explicit interpretive equilibrium is achieved—for a community whose members have access to the full expres-

sive resources supplied by logical vocabulary. They can adopt the *explicit* discursive stance toward one another. Each scorekeeper can explicitly take the others to be doing just what that scorekeeper is doing: attributing discursive deontic attitudes, including that very sort of attribution. Such discursive practitioners have available as topics for explicit discussion the doings that underwrite their sayings, the practices in virtue of which anything can be explicit to or for them at all, and the interpretive stance they adopt to each other.

To the Kantian dictum that judgment is the form of consciousness has been added the claim that logic is the expressive organ of *self*-consciousness. Judging has been construed here as the practical attitude of acknowledging a certain kind of inferentially articulated commitment. Logical vocabulary then supplies the expressive resources needed to make explicit—to put in judgeable form—the semantic and pragmatic bases of judgment. By its means we come to be able to talk about proprieties of *inference*, about the structures of scorekeeping attitudes within which a performance can be accorded the significance of acknowledging or undertaking a *commitment*, and about the relations between these characteristics of specifically *discursive* practice as such. The complete and explicit interpretive equilibrium exhibited by a community whose members adopt the explicit discursive stance toward one another is social self-consciousness.<sup>35</sup> Such a community not only *is a we*, its members can in the fullest sense *say 'we'*.

### 3. Saying 'We'

Such a community-constitutive '*we*'-saying attitude is also the one adopted by those external interpreters who attribute to a community both original intentionality and the use of logical vocabulary. In the weakest sense, we treat others as among us by attributing to them, and interpreting their performances in terms of, propositionally contentful practical and doxastic commitments—that is, by adopting the simple intentional stance. In a more basic sense, we treat others as among us by taking them in addition to perform speech acts. Keeping discursive score in this fuller sense is implicitly treating them as rational scorekeeping creatures who can appreciate the inferentially articulated pragmatic significance not only of their own nonlinguistic performances but also of their claims and of the actions and speech acts of others. At the next level, explicitly keeping discursive score on the members of a community—by ascribing not only acknowledgments but attributions of propositionally contentful commitments—is attributing original intentionality. This is explicitly treating the members of a community as among us, in the sense of being rational linguistic creatures. The richest sort of '*we*'-saying is then taking those others to be in addition *logical* creatures—treating them as able to adopt, toward each other and at least potentially toward us, just the attitude we are adopting toward them.

So at the highest levels of ‘we’-saying, interpretive equilibrium is achieved (whether implicitly or explicitly). The interpretive stance attributed to the members of a discursive community approaches that adopted by the interpreter who attributes original intentionality to that community. Finally, the sort of scorekeeping that is—according to the interpreter outside the community—internal to and constitutive of the community being interpreted comes to coincide with the scorekeeping of the interpreter who attributes discursive practices to the members of that community. External interpretation collapses into internal scorekeeping. Thus attributing discursive practices to others is one form or another of ‘we’-saying. It is recognizing them as us.<sup>36</sup>

This assimilation of the external to the internal interpretive point of view means that the question of what it is to interpret the members of a community as engaged in discursive practices—what it is in this fundamental sense to say ‘we’ to them—has been answered by showing how the deontic scorekeeping model can be elaborated so as to make available the expressive power of logical locutions (in particular ascriptions, conditionals, and normative vocabulary). The next question dictated by the methodological strategy of normative phenomenism about discursive norms is then, When is it *proper* or *appropriate* to adopt such an interpretive stance? When is it appropriate to say ‘we’ in the sense of making what others do intelligible as the acknowledging and attributing of propositionally contentful doxastic and practical commitments? When is it appropriate to *interpret* their antics, as we do for each other, rather than merely to *explain* them, as we do for nonsapient?

The collapse of the external explicit discursive interpretive stance into scorekeeping within our own expressively sophisticated practices transforms this from an abstract theoretical question into a concrete question about our own practices. Understood that way, the proper answer would seem to be latitudinarian (as suggested in the opening paragraphs of Chapter 1): one *ought* to adopt the discursive scorekeeping stance whenever one *can* adopt it. For on the one hand, the detailed requirements one must satisfy in order to count as adopting such an interpretation are stringent. Not just any group of interacting organisms can be made out to be attributing to each other commitments whose inferential and social articulation suffices to confer genuinely propositional contents on their performances. So there is little danger of such a generous policy leading to the facile or promiscuous extension of the franchise of sapience to those undeserving of it. And on the other hand, the rewards for adopting the discursive scorekeeping attitude wherever it is possible are great. Conversation is the great good for discursive creatures. Extending it increases our access to information, our knowledge, and our understanding—our semantic and pragmatic self-consciousness. Those who *can* be understood as fellow strugglers in the enterprise of making it explicit *should* be so understood.

Adopting such an inclusive demarcational attitude is saying ‘we’ to whoever can be understood as adopting demarcating practical attitudes—as them-

selves distinguishing by their scorekeeping a 'we' of rational agents and knowers, inhabiting a normative space of giving and asking for reasons, from an 'it' that comprises what does not live and move and have its being in such a space. Establishing entitlement to such a commitment with respect to demarcation in general would not, however, resolve the more specific issue of the status of the norms that govern the selection of one particular discursive interpretation rather than another. For the underdetermination of normative interpretation by behavior and dispositions specified in nonnormative terms means that whenever what a community does supports an interpretation of its members as engaging in discursive practices in which one set of conceptual norms is implicit, that behavior also supports rival gerrymandered interpretations of them as engaging in discursive practices in which different sets of conceptual norms are implicit. When it is possible to offer some such interpretation, how is it settled which one is most appropriate?

#### *4. Semantic Externalism Begins at Home*

The previous issue was a global one, concerning the propriety of attributing discursive scorekeeping practices at all. The present issue is a local one: assuming the global question settled in the affirmative, what is involved in choosing among various specific alternatives? Deciding to treat each of the members of some alien community as one of us (in the sense of treating them as adopting deontic attitudes, attributing and acknowledging propositionally contentful commitments) by no means settles what those contents and commitments should be taken to be. Their speech acts will typically differ in their nonnormative characteristics; they will utter different noises, make different marks (or, for all that it matters to the abstract scorekeeping model of discursive practice, turn colors, emit odors, shift voltages). What about the conceptual contents they express? What the contents of their commitments and expressions are depends on their inferential practices and on the noninferential perceptual circumstances of application and practical consequences of application implicit in their scorekeeping practices. These may differ from ours in a myriad of details and still be intelligibly interpretable. How radically different might they be?

Both the question of what makes a better discursive interpretation and the question of how different from ours the practices of the others might be taken to be before it becomes impossible to offer an intelligible interpretation of them as in the same discursive line of work as we are—as scorekeeping by changes of deontic attitude with the right social and inferential structure to confer propositional content—are questions that can be addressed only by appeal to our actual practices of interpretation in conversation. Because in a community with sufficient expressive resources the tasks of external discursive interpretation and of internal communicative interpretation are tasks of the same kind, looking at the dynamics of intralinguistic interpretation in

ordinary conversation reveals the essential features that determine also the dynamics of interlinguistic interpretation.<sup>37</sup> This is to say that there is no usefully general answer to the more specific interpretive question. The coalescence of external and internal discursive interpretation dictates a regress to the background language, to *our* discursive practices. The norms that determine the propriety of choices as to which discursive practices, and so which implicit conceptual norms, to attribute to those we take to be talkers are not available in advance as a set of explicit principles. They are implicit in the particular practices by which we understand one another in ordinary conversation.

The question the interpreter faces is to determine what discursive norms the members of a community have instituted, what conceptual contents they have conferred, by their linguistic practices and deontic attitudes. According to the scorekeeping model, two sorts of attribution are involved in such interpretation. The concepts according to which the truth of their claims and the success of their actions (and so their reliability as empirical reporters and practical agents) should ultimately be assessed are the ones they have committed themselves to (a matter of deontic status) by their dispositions to acknowledge some commitments in their linguistic and nonlinguistic behavior (a matter of deontic attitude). According to the interpreter, the conceptual contents practitioners have bound themselves by can outrun their discriminative dispositions to acknowledge their commitments. For this reason, objective, shared concepts can be understood as projecting beyond the dispositions to apply them of those whose concepts they are. The collapse of external interpretation into internal scorekeeping shows that this semantic externalism is just a special case of the sort of perspectival scorekeeping that has been in play all along: the commitments a scorekeeper *attributes* to someone outrun those that individual *acknowledges*. In acknowledging one discursive commitment, one is in general undertaking others, whether or not one knows what they are. This is the pragmatic (scorekeeping) significance of the *inferential* articulation of their semantic contents.

So the job of an external attributor of linguistic practices is just a special case of the job any discursive scorekeeper has: each must keep two sets of books, distinguishing and correlating the commitments interlocutors are disposed to acknowledge by overt performances, on the one hand, and those they undertake thereby, on the other. These correspond to two ways of specifying the contents of their claims—those made explicit in *de dicto* and *de re* ascriptions, respectively. For recall that *de dicto* specifications extract inferential consequences only with respect to auxiliary hypotheses (including those inferential commitments that would be made propositionally explicit in the form of conditional claims) the ascriptional target acknowledges as collateral commitments. *De re* specifications extract those consequences by appealing to auxiliary hypotheses (including inferential ones codifiable as conditionals) that are (according to the ascriber) *true*. Ordinary intralinguistic

communication—the ability to carry on a conversation across the most ordinary differences in doxastic perspective—requires that scorekeepers be able to move back and forth between these two sorts of specifications of the contents of the commitments they attribute. The contents of the commitments it is appropriate to attribute to another depend both on the commitment-acknowledging performances (both linguistic and nonlinguistic) the ascriptional target is disposed to perform and on how things actually are with the objects being talked about. Mastering our practices of attributing conceptually contentful commitments is learning how in particular cases to adjudicate the claims of these two sources of content. Semantic externalism—the way in which what we mean depends on how things actually are, whether we know how they are or not—is a feature of the perspectival character of propositional content.

So semantic (perspectival) externalism begins at home. The contents of the commitments attributed to others, the concepts they have bound themselves by, cannot be specified apart from reference both to what they are disposed to do and say and to what is true of what they are making claims about. For what actually follows from what (according to a scorekeeper = interpreter) depends on the facts (according to that scorekeeper = interpreter). The point that matters here is that once the task of external interpretation is recognized as a special case of internal interpretation (scorekeeping), the practical norms that govern the attribution of one set of conceptually contentful commitments rather than another can be recognized as just one more instance of deciding what others of us are talking about and what they are saying about it. Our norms for conducting ordinary conversations among ourselves are the ones we use in assessing interpretations. There is never any final answer as to what is correct; everything, including our assessments of such correctness, is itself a subject for conversation and further assessment, challenge, defense, and correction. The only answer to the question of what makes one interpretation better than another is what makes one conversation better than another. The answer is a matter of our practical norms of understanding one another here at home.

So the norms governing the use of the home idiom determine how to project the concepts used to specify the content of the stranger's attitudes (which determine how it would be proper to apply those same concepts in novel situations) in the same way they do for the ascriber's own remarks. This is so even in the case where the stranger is best made intelligible by attributing concepts that differ from those used in the home community. Thus the collapse of external into internal interpretation means that the problem caused by the existence of gerrymandered alternatives to any particular discursive interpretation of another community from the outside is displaced to the context of interpretation and projection within our own community. This regress to our own interpretive practices dissolves, rather than solves, the gerrymandering problem concerning the relation between

regularities and norms. For there is no general problem about how, from *within* a set of implicitly normative discursive practices, what we do and how the world is can be understood to determine what it would be correct to say in various counterfactual situations—what we have committed ourselves to saying, whether we are in a position to get it right or not. The account of the use of *de re* ascriptions of propositional attitude shows explicitly just what is involved in such a determination.

For our own practices come to us with the norms in; we do not just utter noises, we undertake commitments, adopt normative statuses, make pragmatically significant moves in the game of giving and asking for reasons. That there is a vocabulary, for instance any nonnormative one, that does not have sufficient expressive power to make it possible to specify our practices, make the distinctions we make, project in the way we do, has, from within our practices, no particular significance. We are always already inside the game of giving and asking for reasons. We inhabit a normative space, and it is from within those implicitly normative practices that we frame our questions, interpret each other, and assess proprieties of the application of concepts.

The account being offered is embodied in the trajectory described by attempts to answer the question, Where are the norms?

The normative first appears in the story in the guise of deontic statuses, of commitments and entitlements. Thought and talk are presented as structures of commitments and entitlements, with particular expressions having the conceptual contents they do because of the role they play in an inferentially articulated structure of such deontic statuses.

Talk of deontic statuses is then traded in, however, for talk of the attitudes of taking or treating people as committed or entitled. Deontic statuses are revealed as scorekeeping devices used for identifying and individuating deontic attitudes. In this sense the first set of norms turns out to be in the eyes of their beholders. This does not amount to a reduction of the normative to the nonnormative, however, because not only actual attitudes, acknowledgments and attributions of deontic status, but also practical *proprieties* governing the adoption and alteration of such attitudes are invoked in explaining the institution of deontic statuses by discursive scorekeeping practices.

At the next stage, these proprieties themselves are removed to the eye of the discursive interpreter, who takes a community to exhibit original intentionality by attributing to it discursive practices socially and inferentially articulated in such a way as to confer propositional contents. Once again, however, it is not only the *actual* attitudes adopted by external interpreters that must be considered but also *proprieties* governing the adoption of the discursive stance and commitment to a particular interpretation.

With the collapse of external into internal interpretation—its revelation as a special case of the sort of interpretation that goes on all the time within the practices of a discursive community—those proprieties are assimilated

to the ordinary scorekeeping proprieties in play in *our own* discursive practices. The norms turn out to be . . . here.

### 5. Making It Explicit

So the theoretical attempt to track down the 'source' of the normative dimension in discourse leads us right back to our own implicitly normative practices. The structure of those practices can be elucidated, but always from within normative space, from within our normative practices of giving and asking for reasons. That is the project that has been pursued in this work. Its aim is not reductive but expressive: making explicit the implicit structure characteristic of discursive practice as such.

The irreducibly normative pragmatics (theory of social practice) presented here is elaborated in terms of the basic deontic statuses of *commitment* and *entitlement* to commitments, and the essentially perspectival scorekeeping attitudes of *attributing* and *acknowledging* those deontic statuses. The semantics, or theory of the sorts of conceptual content that can be conferred by such deontic scorekeeping practices, takes the form of an account of the *inferential*, *substitutional*, and *anaphoric* articulation that distinguishes specifically *discursive* commitments. The result is a use theory of meaning—a specification of the social-functional roles that doxastic and practical commitments and the speech acts that express them must play in order to qualify as semantically contentful. The sorts of content addressed are those traditionally grouped together under the heading of 'intentionality'. Saying what pragmatic scorekeeping significance speech acts must have to count as assertions makes it possible to explain *propositional* contentfulness in turn as what can in that sense be made explicit—as what can in the first instance be *said* (as well as believed or meant or done). *Empirical* and *practical* contributions to such propositional (assertable, and so believable) contents are explained in terms of their conceptually articulated incorporation of the appropriate causal antecedents (in perception) and consequents (in action) of acknowledgments of discursive commitments. The *representational* dimension of propositional contents is explicated in terms of the social-perspectival character of discursive scorekeeping and the substitutional substructure of its inferential articulation. In this way it is possible to understand what is involved in assessments of judgments as *objectively* true or false—as correct or incorrect in a sense that answers to the properties and relations of the objects they are *about*, rather than to the attitudes of any or all of the members of the community of concept users.

One of the leading ideas of this enterprise is that developing an account of how semantics is rooted in pragmatics (meaning in use, content in social-functional role) is an exercise not only in the philosophy of language and the philosophy of mind but also in the philosophy of *logic*. Discursive practice is understood in terms of reasoning and representing, but above all in terms

of *expressing*—the activity of making it explicit. The expressive role distinctive of logical vocabulary is its use in making explicit the fundamental semantic and pragmatic structures of discursive practice, and hence of explicitness and expression. Pursuing the ideal of expressive completeness requires working out an account of the practices of using various particular logical locutions—paradigmatically those used to express inferential, substitutional, and anaphoric commitments and those used to ascribe discursive commitments to others.

In the end, though, this expressive account of language, mind, and logic is an account of who *we* are. For it is an account of the sort of thing that constitutes itself as an expressive being—as a creature who makes explicit, and who makes itself explicit. We are sapients: rational, expressive—that is, discursive—beings. But we are more than rational expressive beings. We are also *logical, self*-expressive beings. We not only make *it* explicit, we make *ourselves* explicit as making it explicit.

---

**Abbreviations**

**Notes**

**Index**

---



---

# Abbreviations

---

The following abbreviations appear frequently throughout the Notes:

- BGS* Gottlob Frege. *Begriffsschrift*. 1879.
- "*EPM*" Wilfrid S. Sellars. "Empiricism and the Philosophy of Mind." Reprinted in *Science, Perception, and Reality*. London: Routledge and Kegan Paul, 1963.
- FPL* Michael Dummett, *Frege's Philosophy of Language*. New York: Harper and Row, 1973.
- GL* Gottlob Frege. *Die Grundlagen der Arithmetik*. 1884. English translation, *Foundations of Arithmetic*, by J. L. Austin. Evanston, Ill.: Northwestern University Press, 1959.
- PI* Ludwig Wittgenstein. *Philosophical Investigations*. Translated by G. E. M. Anscombe. Oxford: Basil Blackwell, 1953.
- PPPW* J. Sicha, ed. *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars*. Reseda, Calif.: Ridgeview Publishing, 1980.
- PW* H. Hermes, F. Kambartel, and F. Kaulbach, eds. *Gottlob Frege: Posthumous Writings*. Chicago: University of Chicago Press, 1979.
- RFM* Ludwig Wittgenstein. *Remarks on the Foundations of Mathematics*. Edited by G. H. von Wright, R. Rhees, and G. E. M. Anscombe. Translated by G. E. M. Anscombe. Cambridge: MIT Press, 1972.
- "*STSSD*" Robert Brandom. "Singular Terms and Sentential Sign Designs." *Philosophical Topics* 15, no. 1 (Spring 1987): 125–167.
- "*ÜSB*" Gottlob Frege. "Über Sinn und Bedeutung." In *Philosophical Writings of Gottlob Frege*, ed. P. Geach and M. Black, pp. 56–78. Oxford: Blackwell, 1969.



---

# Notes

---

## 1. Toward a Normative Pragmatics

1. The particular way in which Kant understands what theoretical and practical concepts are rules for doing—namely manipulating representations either by synthesizing many under one or as determinations of the will—depends on further, independent commitments that are not here in question.
2. That is the lesson of his “Was ist Aufklärung?” Indeed, for this reason the line between Cartesian and Kantian approaches should not be drawn so sharply as to imply that Descartes had no inkling of the significance of normativity, which becomes an explicit concern for Kant. His idea of the mental as a special stuff can be seen as a response to those issues, as yet only dimly appreciated. Descartes’s sense of the mental as special is precisely an inchoate awareness that its essence lies in rational, hence normative, interconnectedness. This makes it impossible to fit into what we now think of as nature, according to a conception of nature that was being formed around Descartes’s time. (Thanks are due to John McDowell for emphasizing this important point.)
3. In the unpublished 1897 draft of “Logic,” in *PW*, p. 147.
4. From another fragment on logic, *ibid.*, p. 4.
5. *Ibid.*, p. 145.
6. *Ibid.*, p. 144.
7. *Ibid.*, p. 145.
8. *Ibid.*, p. 128.
9. *Ibid.*, p. 4. Sometimes the point is put in terms of reasons, correct inference, or

justification: "Logic has a closer affinity with ethics [than psychology] . . . Here, too, we can talk of justification, and here, too, this is not simply a matter of relating what actually took place or of showing that things had to happen as they did and not in any other way" (*ibid.*).

10. Crispin Wright calls this Wittgenstein's "contractual" model of meaning and understanding (though for reasons that will emerge, the overtones of explicitness incorporated in this way of talking about the normative dimension are less than happy) (*Wittgenstein on the Foundations of Mathematics*, [Cambridge: Harvard University Press, 1980], p. 19). John McDowell describes it as the idea that we are "committed to certain patterns of linguistic usage by the meanings we attach to expressions" ("Wittgenstein on Following a Rule," *Synthese* 58 [1984] 325–363). As Saul Kripke puts it: "The relation of meaning and intention to future action is normative, not descriptive" (*Wittgenstein on Rules and Private Language*, [Cambridge: Harvard University Press, 1982], p. 37).
11. For example, at *PI*, § 146.
12. For example, at *RFM*, II121.
13. *PI*, § 193.
14. *RFM*, II116. One crucial difference is that the *laws* of society are explicit—they say what is correct and what is not. The assumption that all the laws of inference (not just those of logic) are explicit in this sense generates a regress, discussed in the next section.
15. *PI*, § 195.
16. *Ibid.*, § 217.
17. Two sorts of norms have been pointed out as involved in attributions of intentional states. On the one hand, intentional states stand in normative relations to each other: acquiring one belief commits one to believing its inferential consequences, intending to make-true a certain claim commits one to intending the necessary means, having certain constellations of beliefs and desires can commit one to form corresponding intentions, and so on. On the other hand, intentional states stand in normative relations to states of affairs that are not intentional states: there is a certain sort of normative *accord* between a belief and the state of affairs that must obtain for it to be true, between a desire and the states of affairs that would satisfy it, between an expectation and the states of affairs that would fulfill it, between an order and the performances that would count as obeying it. The first sort of normative relation is broadly *inferential*, the second is broadly *referential*. Although Wittgenstein invokes both sorts, his primary concern is with the latter. The strategy of this work is to start with the former kind of norm and to explain the latter kind in terms of it.
18. Thus the norms incorporated in the content of a belief concern not only what other beliefs one is committed to by having that belief (and in the context of other intentional states, how one is committed to act) but also how one thereby is committed to the world's being—to be assessed by determining what objects one's belief is *about*, and what is *true* of them.
19. The more general Kantian view at stake is that concerning the normative character of concept use. The more specific view is the understanding of norms as having the form of explicit rules. The juridical idiom he employs systematically obscures the distinction between these two commitments.
20. This is a different sense from the one that Sellars, whose views are discussed below, attaches to this expression.

21. *PI*, § 201.
22. *Ibid.*, § 84.
23. *Ibid.*, § 198.
24. *Ibid.*, § 201.
25. *Ibid.*, § 289 and *RFM*, V33.
26. *PI*, § 202.
27. *Ibid.*, § 199.
28. *Mind* 4 (1895): 278–280.
29. Gilbert Ryle, *The Concept of Mind* (New York: Barnes and Noble, Harper and Row, 1949), chap. 2.
30. *PI*, § 78.
31. What matters for the present project is the opposition between these two orders of explanation. But since they have been set out in connection with actual historical figures, as the lesson Wittgenstein has to teach Kant, it should at once be acknowledged, if only parenthetically, that when one looks at the details, Kant is somewhat better off than he appears in this sketch, for he does appreciate the point that Wittgenstein is making. Kant's acknowledgment of the possibility of a regress of rules appears in his discussion of the faculty of *judgment* (*Urteilskraft*): "If understanding in general is to be viewed as the faculty of rules, judgment will be the faculty of subsuming under rules; that is, of distinguishing whether something does or does not stand under a given rule (*casus datae legis*). General logic contains and can contain no rules for judgment . . . If it sought to give general instructions how we are to subsume under these rules, that is, to distinguish whether something does or does not come under them, that could only be by means of another rule. This in turn, for the very reason that it is a rule, again demands guidance from judgment. And thus it appears that, though understanding is capable of being instructed, and of being equipped with rules, judgment is a peculiar talent which can be practised only, and cannot be taught" (*Critique of Pure Reason*, A132/B171). The regress-of-rules argument is here explicitly acknowledged, and the conclusion drawn that there must be some more practical capacity to distinguish correct from incorrect, at least in the case of applying rules. Very little is made of this point in the first two Critiques, however. Kant's own development of this appreciation of the fundamental character of this faculty of acknowledging norms implicit in the practice of applying explicit rules, in the third Critique, has an immense significance for Hegel's pragmatism, but only his formulation of the issue seems to have influenced Wittgenstein's. The Appendix to this chapter discusses Wittgenstein's use of 'rule' in more detail.
32. P. 60 of "Realism and the New Way of Words," in *PPPW*, pp. 219–256. Another early paper that is important in this connection is "A Semantical Solution to the Mind-Body Problem" (also in *PPPW*), which argues for the paired claims (1) that mental concepts are semantic, metalinguistic concepts and (2) that semantic concepts are normative concepts.
33. "Some Reflections on Language Games," in *Science, Perception, and Reality* (London: Routledge and Kegan Paul, 1963), p. 321.
34. *Ibid.*
35. From "Language, Rules, Behavior," *PPPW*, p. 155. In a similar vein he says: "The mode of existence of a rule is as a generalization written in flesh and blood, or nerve and sinew, rather than in pen and ink" (from the same essay, p. 139). Talk

of rules as generalizations, even incarnate ones, is dangerous in this connection, however, for it flirts with a reductive regularism (about which more below) that identifies proprieties of practice with regularities of conduct.

36. Sellars, "Some Reflections on Language Games," p. 322.
37. *Logic*, trans. R. S. Hartman and W. Schwarz (New York: Dover Publications, 1974], p. 3.
38. Kripke, *Wittgenstein on Rules*.
39. McDowell, "Wittgenstein on Following a Rule," p. 342. It should be acknowledged that McDowell construes the structure of Wittgenstein's argument differently from the way it is presented here. He takes it that the identification of understanding with interpreting presents two unacceptable alternatives: either the regress of rules does not stop, in which case the norms evanesce, since every action is in accord with any given norm on some interpretation and fails to accord on some other, or platonistic, self-applying norms are imagined as the *last* interpretation. Here this platonistic "rails laid out to infinity" misconstrual was presented as arising independently of identifying understanding with interpreting—as a way of misunderstanding norms on a quasi-causal model.
40. It should be clear that to insist on this point is *not* to claim that one *cannot* explicitly say what ought to be done, say by promulgating a rule or giving an order. Nor is it to claim that where one *does* follow such an explicit rule, one must be interpreting it. Precisely not. In the typical case the understanding of what is explicit, the following of a rule, is itself practical—the exercise of implicit understanding or "know-how." One of the central tasks of this work is to say what one must be able to *do* in order to count as in this sense understanding an explicit claim, rule, or order.
41. This usage of 'discursive' is Kant's. See for instance pp. 21, 34, 82 of his *Logic*.
42. Making out this distinction is really the subject of the whole of the *Critique of Practical Reason*. The rational will is defined this way in Section 7 of Part I, Book 1, Chapter 1, p. 32 of the Akademie Textausgabe.
43. "Heidegger on Being a Person," *Nous* 16 (1982): 16.
44. This point is related to McDowell's criticism, discussed below, of social regularity theories of the sort Kripke and Wright attribute to Wittgenstein, which make the community of assessors incorrigible.
45. Although this seems the natural way to elaborate the picture, it is not evidently incoherent to imagine one organism shaping its own behavior by responding to its responses with positively and negatively reinforcing behavior. What makes such a suggestion odd is that one would think that the capacity to distinguish correct from incorrect performance that is exercised in the postulated responsive disposition to assess would also be available at the time the original performance is produced, so that no behavior-shaping ground would be gained by the two-stage procedure. But this need not be the case; the assessment might be addressed toward the performance as characterized by its consequences, discernible more readily in the event than the advent. It is no doubt more difficult to tell a story about how such self-reinforcing patterns of behavior might come about in one animal than in a group, for the behavior-shaping in question is not here, as it is in the regulist versions, *deliberate*, a matter of explicitly expressible intentions. Yet the issue of what it would be for there to be norms implicit in practice ought to be kept distinct from the issue of how such practices might in

fact plausibly arise. If the intra-organism reinforcement story is coherent, then regularity versions of the sanctions approach to implicit norms need be social only in the sense that they essentially involve the distinction of perspective between producing performances and assessing them. This contrasts with Sellar's story, in which the behavior shaping by reinforcement is deliberate and the regularity of conduct aimed at is accordingly explicitly expressible by the assessors, even though the assessors and the assessed may be time-slices of the same organisms (and full-fledged membership in the community may require playing both roles at some time). That account seems genuinely to require that there be performances where the assessing individual and the individual producing the performance being assessed are distinct. For only cases of this sort can be appealed to in making intelligible the norms implicit in grasp of a concept in such a way as to have any leverage at all against the regress-of-interpretations argument that motivates this approach. So the diachronic regulist sanctions theory sees norms as implicit in specifically *social* practice in a stronger sense.

46. For instance, McDowell "Wittgenstein on Following a Rule," p. 350: "If regularities in the verbal behaviour of an isolated individual, described in norm-free terms, do not add up to meaning, it is quite obscure how it could somehow make all the difference if there are several individuals with matching regularities."
47. Construing communal assessment regularity theories (paradigmatically those Kripke and Wright attribute to Wittgenstein) as offering an account of what it is for norms to be implicit in practice is implicitly disagreeing with one of McDowell's central criticisms of the relevance of such theories to Wittgenstein's text. McDowell objects: "The fundamental trouble is that Kripke makes nothing of Wittgenstein's concern to reject the assimilation of understanding to interpretation" (*ibid.*, p. 343). He is right that neither Kripke nor Wright makes anything of this crucial motivating line of thought, which is rehearsed in Section II of this chapter. But he overlooks the fact that the theory they do elaborate can nonetheless be understood as an attempt to address just the considerations that are motivated by the regress-of-interpretations argument. For they can be seen as concerned to provide a notion of what it is for norms to be implicit in practice or for practice to be implicitly norm-governed rather than explicitly rule-governed. As will emerge, this repudiation is consistent with endorsement of McDowell's other criticisms of this line of thought and interpretation.
48. Kripke, *Wittgenstein on Rules*, p. 108.
49. Wright, *Wittgenstein on the Foundations of Mathematics*, p. 220. Although Wright explicitly addresses only the significance of linguistic performances, his point applies more generally to acting correctly according to one's intentional states. A communal assessment regularity theory is also put forward in the author's "Freedom and Constraint by Norms," *American Philosophical Quarterly*, April 1977, pp. 187–196, reprinted in *Hermeneutics and Praxis*, ed. R. Hollinger (Notre Dame, Ind.: University of Notre Dame Press 1985).
50. Davidson is a notable exception, taking linguistic practice and therefore intentionality to be essentially social only in the sense that it can be made intelligible only in the context of mutual interpretation—an *I-thou* relationship, in the current terminology.
51. One example of how this demand could be met by defining community membership in such a way as to preserve the distinction between those governed by

the practice and those whose practice it is they are governed by, without disjoining the groups, is provided by Sellars's account of pattern-governed practice. This is what going intergenerational does for him—the judgments of the assessors who train new community members are authoritative, and those they assess and train are the community members subject to their authority.

52. These correspond to the two sorts of objections to individual regularity or dispositional theories that Kripke (*Wittgenstein on Rules*) offers. McDowell ("Wittgenstein on Following a Rule") argues that the social regularity theory Kripke then suggests Wittgenstein endorses in response is subject to an objection of the second sort, namely that it fails to distinguish between a claim's being correct (normative status) and its being taken to be correct (normative attitude) by the community as a whole. It is argued above that this approach also falls foul by importing illicit notions of communal assessment, normative statuses such as community membership (being subject to communal authority), and expertise (exercising communal authority).
53. Some account along these lines has been a popular post-Enlightenment reading of what is being allegorically communicated by the supernatural retributive strand in Christian ethical theory.
54. Mill, in *Utilitarianism* (reprinted in *Essential Works of John Stuart Mill*, ed. Max Lerner [New York: Bantam Books, 1965], 3:215), introduces the vocabulary of internal and external sanctions, but to point to a different distinction than that intended here. His "internal" sanctions are internal to the individual (rather than to the space of norms). A paradigm would be feelings of shame or guilt.
55. *PI*, § 201.
56. Typically, though not in every case, by not letting it begin—since in the commonest cases we understand explicit claims, rules, principles, orders, and so on without interpreting them.
57. The theory developed in this work incorporates both of these suggestions. But at this point in the exposition no specific interpretation of either has yet been endorsed.
58. Wright (*Wittgenstein on the Foundations of Mathematics*) and Kripke (*Wittgenstein on Rules*) offer interpretations along these general lines.
59. This reading is closely related to McDowell's criticism of readings of passages such as these: "If I have exhausted the justifications I have reached bedrock, and my spade is turned. Then I am inclined to say: 'This is simply what I do'" (*PI*, § 217); "When I obey a rule, I do not choose. I obey the rule blindly" (§ 219); "How do I know [how I intend the pattern to be continued]?—If that means 'Have I reasons?' the answer is: my reasons will soon give out. And then I shall act, without reasons" (§ 211). These succumb to the temptation to conclude that, "at the level of 'bedrock' (where justifications have come to an end), there is nothing but verbal behavior" (McDowell, "Wittgenstein on Following a Rule," p. 341). That is to think of the bedrock of unreflective practice exclusively in nonnormative terms of behavioral dispositions and regularities. But as McDowell points out, one should not conclude that where justification has run out, normative assessments no longer apply. That is just what the regress-of-rules argument for the existence of norms implicit in practice shows. Wittgenstein says, in a claim important enough to appear verbatim in both *PI* and *RFM*: "To use the word without a justification does not mean to use it wrongfully [zu]

*Unrecht gebrauchen]*" [RFM, V33; PI, § 289]. As McDowell says: "It seems clear that the point of this is precisely to prevent the leaching out of norms from our picture of 'bedrock'—from our picture, that is, of how things are at the deepest level at which we may sensibly contemplate the place of language in the world" ("Wittgenstein on Following a Rule," p. 341).

60. P. 5, par. 2, of *On the Law of Nature and of Nations*, trans. of 1688 ed. C. H. Oldfather and W. A. Oldfather, vol. 2 in the Classics of International Law series (reprint, New York: Oceana Publications; London: Wiley & Sons, 1964).
61. *Ibid.*, par. 3.
62. *Ibid.*, par. 4.
63. *Ibid.*, p. 6, par. 5.
64. One of Heidegger's central concerns in *Being and Time* (trans. J. Macquarrie and E. Robinson [New York: Harper and Row, 1963]) is to deny this characteristic Enlightenment thought, by describing how the value-free presence-at-hand (*Vorhandenheit*) studied by the physicist is abstracted from the value-laden readiness-to-hand (*Zuhandenheit*) of everyday life. Here is a characteristic statement: "In interpreting we do not so to speak, throw a 'signification' over some naked thing which is present-at-hand, we do not stick a value on it" (sec. 32, p. 190). This view is discussed in detail in the author's "Heidegger's Categories in Being and Time," *Monist* 66, no. 3 (July 1983): 387–409.
65. The evolution of physics from its "atoms in the void" conception has not appreciably altered the difficulty of fitting norms into the natural scientific world-picture. It is this difficulty that motivates both the Kantian dualism of norm and fact and the Kantian normative idealism that subordinates the latter to the former. Since the normative force of the better reason is not easily understood in terms of the sort of causal forces invoked by Newton, a normative conception of the way in which the necessity codified in laws outruns mere regularities is called in to support an understanding of causes in terms of proprieties governing the employment of concepts. Whatever one thinks of this heroic inversion strategy for reuniting the disparate elements of the Kantian dualism, its motivation underscores the difficulty of accommodating the normative within the natural.
66. It is somewhat disingenuous to characterize his view in terms of *our* attitudes. Although his primary concern in this work is with the moral attributes instituted by human beings, Pufendorf also acknowledges (as which seventeenth-century philosopher did not?) that God is also an intelligent being and can also impose or institute moral attributes. In this sense, God is treated as one of us. But even for God, in this respect *primus inter pares*, creation of the physical world is one thing, imposing moral attributes on it something else.
67. *Leviathan*, chap. 6, p. 24.
68. David Gauthier, *Morals by Agreement* (New York: Oxford University Press, 1986), p. 21.
69. He does not endorse a corresponding thesis for the merely prudential or instrumental norms according to a conception of which we are also capable of acting. But he also holds that there could not be a being that had a rational will in the sense of being able to act according to a conception of a prudential rule or maxim, but did not have a rational will in the sense of being subject to moral norms. So even though not all the rules we acknowledge or act according to

conceptions of are moral rules, we can still be demarcated as the ones who act according to moral rules, for which he does endorse a version of the thesis being discussed.

70. Kant acknowledges his most immediate debt to Rousseau. (It has seemed incongruous to some that a portrait of that wild, intemperate, irregular figure should have provided the sole adornment in the study of the excruciatingly continent and excessively rule-governed Kant.) This tradition is treated as the organizing theme of the Enlightenment in Kant's "Was ist Aufklärung?"
71. Pufendorf, *Law of Nature*, chap. 2, par. 6, p. 27.
72. Ibid., chap. 5, par. 4, p. 89.
73. Ibid., par. 9, p. 95.
74. Ibid., par. 14, p. 107.
75. This is, of course, just as one would expect for an approach that takes its point of departure in construing norms from the example of explicit positive law. "So there are two parts of a law, one defining the offence, and one setting the penalty or the penal sanction; two parts, I say, and not two kinds of laws. For it is idle to say, 'Do this', if nothing follows; and it is equally absurd to say 'You will be punished', if the reason is not added, why punishment is deserved. It must, therefore, be borne in mind that the whole power of a law properly consists in its declaring what our superior wishes us to do or not to do, and what penalty awaits its violators" (*ibid.*). Thus the superior must have "the strength to threaten some evil against those who resist him" (par. 9, p. 95). Besides construing authority in terms of sanctions, Pufendorf also endorses two other central theses considered in the previous section of this chapter. For he takes it that by an obligation "we are bound by the necessity of doing something, for by it some moral bridle, as it were, is slipped over our liberty of action, so that we cannot rightly turn to any other quarter than that to which it directs. An obligation, however, can in no way so bind the will that it cannot, indeed, go contrary to it, although at its own peril" (chap. 5, par. 5, p. 90). Thus his conception of norms treats as essential the possibility of a distinction between what is in fact done and what ought to be done. Perhaps more remarkably, he develops his retributive picture of the practical expression of assessments by endorsing the idea of normatively *internal* sanctions: "an obligation affects the will morally, and fills its very being with such a particular sense, that it is forced of itself to weigh its own actions, and to judge itself worthy of some censure, unless it conforms to a prescribed rule . . . Again, an obligation differs in a special way from coercion, in that, while both ultimately point out some object of terror, the latter only shakes the will with an external force, and impels it to choose some undesired object only by the sense of an impending evil; while an obligation in addition forces a man to acknowledge of himself that the evil, which has been pointed out to the person who deviates from an announced rule, falls upon him *justly*" (*ibid.*, emphasis added).
76. Ibid., par. 12, p. 101.
77. A contemporary version of this view—in particular of the sort of positive freedom (freedom to do new sorts of things, rather than freedom *from* constraint) that results from constraining oneself by specifically *linguistic* norms—is presented in the author's "Freedom and Constraint."
78. *PI*, § 258.

79. Wittgenstein on the Foundations of Mathematics.
80. Ibid., p. 220.
81. McDowell, "Wittgenstein on Following a Rule," pp. 333–334.
82. Most of the discussion of Dennett refers to views propounded already in his "Intentional Systems," *Journal of Philosophy* 68, no. 4 (1971): 87–106; reprinted in *Brainstorms: Philosophical Essays on Mind and Psychology* (Montgomery, Vt.: Bradford Books, 1978); page references are to the reprint edition.
83. Ibid., p. 221.
84. Davidson calls constellations of beliefs and pro-attitudes of this sort "primary reasons" for action, originally in "Actions, Reasons, and Causes," *Journal of Philosophy* 60 (1963), reprinted in *Actions and Events*, (New York: Oxford University Press, 1980), p. 4. This sort of intentional explanation is discussed in more detail in Chapter 4.
85. "Intentional Systems," p. 13.
86. Ibid., p. 17.
87. John Searle, *Intentionality* (Cambridge: Cambridge University Press, 1983).
88. Berkeley: California University Press, 1969.
89. To say that the norms implicit in practices confer conceptual content is to say that having such content just consists in being governed by those proprieties.
90. *PI*, § 54. This catalog might be taken to refer only to what he calls "definite" rules, as also in *ibid.*, § 81.
91. As he does at *ibid.*, §§ 224 and 225.
92. For example, at *ibid.*, § 142.
93. At *ibid.*, § 198, but his better wisdom may be expressed rather in the converse proposition at § 85.
94. *Ibid.*, § 653.
95. For example, at *ibid.*, § 237.
96. *Ibid.*, § 199.

## 2. Toward an Inferential Semantics

1. This choice of terminology follows Hegel's use of *anerkennen* in his *Phenomenology*.
2. Franz Brentano, *Psychology from the Empirical Standpoint*, trans. D. B. Terrell, quoted on pp. 119–120 in *Introduction to the Philosophy of Mind: Readings from Descartes to Strawson*, ed. H. Morick (Glenview, Ill.: Scott, Foresman, 1970).
3. John Searle, *Intentionality* (Cambridge: Cambridge University Press, 1983), p. 2.
4. Robert Stalnaker, *Inquiry* (Cambridge: MIT Press, 1984), p. 2.
5. Searle, *Intentionality*, p. 17.
6. It is because of the distinction indicated by Brentano's reservation that the hyphenated phrase "object-representing" is used here, rather than more committal talk of "representing objects," to mark the categorial contrast with propositional contentfulness.
7. Cited by Roderick Chisholm, on p. 140 in *Introduction to the Philosophy of Mind*, ed. H. Morick (Glenview, Ill.: Scott Foresman, 1970).
8. "The regular connexion [Verknüpfung] between a sign, its sense, and its reference is of such a kind that to the sign there corresponds [entsprechen] a definite

sense and to that in turn a definite reference" ("ÜSB," p. 58). As an abbreviation only, Frege also allows talk of the expression, rather than the sense it expresses, designating or referring to what it represents: "To make short and exact expressions possible, let the following phraseology be established: A proper name (word, sign combination, expression) expresses its sense, stands for or designates its reference. By means of [mit] a sign we express its sense and designate its reference" (p. 61).

9. "Artificial Intelligence as Philosophy and as Psychology," in *Brainstorms: Philosophical Essays on Mind and Psychology* (Montgomery, Vt.: Bradford Books, 1978), p. 122.
10. It may be helpful in clearing up an incipient misunderstanding to remark here that in the official idiom to be developed and employed in this work, linguistic *expressions*, in the sense of marks and noises, do not need to be separately mentioned at this point. For it is not tokens but tokenings that are in the first instance considered as contentful. Sign-designs, the linguistic vehicles of content, are meaningful only at one remove, in virtue of their involvement in linguistic performances that express intentional states and attitudes.

The token/tokening distinction can often be overlooked (so that the theoretical decision as to explanatory priority alluded to here does not even arise) in the case of evanescent tokens such as utterances; the uttering/uttered ambiguity need not be resolved. The issue becomes more evident if one thinks about more permanent tokens, as when the religious enthusiast walks around the city with a sign in the shape of an arrow, inscribed "YOU are a sinner!" and points it at various passersby. In such a case the different tokenings have different contents, even though only one token is involved. The payoff of the policy mentioned here accordingly comes when the use of indexicals and other tokenings that are in principle unrepeatable becomes a topic, in Chapters 7 and 8.

11. *Critique of Pure Reason*, A97.
12. Bertrand Russell, in *The Philosophy of Logical Atomism*, ed. D. Pears (LaSalle, Ill.: Open Court, 1985), is a case in point.
13. *Critique of Pure Reason*, A69/B94.
14. Ibid.
15. Ibid., A68/B93.
16. Ibid., A69/B94.
17. Ibid., A126.
18. Ibid., A79/B104–105. The "transcendental element" introduced in this way is just reference to objects.
19. "Notes for L. Darmstädter," in *PW*, p. 253.
20. "Boole's Logical Calculus and the *Begriffsschrift*," in *PW*, pp. 16–17.
21. The concept of substitution and its significance in such a decompositional semantic program are investigated in detail in Chapter 6.
22. *GL*, sec. 60. The claim that "only in the context of a proposition [*Satz*] does a name have any meaning" is enunciated also in the Introduction (p. x), and in secs. 46 and 62.
23. *PW*, p. 232.
24. "ÜSB," p. 63.
25. Ibid., pp. 57 and 58.
26. Ibid., pp. 62–63.

27. *PW*, p. 144.
28. "My Basic Logical Insights," *PW*, p. 252. See also the opening pages of the "Logic" of 1897, beginning at *PW*, p. 129.
29. This is what Kant is getting at in seeing the "transcendental element" of referring to objects as introduced into representations by their role in judgment, in the passage quoted above.
30. Desiring a mouse or desiring relief from hunger are best thought of as elliptically specified desires that one have or eat a mouse, that one's hunger be relieved. This should become apparent in the initial discussion of practical reasoning in Chapter 4. In any case, it is sufficient for the point being made here that the contents of the corresponding beliefs must be specified by sentential clauses.
31. A clue that is exploited in the account of this relation endorsed further along can be gleaned from looking at how to make explicit what a theorist becomes committed to in taking one complex object (for instance, a map) to be a representation of another (for instance, terrain). The theorist is claiming that from a certain kind of fact about the representing object (corresponding to a privileged vocabulary for describing it), it is possible to *infer* a certain kind of fact about the represented object. Thus from the fact that the blue squiggly line passes between a round dot and a square one, it is possible to infer that there is a river between a city whose population is less than 100,000 and one whose population is greater than 100,000. (This is not to say that when the representational relation is acknowledged only implicitly in the practice of someone using a complex object as a representation of another, the practitioner must be able to state explicitly the premises and conclusions of these inferences. See below at Chapter 8, Section II, Subsection 4 [such cross-references are abbreviated hereafter as 8.2.4].)
32. It is worth pointing out that this is not a difficulty that automatically confronts any theory that invokes a special ontological category of propositions in its account of claiming, judging, and believing. Obviously such accounts can accommodate the special status of propositional contents. A theory such as Stalnaker's, which understands propositions as sets of possible worlds and construes the attribution of propositionally contentful intentional states in terms of the use of the structure of possible worlds to *measure* those states for the purpose of explaining actions, is not vulnerable to the charge of semantic nominalism, of being in thrall to the model of designation. Such theories need involve no inappropriate assimilation of propositionally contentful states, attitudes, and performances to representings thought of as naming what they represent. They can respect the primacy of the propositional. They can do so precisely because they begin with the idea of an utterance expressing a proposition or a state exhibiting a propositional content. The question that then arises is what *expressing* a proposition has to do with *representing* anything. An account is required in any case of the relation between propositional contentfulness and object-representing contentfulness (purporting to represent objects). But only obfuscation results from talking in addition of sentences not only as expressing propositions and beliefs as having propositional contents but also of their *representing* propositions.
33. Although the point is put here in terms of cognition, a parallel point can be made on the side of rational action. For Kant understands the rational will as a faculty

- of causally determining particular acts through the conception of a general rule (*Critique of Practical Reason*, sec. 7).
34. It is abstracted by a comparative analysis, the forerunner of Frege's substitutional or functional method of analysis of the conceptual contents of judgments, which is the concern of Chapter 6 of this work.
  35. *Phenomenology*, par. 109. The erotic theory of classificatory consciousness arises in the order of exposition of the *Phenomenology* as the introduction to the theory of self-consciousness. Heidegger's successor concept of understanding in terms of taking something as something in practice is discussed in the author's "Heidegger's Categories in Being and Time," *Monist* 66, no. 3 (July 1983): 387–409.
  36. To make this point is not to claim that Hegel's erotic model does not have more resources (for instance for funding a distinction between correct and incorrect taking of something to be food) than are made available in the inorganic case. Consideration of inanimate objects suffices for the contrast of interest here, however.
  37. P. 262 of "Inference and Meaning," reprinted in *PPPW*.
  38. This brief sketch can no more than gesture at the rich development of these ideas in the *Phenomenology*. A fuller discussion of this important chapter in the tradition inherited by the approach pursued here lies outside the scope of this work. (It will be pursued on another occasion.) The few cryptic characterizations offered here are intended to serve only as placeholders, whose significance will become somewhat clearer as the way in which material contents can be construed in terms of inference and incompatibility and expressed by means of logical vocabulary are filled in as this chapter and the rest of the work proceed.
  39. In the autobiographical sketch in *Action, Knowledge, and Reality*, ed. H.-N. Castañeda (Indianapolis: Bobbs-Merrill, 1975), p. 285.
  40. A detailed accounts of their efforts, understood along these lines, is offered in the author's "Leibniz and Degrees of Perception," *Journal of the History of Philosophy* 19, no. 4 (October 1981): 447–479; and "Adequacy and the Individuation of Ideas in Spinoza's Ethics," *Journal of the History of Philosophy* 14 (April 1976): 147–162.
  41. *BGS*, sec. 3. Frege's word *richtig* here is usually misleadingly translated as 'valid'. The discussion below of the relation between materially and formally good inferences is intended to explain why 'correct' is a better translation here.
  42. *Ibid.*, sec. 2.
  43. "Boole's Logical Calculus and the *Begriffsschrift*," *PW*, pp. 16–17.
  44. *BGS*, sec. 3.
  45. *Logical Syntax of Language* (London: Routledge, Kegan, Paul, 1964), p. 175. Sellars's reference is in "Inference and Meaning," *PPPW*, p. 266.
  46. *FPL*, p. 432.
  47. *Ibid.*, p. 433.
  48. As will become clear, the idiom of material *inference* is not to be understood in relation to the use of the so-called material *conditional*.
  49. "Inference and Meaning," in *PPPW*, p. 261.
  50. Reprinted in *Brainstorms* (Montgomery, Vt.: Bradford Books, 1978), pp. 10–11.
  51. "Intentional Systems," p. 11.
  52. *Ibid.*

53. "Inference and Meaning," in *PPPW*, p. 265.
54. Ibid., p. 284. This talk about the "framework" of logical transformation rules is just one expression of the attitude toward the relation between formal and material inference considered here. It would not be underwritten by the approach endorsed below, where logical vocabulary is picked out by its expressive role and then used to derive a notion of formal validity from material correctnesses of inference.
55. Ibid., pp. 270–271.
56. Ibid., p. 273.
57. Ibid., p. 274.
58. It should be noticed that the point being made here has nothing to do with the relation in mathematical logic between proof theory and model theory. In particular, it is *not* being claimed that one need be concerned only with the former, to the exclusion of the latter. The concepts of arithmetic cannot be fully specified by finitely stateable rules of inference. Nevertheless, we do grasp those concepts. But this is just to say that we do in fact understand their inferential significance. To make explicit the inferences that articulate the concepts of arithmetic, we must employ model-theoretic metalanguages. This fact in no way impugns the inferential conception of conceptual content; it merely shows that traditional proof-theoretic metalanguages are not sufficiently expressively powerful to make such inferential roles explicit. The additional (inferential) expressive power added by metalanguages that employ the traditional semantic vocabulary of truth, denotation, and satisfaction is discussed in Part 2 below.
59. "Language, Rules, and Behavior," in *PPPW*, p. 136 n. 2.
60. From "Concepts As Involving Laws, and Inconceivable without Them," in *PPPW*, p. 122. The remark of A. J. Ayer's referred to is from p. 17 of *Language, Truth, and Logic* (New York: Dover Publications, 1952).
61. "Boole's Logical Calculus and the *Begriffsschrift*," in *PW*, pp. 12–13.
62. Ibid., p. 13.
63. Ibid., p. 46.
64. BGS, Preface, in *From Frege to Gödel*, ed. Jean van Heijenoort (Cambridge: Harvard University Press, 1967), p. 7.
65. Ibid.
66. "Boole's Logical Calculus and the *Begriffsschrift*," in *PW*, p. 16.
67. This is a reason to reject the quasi-Tractarian view according to which nothing can count as claiming or asserting (and so nothing can count as inferring) unless the repertoire already contains logical vocabulary, so that the simplest claiming (the making explicit of anything) already presupposes the whole of logic.
68. See n. 28 above.
69. It will emerge in Chapter 3 that *entitlement-preserving* inferences are also important. They correspond roughly to inductive inferences in the same way that commitment-preserving ones correspond to deductive inferences.
70. In his fragment "Logic," Frege seems to endorse this order of explanation. He says: "To make a judgment because we are cognisant of other truths as providing a justification for it is *inferring*. There are laws governing this kind of justification, and to set up these laws of correct [*richtigen*] inference is the goal of logic . . . It would not perhaps be beside the mark to say that the laws of logic are nothing other than an unfolding of the content of the word 'true'" (*PW*, p. 3).

71. Only the sentential logical connectives are being addressed here. Identity and quantification, which raise special formal and philosophical difficulties, are discussed in later chapters.
  72. 'Extensional' in this context can be made sense of in purely substitutional terms, without having to appeal to the sort of representational concepts in terms of which it is usually explicated (see 6.2 below).
  73. The original investigation is in the author's "Varieties of Understanding," in *Reason and Rationality in Natural Science*, ed. N. Rescher (Lanham, Md.: University Press of America, 1985), pp. 27–51. The treatment there is cleaned up, corrected, and substantially extended in Mark Lance's "Normative Inferential Vocabulary: The Explication of Social Linguistic Practice" (Ph.D. diss., University of Pittsburgh, 1988), where the relevant completeness results are proven. The most interesting logical systems result from a semantics that combines pragmatically conferred incompatibility relations with pragmatically conferred entailment relations.
  74. The actual procedure defines the introduction of a connective only as the principal connective in a formula and defines how to eliminate only principal occurrences. Full generality is nonetheless assured by working recursively. It should be remarked that according to the approach developed here, the standard Gentzen-style definitions for logical connectives are still possible for conjunction and disjunction, but the expressive role of conditionals, negation, and many other bits of logical vocabulary requires that they be understood as having quite another sort of introduction rule.
75. *FPL*, p. 453.
76. Noninferential reports are discussed in more detail in Chapter 4.
  77. The *empirical* contribution to conceptual content made by noninferential *circumstances* of application in *perception* and the *practical* contribution to conceptual content made by noninferential *consequences* of application in *action* are discussed in Chapter 4.
  78. The significance of this sort of example is explored in the author's "Truth and Assertibility," *Journal of Philosophy* #73, no. 6 (March 1976): 137–149. Ingredient contents are discussed below at 6.1.2.
  79. *FPL*, p. 455; the following passage is on pp. 453–454.
  80. *Ibid.*, pp. 456–457.
  81. *Ibid.*, p. 455.
  82. *Ibid.*, p. 454. It should be noted that inferential conservativeness is a weaker condition than derivability of circumstances from consequences (or vice versa). Showing how to derive one aspect from the other, using logic or prior inferential commitments, is sufficient but not necessary for conservativeness. I am grateful to Michael Kremer for this point.
  83. N. Belnap, "Tonk, Plonk, and Plink," *Analysis* 22 (1962): 130–134, commenting on A. N. Prior's "Runabout Inference Ticket," *Analysis* 21 (1960–1961): 38–39.
  84. *FPL*, p. 454.
  85. Jonathan Bennett suggested this illustrative anecdote.
  86. *FPL*, p. 455n.
  87. *Ibid.*, p. 358.
  88. In Quine's *From a Logical Point of View* (Cambridge: Harvard University Press, 1953), pp. 20–46.

89. This discussion addresses only versions of the project of offering truth conditions that envisage employing *other* concepts than those expressed by the words appearing in the sentences for which one is offering truth conditions, so as to offer *substantive* explications of those concepts. If truth conditions are conceived *modestly*, so that one is allowed to specify the truth condition for the sentence “Luther was a Schwabian” as simply as that Luther was a Schwabian, then the consequences pointed to in the text do not arise.
90. It should be acknowledged that although the discussion of this chapter has been framed throughout in terms of a stark opposition between two complementary orders of explanation—the representationalist and the inferentialist—these alternatives are not exhaustive. Other possibilities include treating neither representation nor inference as explanatorily prior to the other. One might then go on to explain both in terms of some third notion, which is treated as more fundamental. Or one might eschew reductive explanations in semantics entirely and remain contented with describing the relations among a family of mutually presupposing concepts—a family that includes representation, inference, claiming, referring, and so on.
91. Recall from the discussion in 1.4 above that the most serious objection McDowell levies against the social-practice theories of norms put into Wittgenstein’s mouth by Wright and Kripke is that they have no room for the idea of proprieties of concept use that the whole community could be wrong about. As Wright puts it, these theories jettison the intuitive “ratification independence” of concept use for the special case where the ratifying attitudes of taking particular candidate applications of concepts to be correct or incorrect are those of the community as a whole.
92. The obligations involved in this order of explanation have just been indicated. It was suggested above that the corresponding explanatory demands on the contrary directions of explanation pursued by the intellectualist about norms, the formalist about logic, and the representationalist about content are difficult to meet. The intellectualist about norms has trouble explaining the norms governing the use or application of rules, principles, claims, and concepts. The formalist about logic has trouble explaining nonlogical content. The representationalist has trouble explaining specifically propositional content and its grasp. Of course there are various strategies for meeting or evading these demands. The present assembling of reminders and considerations intends only to sketch an alternative; it does not pretend to offer all-purpose refutations of the various contrary explanatory strategies that might be adopted.
93. This is an important point for Kant as well. His terminology in the *Logic*, where one-premise inferences are called “immediate” and multipremise inferences are called “mediated,” greatly influences Hegel’s use of those central technical terms of the *Phenomenology*.

### 3. Linguistic Practice and Discursive Commitment

1. The same structure is exhibited even if sentences are not taken as the primitive interpreted expressions. For example, if the basic stipulated assignment is of objects to singular terms, and sets of objects to predicates (or sets of sequences of objects, for multiplace predicates), then the results of simple syntactic predi-

- cations may be assigned truth-values as derived interpretants accordingly as the objects (or sequences) corresponding to the term(s) appearing in the predication are or are not included in the sets corresponding to the predicates.
2. *FPL*, p. 413. The view Dummett expounds in this passage differs from the one to be developed in this work in that he is concerned only with the contentfulness of linguistic expressions, not with that of intentional states and attitudes more generally; he considers only truth-conditional semantic interpretation; and he does not make clear the essentially normative character of the linguistic practices that constitute the use or working of the language.
  3. A paradigm of the recognition of this promissory note implicit in, for example, possible-worlds semantics is David Lewis's "Languages and Language," in *Language, Mind, and Knowledge*, Minnesota Studies in the Philosophy of Science, vol. 7, ed. Keith Gunderson (Minneapolis: University of Minnesota Press, 1975), pp. 3–35; reprinted in Lewis's *Philosophical Papers*, vol. 1 (New York: Oxford University Press, 1983), chap. 11. Lewis's account of what it is for semantic interpretants to be appropriately associated with expressions by the use of language turns on his notion of *convention*, which appeals to propositionally contentful *intentions* and beliefs and so is not suitable to be extended to an account of the pragmatics corresponding to the contentfulness of such intentional states. In *Inquiry* (Cambridge: MIT Press, 1984), Robert Stalnaker provides an account of what it is for sets of possible worlds to be associated as the propositional contents of intentional states such as belief, appealing only to the possibility of intentional interpretation of intelligent behavior.
  4. H. P. Grice, "Meaning," *Philosophical Review* 66 (1957): 377–388. Also see Stephen Schiffer, *Meaning* (New York: Oxford University Press, 1972).
  5. David Lewis, *Convention: A Philosophical Study* (Cambridge: Harvard University Press, 1969), and "Languages and Language"; Jonathan Bennett, *Linguistic Behavior* (Cambridge: Cambridge University Press, 1976); John Searle, *Speech Acts* (Cambridge: Cambridge University Press, 1969); and *Expression and Meaning* (Cambridge: Cambridge University Press, 1979).
  6. Davidson takes this model of communication to entail that sharing a language is a merely practical, hypothetical necessity—it is convenient for members of a linguistic community to use the same noises to express the same thoughts because it minimizes the need for explicit theorizing about the intentions with which the noises are produced. The expectations or customs of other speakers, however, have no authority over how any individual is correctly understood. What matters is how the speaker intends to be understood or interpreted. Davidson's subtle position is different from the others mentioned in this connection in important ways, however, as he does not take it that the contents of these communicative intentions can be made sense of antecedently, in abstraction from interlocutors' interpretation of one another. Put another way, the intention to be interpreted one way rather than another that Davidson rightly takes to be essential to the meaningfulness of ordinary discourse can be understood to be implicit rather than propositionally explicit, an intention-in-action rather than a separately individuable prior intention. That is, its involving such an intention can be conceived as an automatic compliment paid to a performance in virtue of the fact that were the issue to be raised and the speaker sincerely to disavow the intention to be understood in a certain way, that would be evidence

that it had been misinterpreted. Such a view is importantly different from the Gricean picture of meanings as imposed on utterances by the antecedently contentful intentions of speakers, although Davidson is not always careful to register the distinction.

7. Jay Rosenberg, *Linguistic Representation* (Dordrecht: D. Reidel, 1975), chap. 2.
8. This thought leads John Searle (*Intentionality* [Cambridge: Cambridge University Press, 1983]) to insist that the intentionality of intentional states, unlike that of linguistic expressions, must be *intrinsic*. That is, Searle endorses the view, characteristic of agent semantics, that "I impose Intentionality on my utterances by intentionally conferring on them certain conditions of satisfaction which are the conditions of satisfaction of certain psychological states." (p. 28). Accordingly, he is committed to a strong distinction between the contentfulness of utterances and that of the intentional states they express: "Since sentences—the sounds that come out of one's mouth or the marks one makes on paper—are, considered in one way, just objects in the world like any other objects, their capacity to represent is not intrinsic, but is derived from the Intentionality of the mind. The Intentionality of mental states, on the other hand, is not derived from some more prior forms of Intentionality but is intrinsic to the states themselves. An agent uses a sentence to make a statement or ask a question, but he does not in that way use his beliefs and desires, he simply has them" (pp. vii–viii). Thus, "That the belief has those conditions of satisfaction is not something imposed on the belief by its being used at all. A belief is intrinsically a representation in this sense: it simply consists in an Intentional content and a psychological mode . . . [It is false that] in order for there to be a representation there must be some agent who uses some entity as a representation. This is true of pictures and sentences, i.e. of derived Intentionality, but not of Intentional states" (p. 22). The doctrine that the intentionality or contentfulness of intentional states and attitudes is *intrinsic* is in some ways the correlate in the domain of philosophical semantics of the method of *stipulation* in formal semantics. The theorist takes it that intentional states are just the sort of thing that comes with an intentional (paradigmatically propositional or representational) content. This is Descartes's strategy—the mental is distinguished as being naturally *about* other things (it is the sort of thing other things have *objective* reality in). The theorist can talk about the consequences of such contentfulness but cannot be expected to have anything substantive to say about what that contentfulness consists in, apart from those consequences. Thus Searle denies that it is possible to give an analysis of intentionality, taking it rather that "Intentionality is, so to speak, a ground floor property of the mind" (pp. 14–15).
9. Robert Stalnaker, *Inquiry* (Cambridge: MIT Press, 1984), p. 4.
10. *Ibid.*, p. 5.
11. *Ibid.*, p. 6.
12. In spite of the acknowledgment of these alternatives implicit in his disjunctive formulation, however, Stalnaker restricts his consideration (and his arguments against the linguistic approach) to resemblance theories. Thus he concludes his description of the linguistic approach: "It is not essential to the linguistic picture that every thinking creature be capable of outward speech or that every one of our thoughts be expressible in our public language. All that is essential

is that thought be explained by analogy with speech" (*ibid.*, p. 5). This is true of resemblance theories—but not relation theories—of the significance of linguistic practice to intentional states. His further subdivision of the linguistic picture is phrased so as to accord with the restriction of consideration to analogical theories, in which thought is understood on the model of speech. "The development of the linguistic picture leads in two quite different directions which emphasize different analogies between speech and thought. One hypothesizes a language of thought, which may be different from any language used for communication; the other argues for the dependence of thought on the social activities of speech" (pp. 5–6). But in fact thought or the possession of contentful intentional states might be taken to depend on the social activities of speech *either* because the contents of intentional states must be modeled on and so understood in terms of the contents expressed by public speech acts in general or because it is essential to a state's having such content that the state can issue in a speech act by which it is publicly expressed—that is, either according to a linguistic theory of intentionality structured by resemblance or according to one structured by relation.

13. "Thought and Talk," in *Inquiries into Truth and Interpretation* (New York: Oxford University Press, 1984), p. 156.
14. This and the subsequent passages are from *ibid.*, p. 170.
15. *FPL*, p. 362.
16. Hartry Field, "Mental Representation," *Erkenntnis* 13 (1978): 9–61. This strategy was put forward by Sellars already in 1953, in his "Semantical Solution to the Mind-Body Problem" (in *PPPW*, pp. 219–256; see especially secs. IV–VI), and remained at the center of his approach thereafter.
17. *Beurtheilbarer Inhalt*, introduced in the *BGS*, sec. 2.
18. This is not to say that in the full-fledged language game any particular move that is taken to be correct by the practitioners—even *all* the practitioners—thereby counts as correct. (That this promissory note is eventually redeemed is demonstrated by the objectivity proofs presented below in 8.6.5.) One of the primary tasks of this work is to begin to explain the way in which the linguistic community can institute incompatibilities relating subpractices, which then constrain its own assessments where those subpractices interact—that is, the way in which noninferential reporting practices and deliberate actions performed as a result of practical reasoning conspire to confer *objective* empirical content on the concepts they are inferentially linked to—the way in which what is said can come to answer for its correctness not to the ones using the language but to what they use it to talk *about*.
19. (Formal) inconsistency is to (material) incompatibility as formal logical validity of inference is to material correctness of inference. In each case the former should be defined in terms of the latter, for the reasons discussed above in 2.4.4. (The official interpretation of the 'scare quotes' employed here is presented below in 8.4.5.)
20. In order to get a more realistic model, a shell might be added around this practice. A man, for example, who routinely fails to fulfill promises might be held responsible not only for failing to recognize himself as having undertaken a commitment (not recognizing his entitlement to entitle others to rely on him) but also for abusing the community practice of promising by his repeated attempts to claim or pretend to an authority that he has then withheld. Holding

such a man responsible in this way might consist in shunning him or in beating him with sticks, as in the original case. It is possible, in other words, to combine a general, externally defined sanction, with a specific, internally defined one. The former becomes a shell around the latter. The general sanction might be an instance of holding responsible for any abuse of linguistic authority, which might eventually lead to expulsion from the linguistic community (being treated like a parrot or a pariah). (This addition was suggested by Michael Kremer.)

21. The language of scorekeeping is suggested by David Lewis's "Scorekeeping in a Language Game," reprinted as Chapter 13 of his *Philosophical Papers*, pp. 233–249. More is made of this notion below, in Section IV.
22. The responsibility characteristic of action and the authority characteristic of perception are discussed in Chapter 4.
23. Mark Lance noticed that this definition permits one to consider *asymmetric* incompatibility relations as well as symmetric ones. He exploits this possibility formally in his development (in "Normative Inferential Vocabulary: The Explication of Social Linguistic Practice" [Ph.D. diss., University of Pittsburgh, 1988]) of the sort of incompatibility semantics originally suggested in the author's "Varieties of Understanding," in *Reason and Rationality in Natural Science*, ed. N. Rescher (Lanham, Md.: University Press of America, 1985), pp. 27–51. *Inconsistency* is the formal correlate of incompatibility. It is a logical notion, to be understood in terms of *negation*. But what makes a bit of vocabulary express negation is itself to be understood in terms of its relation to material incompatibility. The negation of a claim is defined as its minimal incompatible, the inferentially weakest claim that is entailed (in the commitment-preserving sense) by everything that is incompatible with the original claim.
24. As the examples discussed in the previous section indicate, it is not impossible to have authority without responsibility, or responsibility without authority. As students of organizational behavior will attest, however, it is a basic principle of social engineering that the stability and effectiveness of a practice are undercut if the authority accorded to some practitioners outruns their corresponding responsibilities, or vice versa. Linguistic practice as here construed is well designed in this respect.
25. Kurt Baier, "Responsibility and Freedom," in *Ethics and Society*, ed. R. T. DeGeorge (New York: Anchor Books, 1966), pp. 49–84.
26. John Searle, *Speech Acts*, p. 96.
27. Justificatory practices depend on entitlement-preserving inferences. But commitment-preserving inferences are also entitlement-preserving (though not conversely). If anyone who is committed to *p* is thereby committed to *q*, the only case in which entitlement to *p* plausibly would not carry with it entitlement to *q* is one in which the interlocutor is precluded from entitlement to *q* by concomitant commitment to something incompatible with it. But if *p* commitment-entails *q*, anything incompatible with *q* is incompatible with *p*, so under the circumstances described, the interlocutor could not be entitled to *p*.
28. The justificatory mode of entitlement inheritance requires that one invoke claims with *different* contents, for otherwise the 'stuttering' inference, from *p* to *p*, would count as a justification of *p*.
29. This is, as will become clear in terms of the model, compatible with the possibility in the fully developed practice of an interpreter correctly taking the entire community to be wrong about what commitments they are entitled to—but

such a judgment will always be that they are in some sense wrong by their own lights, that is, wrong given how they have committed themselves to its being proper to settle such questions and assess the answers.

30. Lewis, "Scorekeeping in a Language Game," p. 236.
31. Ibid.
32. Ibid., p. 237.
33. Ibid., p. 238.
34. Chapter 9 (Sections II and III) discusses the essential role played in such objective constraints by the fact that the interpreter who attributes discursive deontic scorekeeping practices to a community can use nonscorekeeping vocabulary with an antecedent use in specifying those practices.
35. Lewis, "Scorekeeping in a Language Game," p. 239.
36. *FPL*, p. 361.
37. Ibid.
38. "To begin with" because of the contribution made to semantic contents by the role sentences play in noninferentially elicited (but inferentially articulated) observation reports and in their role in giving rise to actions.
39. It is assumed throughout (though this requirement could be relaxed) that incompatibility is a symmetric relation. Also, if it is assumed that commitment-preserving inferences are entitlement preserving, in the absence of incompatible defeasors it follows that if everything incompatible with  $q$  is incompatible with  $p$ , then the inference from  $p$  to  $q$  is good both committively and permissively.
40. The correctness of such an inference according to  $A$  depends not only on the commitments and entitlements to commitments that  $A$  attributes to  $B$  but also on the commitments that  $A$  undertakes. When the expressive resources are available for explicit challenges to reliability inferences, these background premises of  $A$  will be cited in justification of those inferences and their conclusions. Under those circumstances they will be cited as constituting *standard conditions* for  $A$ 's observational authority with respect to this sort of content. This point is adverted to in the text further along.
41. Such a mapping directly characterizes what corresponds to the appropriate consequences of application of the expression. The circumstances of appropriate assertion (according to the scorekeeper in question) can be recovered from the full mapping, however, as the subset of initial scores in which the consequences of assertion include attribution, not only of commitment, but of entitlement to the assertion (according to the scorekeeper in question). The way in which entitlement is attributed to noninferential reports (thereby treating the reporter as reliable) shows that this class of appropriate circumstances of assertion is wider than what would result from taking the assertion to be in order only when the prior score already included an attribution of entitlement.
42. Philip Kremer and Mark Lance present some fascinating results concerning the explicit codification of commitment consequences in the form of logical conditionals in "The Logical Structure of Linguistic Commitment, I: Four Systems of Non-Relevant Commitment Entailment" and "The Logical Structure of Linguistic Commitment, II: Relevant Commitment Entailment," both forthcoming in the *Journal of Philosophical Logic*.
43. If one is entitled to  $p$  and  $p$  commitment-entails  $q$ , one is entitled to  $q$ —any entitlement-defeating incompatibilities to  $q$  equally defeat entitlement to  $p$ .

44. One can be (taken to be) entitled to claims one is not (taken to be) committed to—these are conclusions one is entitled to draw but has not yet committed oneself to. In this way one may be entitled to each of two mutually incompatible claims, so long as neither has been endorsed and commitment to it undertaken. So one might have good inductive reasons for believing that the barn is on fire (smoke, the particular noises that would usually accompany it, and so on) and a different set of good inductive reasons for believing it is not on fire (the alarm has not rung, it is pouring rain, the barn was just inspected, and so on). Either conclusion by itself could be defended, though one would cease to be entitled to it if already committed to the conclusion of the other argument. Attribution of entitlement to the consequences of conjoining incompatible contents to which one is severally entitled (because not committed to either) is avoided by closing entitlements without commitment only under commitive inferences (if not defeated by incompatibilities) and closing only entitlements to commitments actually undertaken (according to the one keeping score) under permissive inferences.
45. Under suitable conditions. Considerations having to do with the possibility of coercion, insincerity, shyness, and so on are systematically suppressed in presenting the model of assertional practice, on the grounds that they are intelligible only against a background of propositional contents conferred by the sorts of interactions considered here. Real-world phenomena such as these (which presumably will be present at every stage in the development of actual practice) create play between the proprieties of practice (especially scorekeeping practice) an interpreter takes a community to be bound by and their actual behavior. The status of this discrepancy—a normative generalization of the competence-performance distinction—is discussed in Chapter 9.
46. Daniel Dennett, "Intentional Systems," in *Brainstorms* (Montgomery, Vt.: Bradford Books, 1978), p. 19.
47. *Ibid.*, p. 20.
48. *Ibid.*
49. There are subtleties that require qualifying this formula, some of which are discussed below in 8.5.2.

#### 4. Perception and Action

1. The further conditions that have been suggested in response to examples of the sort Gettier first presented are not discussed here. Those impressed by the significance of these counterexamples to the sufficiency of JTB analyses of "S knows that *p*" may want to treat the term 'knowledge' as it appears in the rational reconstruction presented here as really invoking knowledge\* (which is defined simply as justified true belief). The considerations motivating many of the proposals for a fourth condition on knowing can be straightforwardly transposed into the idiom of this work, so those who believe that the status reconstructed here must be further specified in order to deserve the central role it is given might try the experiment of seeing how their favorite candidate looks in deontic scorekeeping guise. The basic elements of the social practice model's construal of the statuses corresponding to justification, truth, and belief do not turn on the kind of niceties concerning their interaction that attempts to for-

mulate a fourth condition must address. It is explained below why knowledge\* deserves to be accorded a fundamental explanatory role, regardless of its relation to what is expressed by the English word ‘know’.

2. Of course, there is no necessity to adopt either direction of explanation; it may be that neither term is intelligible apart from its relation to the other.
3. This claim about the conceptually basic level of practices is compatible with the institution of a distinction in more sophisticated practices between *claiming* that *p* and claiming to *know* that *p*. This point is discussed further along.
4. The distinction being made here between the present approach and standard ones is different from, though intimately related to, the fundamental difference between understanding belief, justification, truth, and so knowledge as kinds of *normative status* rather than as kinds of *natural state*, so that one looks for *proprieties*, rather than *properties*, corresponding to them. The connection between the issues consists in the difficulty of appreciating the significance of the social distinction of attitude between acknowledging and attributing unless one is already thinking of what these are attitudes toward in normative terms, as commitments and entitlements—that is, deontic statuses, not descriptive states.
5. Being a *logical* being—having the expressive resources to make propositionally explicit crucial semantic and pragmatic features of the social practices in virtue of which one is a linguistic, rational, cognitive being—is a further, optional stage of development, which presupposes this fundamental one.
6. Insofar as it addresses sapience rather than sentience—that is, insofar as it is concerned with intentionality in the sense defined by possession of *propositional* attitudes.
7. One source of these views is the “thermometer view of knowledge” put forward by D. Armstrong in *Belief, Truth, and Knowledge* (Cambridge: Cambridge University Press, 1973). F. Dretske defends such a view in *Knowledge and the Flow of Information* (Cambridge: MIT Press, 1981), and more recently in “The Need to Know,” in *Knowledge and Skepticism*, ed. M. Clay and K. Lehrer (Boulder, Colo.: Westview Press, 1989). A. Goldman’s development of a reliabilist theory is expounded in *Epistemology and Cognition* (Cambridge: Harvard University Press, 1986), and again in “Précis and Update of Epistemology and Cognition,” in *Knowledge and Skepticism*, ed. Clay and Lehrer. R. Nozick puts forward a version of reliabilism in *Philosophical Explanations* (Cambridge: Harvard University Press, Belknap Press, 1981). M. Swain has a version that takes reasons more seriously than the others, in *Reasons and Knowledge* (Ithaca: Cornell University Press, 1981), and more recently in “Justification, Reasons, and Reliability,” *Synthese* 64, no. 1 (1985): 69–92. This list is intended only to provide a reasonably representative sample.
8. This sort of talk involves a promissory note—only in Chapter 6 is the official account developed to the point that it explains (in terms of substitution inferences) the relation between applying a *predicate* and making a *claim*.
9. Alvin Goldman, “Discrimination and Perceptual Knowledge,” *Journal of Philosophy* 73, no. 20 (1976).
10. J. L. Austin “Other Minds,” in his *Philosophical Papers* (New York: Oxford University Press, 1961).
11. Not objective “to begin with” because, as was acknowledged already in Chapter 1, it is a critical criterion of adequacy of any account of the use of empirical

concepts that it be able to explain how *in the end* objective proprieties governing that use can come into play—how claims can be understood as true or false regardless of whether anyone or everyone takes them to be so, depending rather on how things are with what the claims are about. Not until Chapter 8 will assembly be complete of the raw materials necessary to explain what is involved in this sense of objectivity. One important clue has been put on the table already, however. For it has been pointed out how a distinction can arise between what someone to whom a commitment is attributed is *justified* or *entitled* to believe, and what is in fact *true*. This is just the social-perspectival distinction of attitude between *attributing* deontic statuses and *undertaking* them. It is in terms of this fundamental *social* articulation of deontic attitudes that the possession by claims and concepts of *objective* representational content is eventually to be understood.

12. "EPM," sec. 32.
13. *Ibid.*, sec. 34.
14. *Ibid.*, sec. 35.
15. *Ibid.*, sec. 32.
16. *Ibid.*, sec. 35.
17. *Ibid.*
18. *Ibid.*, sec. 36.
19. This possibility was floated in the author's original discussion in "Asserting."
20. It is a commonplace among teachers of mathematics that students often profess to be completely unable to deal with problems of a certain sort long after they are in fact able to solve them reliably.
21. Assuming that the attributor of knowledge both attributes to Monique commitment to the claim and undertakes such commitment (that is, endorses the claim and so takes it to be true).
22. Sellars, "Some Reflections on Language Games," in Sellars, *Science, Perception, and Reality* (New York: Routledge and Kegan Paul, 1963).
23. "EPM," sec. 32.
24. "There are two ways in which a sentence token can have credibility: (1) The authority may accrue to it, so to speak, from above, that is, as being a token of a sentence type all the tokens of which, in a certain use, have credibility, e.g. ' $2 + 2 = 4$ '. In that case, let us say that token credibility is inherited from type authority. (2) The credibility may accrue to it from the fact that it came to exist in a certain way in a certain set of circumstances, e.g. 'This is red'. Here token credibility is not derived from type credibility" (*ibid.*).
25. "Epistemology Naturalized," in *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969).
26. As they are called in the author's "Asserting," p. 643. Cf. Hegel's remark in a related context that "one barren assurance is of just as much worth as another," in the Introduction to the *Phenomenology* (par. 76).
27. Recall from the earlier discussion that Stalnaker introduces the linguistic approach in a way that leaves room for relational linguistic accounts such as Davidson's and the one pursued here, but he then discusses and argues against only the reductive versions.
28. Though just how this latter possibility should be understood is not officially addressed until Chapter 8.

29. As will become clear, it would be another sort of reductive mistake to *identify* rational agency with preference-maximizing.
30. David Lewis, *Convention: A Philosophical Study* (Cambridge: Harvard University Press, 1969), also "Languages and Language," in *Language, Mind, and Knowledge*, Minnesota Studies in the Philosophy of Science, vol. 7, ed. Keith Gunderson (Minneapolis: University of Minnesota Press, 1975).
31. "Communication and Convention," in *Inquiries into Truth and Interpretation* (New York: Oxford University Press, 1984), p. 280.
32. *FPL*, pp. 298ff., 354–356.
33. "Some Reflections on Language Games."
34. Though as will emerge below, in some cases the nonlinguistic intentional performance just *is* (has the scorekeeping significance of) the acknowledgment of a practical commitment, rather than being a *response* to such an acknowledgment. The difference corresponds to that between intentions-in-action and prior intentions.
35. Section V below discusses some different patterns of availability of practical reasons across interlocutors.
36. Here 'theoretical' is opposed to 'practical', as pertaining to relations exclusively between doxastic discursive commitments. This use ought not to be confused with the sense of 'theoretical' that is opposed to 'observational', within the doxastic sphere. In the latter usage, following Sellars's practice (in "Empiricism and the Philosophy of Mind" and elsewhere), theoretical claims are distinguished as those one cannot become entitled to noninferentially, by the exercise of reliable differential responsive dispositions to acknowledge doxastic commitments. Theoretical vocabulary is then distinguished as that which appears only in claims that are theoretical in this sense. It is this usage that stands behind Sellars's claim that the distinction between the observable and the theoretical is not ontological but only methodological. Neptune was a theoretical entity so long as claims about it could be arrived at only inferentially, as based on its perturbation of the observable orbits of other planets. It became observable, however, once we built telescopes powerful enough to make it subject to noninferential reporting. Something is theoretical or observable in this sense only relative to our practices; nothing is "intrinsically" theoretical.
37. Davidson wants to analyze intention in terms of reasons. A representative formulation is "Someone who acts with a certain intention acts for a reason" ("Intending," reprinted as Chapter 5 of *Actions and Events* [New York: Oxford University Press, 1980], p. 84). Irrational actions accordingly pose a problem for him (which he addresses in "How Is Weakness of the Will Possible?" reprinted as Chapter 2 of *Actions and Events*). According to the present view, he is mistaking a global condition on intention for a local condition, as a result of failing to distinguish commitment from entitlement.
38. This formulation is intended to encompass both committive and permissive inferences. The latter can be thought of as conveying *entitlement* to the commitments that are their premises to *entitlement* to the commitments that are their conclusions. In scorekeeping terms, for an attributor to endorse such a permissive practical inference is to take it that anyone who is committed and entitled to the premises is entitled, though not committed, to the conclusion. Nonetheless, the conclusion can be thought of as a commitment, because what one is entitled to is in the first instance a commitment.

39. This use of 'shall' is based on Sellars's technical usage of 'shall' as an expression of intention. The treatment of action, intention, and practical reasoning presented here owes a great deal to Sellars's seminal work on the topic, "Thought and Action," in *Freedom and Determinism*, ed. Keith Lehrer (New York: Random House, 1966).
40. The point just made about the inheritance by practical commitments of inferential relations from the corresponding doxastic commitments can be illustrated in these terms by noting that if "*I will* wear a necktie" commitment-entails "*I will* wear something around my neck," then "*I shall* wear a necktie" commitment-entails "*I shall* wear something around my neck."
41. Beginning with "Actions, Reasons, and Causes," reprinted as Chapter 1 in *Actions and Events*. Quotations are from p. 4.
42. The expressive role of such ascriptions should be understood by analogy to that of the explicit ascriptions of doxastic commitment discussed below in 8.1–4.
43. "Intending" p. 86.
44. Hempel discusses this feature of inductive arguments in detail in *Aspects of Scientific Explanation* (New York: Free Press, 1965), pp. 394–403. An example he offers (recast in the idiom of this work) is that one can have good evidence both for the inference codified by the conditional "If the barometer falls, it almost certainly will rain" and for the inference codified by "If the sky is red at night, it almost certainly will not rain." Since there are occasions on which one can be entitled to commitment to both of the antecedents, the incompatibility of the conclusions shows that these inferences cannot be commitment preserving. But they can each be entitlement preserving, even though in the situation where one is entitled to both antecedents, assertion of either can serve as a challenge to the conclusion of the other inference.
45. Literally at the end of this story, in Chapter 9, building on the social-perspectival account of objectivity developed in Chapter 8.
46. "Intending," p. 84.
47. Ibid., p. 85.
48. G. E. M. Anscombe, *Intention* (Oxford: Blackwell, 1959), and Davidson, originally in "Actions, Reasons, and Causes."
49. Davidson, "Actions, Reasons, and Causes," p. 8.
50. Ibid., p. xiii.
51. This and the next passage quoted are from John Searle, *Intentionality* (Cambridge: Cambridge University Press, 1983), pp. 84–85.
52. Sellars, "Thought and Action," p. 110.
53. Ibid., p. 109.
54. This is the conclusion of Davidson's "Intending," summarized at pp. 100–101.
55. Sellars's central text on this topic is "Thought and Action." The view is introduced in his earlier "Imperatives, Intention, and the Logic of 'Ought,'" *Methodos* 8 (1956): 228–268, and is developed at greater length in the Tsanoff lectures, delivered at Rice University in 1978, entitled "On Reasoning about Values." Castañeda's treatment dates to his 1952 University of Minnesota master's thesis, "An Essay on the Logic of Commands and Norms." The fullest statement of his view is in *Thinking and Doing: The Philosophical Foundations of Institutions* (Dordrecht: D. Reidel, 1975). A concise summary can be found in "The Two-Fold Structure and the Unity of Practical Thinking," in *Action Theory*, ed. M. Brand and D. Walton (Dordrecht: D. Reidel, 1976), pp. 105–130.

56. Indeed, in his extremely useful summary of Sellars's views on this topic (pp. 149–188 in *The Synoptic Vision: Essays on the Philosophy of Wilfrid Sellars*, by C. F. Delaney, Michael J. Loux, Gary Gutting, and W. David Solomon [Notre Dame, Ind.: University of Notre Dame Press, 1977]), W. David Solomon paraphrases Sellars's view about the intentions expressed by his regimented 'shall' locutions in just this way: "Whereas expressions of intention manifest my commitment to act at some future time (perhaps precisely datable, perhaps not), volitions are commitments on my part to act here and now" (p. 163). Sellars's treatment of pro-attitudes is quite different from that presented here, and in any case, he has no general account of beliefs as commitments, to which intentions might be assimilated.
57. Davidson, "Intending," p. 90.
58. Ibid.
59. Ibid.
60. The analogy is not exact, for 'shall' indicates a kind of pragmatic *force*, the acknowledging of a practical commitment. A 'shall' statement is not an ordinary assertion in which that force becomes explicit as part of the content, for instance in the way in which 'believe' can be used to make the attribution of a doxastic commitment explicit as part of the assertible content of an ascription (as discussed in Chapter 8). As Sellars says: "'Shall', in spite of its logical role, can be said to be a *manner* rather than a *content* of thought" ("Thought and Action," p. 109). For 'shall' statements do not embed in more complex statements—paradigmatically as the antecedents of conditionals—in the way ordinary assertible contents must. Such embedding strips off the pragmatic force associated with the utterance of the embedded sentence. Thus when I say "If I believe that Kant liked turnips, then I believe that he liked some tuber," I am not saying that I believe either claim. The force of the (self-)ascription "I believe that Kant liked turnips" has been lost, but 'believe' still means exactly what it does in unembedded contexts, even though the pragmatic significance of uttering it is different. It is different with 'shall'. "I shall marry, so I won't be a bachelor" is a good inference, but when codified as a conditional, it takes the form "If I should marry, I won't be a bachelor," not the ungrammatical "If I shall marry, I won't be a bachelor," which contains a defective use of 'shall'. The rest of this section indicates why this shift from 'shall' to 'should' is required in embedded contexts. The result is that 'shall' statements do not make acknowledgments of practical commitments explicit as assertible contents in quite the same sense that 'believe' statements make attributions of doxastic commitment explicit as assertible contents.
61. Self-attribution of a deontic state is what is made explicit by certain uses of 'I' and is not equivalent to merely attributing it to oneself. For the latter may be attributing it to someone who is, as a matter fact, though one is not aware of this fact, oneself—as I might attribute a certain commitment to whoever wrote the words appearing on a scrap of paper I find in the street, not realizing that I wrote them a year ago. It is not at all surprising that one can attribute a practical commitment to oneself without acknowledging it; it is more surprising that one can even self-attribute such a commitment without acknowledging it (though not without undertaking it). This issue is discussed in more detail below in 8.5.2.

62. *Foundations of the Metaphysics of Morals*, trans. Lewis White Beck (Indianapolis: Bobbs-Merrill, Library of Liberal Arts, 1959), p. 29; p. 412 of vol. 4 of the Akademie Textausgabe.
63. Ibid., p. 30; p. 413.

## 5. The Expressive Role of Traditional Semantic Vocabulary

1. In an extended sense: intentions have propositionally expressible conditions of satisfaction corresponding to the claims the agent is practically committed to making-true.
2. Permissive or entitlement-preserving inferences, which are of the first importance in justificatory practices—particularly inductive ones—are notoriously difficult to parse in terms of truth conditions. It is because of the hope that the notion of *reliability* can supply what is wanted that that concept assumes the significance it has for epistemologists who understand the contents of knowledge claims in terms of truth conditions. That notion, and the gerrymandering difficulties to which it is subject, are discussed in Chapter 4.
3. "The Concept of Truth in Formalized Languages," in *Logic, Semantics, Metamathematics: Papers from 1923–1938* (New York: Oxford University Press, 1956), pp. 152–278.
4. Thomas Gray, "Elegy Written in a Country Churchyard."
5. An analogy is the way in which, once the expressive role of 'looks' or 'seems' is properly understood, it becomes apparent that the incorrigibility of statements formed by the use of these operators is trivial in a way that makes them unsuitable precisely for the role of foundation of knowledge that Descartes assigned to them. The analysis required for this argument is presented in Sellar's "EPM" and is sketched below in Section II.
6. Part of what is distinctive about the present approach, however, is that what are here treated as *semantic* primitives are themselves explained in terms of a prior *pragmatics*, which in turn appeals to *normative* primitives, themselves made available by mapping the theoretical idiom onto our ordinary talk.
7. Indeed, that is why semantic theorists, as opposed to linguists, have been in general so little interested in this notion—which will be taken here to be of absolutely the first importance to semantic theory. Chastain (whose work is cited below in Section IV) is a notable exception, as is Hintikka.
8. The material presented in this section and the next originally appeared as "Pragmatism, Phenomenalism, and Truth Talk," in *Realism (Midwest Studies in Philosophy 12 [1988])*, pp. 75–93.
9. The point here does not concern merely the *senses* of the contrasted expressions, but the *extensions* they determine. The appropriateness of this question would have to be defended by adducing cases in which a belief apparently "worked" and was not true, or vice versa. Such cases are not far to seek. This sort of argument is considered more carefully below.
10. *Pragmatism* (Cambridge: Harvard University Press, 1978), bound with its sequel, *The Meaning of Truth*, which as here interpreted ought to be titled *The Meaning of Taking-True*. For an important assessment on a larger scale, see R. Rorty's "Pragmatism, Davidson, and Truth," in *Philosophy of Donald Davidson: A Perspective on Inquiries into Truth and Interpretation*, ed. E. LePore (Oxford:

Basil Blackwell, 1986). The essay that originally presented the material from which the first part of this chapter is drawn was written as a tangential response to this piece of Rorty's.

11. For instance, in "Truth," reprinted in *Truth*, ed. G. Pitcher, (Englewood Cliffs, N.J.: Prentice-Hall, 1964), pp. 32–53. Strawson's view is often also referred to as a 'redundancy' account, in the sense of an account focusing on redundancy of force.
12. *Logic: The Theory of Inquiry* (New York: Holt, Rinehart, Winston, 1938).
13. Anthony Appiah, *Assertion and Conditionals*, Cambridge Studies in Philosophy (Cambridge: Cambridge University Press, 1985), p. 103. J. T. Whyte develops this line of thought in "Success Semantics," *Analysis* 50, no. 3 (June 1990): 149–157, and "The Normal Rewards of Success," *Analysis* 51, no. 2 (March 1991): 65–73. The criticisms mentioned in the text are applied to this version in the author's "Unsuccessful Semantics," forthcoming in *Analysis*.
14. It is in many ways more natural to think of the approach considered here as "expressivism," in Allan Gibbard's sense (see *Wise Choices, Apt Feelings* [Cambridge: Harvard University Press, 1990], pp. 7ff. and throughout), rather than as "phenomenalism." There are differences, however, and in any case the account offered here of expression as making explicit forbids the more familiar rubric.
15. In Quine's *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969).
16. As Sellars argues in "Phenomenalism," in *Science, Perception, and Reality* (London: Routledge and Kegan Paul, 1963).
17. "Ascriptivism" and "Assertion," reprinted in Geach's *Logic Matters* (Berkeley: University of California Press, 1972), pp. 250–253 and 254–269.
18. "Reference and Understanding," in *Meaning and the Moral Sciences* (London: Routledge and Kegan Paul, 1978), p. 108.
19. The objectivity proofs below at 8.6.5 show that this difference is preserved by the deontic scorekeeping model of discursive practice.
20. D. Grover, J. Camp, and N. Belnap, "A Prosentential Theory of Truth," *Philosophical Studies* 27 (1975): 73–125.
21. See P. Geach, *Reference and Generality* (Ithaca: Cornell University Press, 1962), pp. 124ff. For some complications, see B. Partee, "Opacity, Coreference, and Pronouns," in *Semantics of Natural Language*, ed. G. Harman and D. Davidson (Dordrecht: D. Reidel, 1972), pp. 415–441.
22. This discussion must be merely preliminary, since singular terms, pronouns, and substitution have not yet been officially discussed—the model of asserting and inferring stays resolutely at the level of sentences. These deficiencies will be remedied further along. The point of this discussion is in part to motivate those later discussions and also to set up some strategic criteria of adequacy for them. Anaphora in the grammatical category of singular terms is considered in greater detail later in this chapter, though it is only in Chapter 7 that it is shown how such a notion can be defined in the official model of linguistic practices. Substitution is introduced, and singular terms are defined, in Chapter 6.
23. This claim is qualified below in 7.3–4, where an indispensable expressive role is discerned even for 'lazy' anaphora—not only in spite of, but because of this redundancy.
24. Grover, Camp, and Belnap, "Prosentential Theory of Truth," p. 87.

25. The authors of the original theory may believe that for any syntactic prosentence type (paradigmatically "That is true") and any declarative sentence tokening there is potentially an anaphorically dependent prosentence tokening of that type that has the declarative tokening as its antecedent. On the disquotational account of lazy prosentences offered in emendation below, that formulation would not hold.
26. Grover, Camp, and Belnap, "Prosential Theory of Truth," p. 91.
27. *Ibid.*, p. 109.
28. The material presented in this section originally appeared as "Reference Explained Away," *Journal of Philosophy* 81, no. 9 (September 1984): 469–492, reprinted in *Philosopher's Annual* "Best Ten Philosophical Articles of 1984."
29. "Truth and Correspondence," *Journal of Philosophy* 59, no. 2 (January 1962): 29–56, reprinted in *Science, Perception, and Reality*; and Chapter 4 of *Science and Metaphysics* (London: Routledge and Kegan Paul, 1968).
30. In *Language, Mind, and Knowledge*, Minnesota Studies in the Philosophy of Science, vol. 7, ed. Keith Gunderson (Minneapolis: University of Minnesota Press, 1975), pp. 194–269.
31. More exactly, it is tokenings that stand in anaphoric relations, though this nicety will continue to be suppressed in the text. Written expressions are particularly vulnerable to reuse of tokens, which is what enforces the distinction. Tokens of pronouns are as vulnerable to this as are tokens of demonstratives (where the same "you are here" sign might be moved from one map to another on campus).
32. Not all elements of anaphoric chains need be understood as singular referring terms. Chastain says that quantificational, modal, and hypothetical contexts are "referentially segregating" and that the syntactically singular expressions that occur inside them should not in general be understood as singular referring terms. The claims made here do not require special treatment of such segregated occurrences.
33. The significance of this intersubstitution test is pursued at length in Chapter 6. It is suggested below at 7.4.3 (and argued in 8.5.4–6 in connection with Kripke's puzzle about belief) that proper names should themselves be understood as functioning anaphorically, an idea that Chastain himself suggests. For the purpose of introducing anaphoric ideas, however, it does no harm to pretend that proper names function in the idealized way they have traditionally been conceived of as functioning.
34. It would be less clumsy to call the expressions not all cotypically tokens of which are coreferential, 'token reflexive'. That policy is not adopted here, because the phrase has an established usage (due to Reichenbach) and is not generally thought of as applying to expressions like 'the Senator', even when such expressions are used as anaphoric dependents, which are among the paradigmatic cotypically nonintersubstitutable term occurrences. It is argued in Chapter 7 that the special substitution conditions applying to anaphorically dependent expressions are more explanatorily fundamental than indexicality of canonical token-reflexives.
35. One difference that might be remarked between ordinary pronouns and those formed by indirect description concerns *backward anaphora*, in which the anaphorically dependent occurrence precedes its antecedent in the discourse.

Such cases are unusual, though by no means always deviant or strained, for ordinary dependents such as 'he' and 'the man'. Consider the intertwining anaphoric chains in the opening sentences of Henry James's "The Beast in the Jungle":

What determined the speech that startled *him* in the course of *their* encounter scarcely matters, being probably but some words spoken by *himself* quite without intention—spoken as *they* lingered and slowly moved together after *their* renewal of acquaintance. *He* had been conveyed by friends an hour or two before to the house at which *she* was staying; the party of visitors at the other house, of which *he* was one, and thanks to whom it was *his* theory, as always, that *he* was lost in the crowd, had been invited over to luncheon. There had been after luncheon much dispersal, all in the interest of the original motive, a view of Weatherend itself and the fine things, intrinsic features, pictures, heirlooms, treasures of all the arts, that made the place almost famous; and the great rooms were so numerous that guests could wander at their will, hang back from the principal group and in cases where they took such matters with the last seriousness give themselves up to mysterious appreciations and measurements. There were persons to be observed, singly and in couples, bending toward objects in out-of-the-way corners with their hands on their knees and their heads nodding quite as with the emphasis of an excited sense of smell. When they were two they either mingled their sounds of ecstasy or melted into silences of even deeper import, so that there were aspects of the occasion that gave it for *Marcher* much the air of the 'look round' previous to a sale highly advertised. (In *Henry James: Selected Fiction*, ed. Leon Edel [New York: G. P. Dutton, 1964], p. 482. There are four more sentences before the initial 'she' is anchored to a token of 'May Bartram', thus allowing 'their' and 'they' to become attached as well.)

Indirect descriptions, in virtue of the explicit way they pick out the tokens they depend upon anaphorically, exhibit no such prejudice for the discursive past and, accordingly, often possess 'antecedents' only in the broader sense of anaphorically inheriting content from another tokening.

36. This last sort of example shows that the term token can be picked out in a variety of ways, in particular, by citing a predicate type to pick out a sentence tokening that is the characterizing or calling and that contains the term token (not otherwise specified) that on the present account is the anaphoric antecedent of the indirect description in the example. It may also be noted that indefinite descriptions can similarly be constructed from indirect anaphoric sortals, as in "A woman the lawyer referred to as 'she who must be obeyed' explained the matter to us," which both initiates a new chain and characterizes the referent of that chain by relation to some antecedent tokening of 'she who must be obeyed' by the lawyer in question. See the discussion of referential predication below.
37. Speakers' reference is discussed briefly below as an anaphoric phenomenon, and at greater length at 7.5.6 and 8.2.2.
38. It may seem that the presence of a sortal restriction on indirect descriptions causes difficulties. The issue can be avoided, as in {t}, however, by using an anaphorically dependent sortal. 'One' anaphora has long been recognized by

linguists as permitting anaphoric proforms as stand-ins for common nouns, as in "There were red pens as well as green ones on the table." In fact philosophers have made up the expression 'referent of *t*' to mean 'the one referred to by *t*'. So the fact that indirect descriptions are sortally restricted, as are ordinary descriptions (or for that matter quantifications) in natural language, adds no new difficulties to an anaphoric analysis of 'refers'.

39. Chapters 6 and 7 make explicit how these intersubstitutability claims are to be understood.
40. The identification of indirect descriptions by the iteration test provides a sense in which one term *type* can be considered to be anaphorically dependent on another *type*, as the iterated indirect description is on the type of its antecedent token. But this notion of type-anaphora is entirely derivative from the basic notion of token-anaphora, a derivation made possible by the existence of operators that form lexically complex pronouns that are invariant under cotypal intersubstitution.
41. It may seem that talk here and elsewhere of 'picking out', 'coreferring', and so on begs important questions. But it will be shown below that such talk can be understood in a way that does not commit one to reference relations.
42. Joe Camp pointed out the need to deal with these relations, which 'piggyback' psychological relations on semantic ones.
43. The significance of this principle for Frege is discussed at greater length in the author's "Frege's Technical Concepts," in *Frege Synthesized: Essays on the Philosophical and Foundational Work of G. Frege*, ed. L. Haaparanta and J. Hintikka, Synthese Library (Dordrecht: D. Reidel, 1986), pp. 253–295. The motivation for it is pursued at some length in Chapters 6 and 7 below.
44. Compare Frege's similarly motivated paraphrase of "Jupiter has four moons" into "The number of Jupiter's moons is (=) four," in *GL*, sec. 57.
45. "Tarski's Theory of Truth," *Journal of Philosophy* 69 (1972): 347–375.
46. *Journal of Philosophy* 74, no. 10 (October 1977). Also in this paper Grover independently states a weaker version of the observation exploited in the present account: "Descriptive phrases such as 'just mentioned', 'have been talking about', 'are referring to', which ostensibly describe discourse may often be used merely to locate an antecedent piece of discourse from which a referent is inherited" (p. 594). See also Grover's "'This is False' on the Prosentialist Theory," *Analysis* 36 (1976): 80–83.
47. The special anaphoric nature of the disquotation involved is most apparent in the case where the indirect description picks up the speaker's reference of its mentioned antecedent.
48. Though it may be recalled that a sense is given above to the claim that "truth is one" (in 4.4.3, appealed to again in 4.5.4).
49. The specification that it is to be a language-*in-use* is intended to indicate that it should be thought of as comprising not only words and sentences but also the practices in virtue of which they mean what they mean.
50. Paul Boghossian, "The Status of Content," *Philosophical Review* 99, no. 2 (April 1990): 157–184.
51. *Ibid.*, p. 181.
52. As Frege says in "The Thought," in *Essays on Frege*, ed. E. O. Klemke (Urbana: University of Illinois Press, 1968), pp. 507–536.
53. In fact, Boghossian does not so much argue for this claim as assume it, for lack

- of an alternative: "For the remainder of the paper, therefore, I will assume that contents just are truth conditions" ("The Status of Content," p. 173).
54. Compare Sellars's claim in "EPM" that once the expressive role of 'looks' or 'seems' is properly understood, its unsuitability for service in characterizing an epistemological foundation for empirical knowledge becomes apparent, in spite of the incorrigibility of 'seems' claims that so impressed Descartes.
55. A long story would be required to distinguish properly between cases where it would be correct to take the proprieties governing the use of the concept to be determined in this way from those where the alien substance should be taken to be a new kind of water. Our use of 'water' involves commitment to all samples of water having the same fundamental physical constitution, and we have sufficiently developed conceptions of what that means to rule out XYZ and H<sub>2</sub>O correctly being put in the same box, even in a case where we could not tell the difference. It is also possible, of course, to describe cases in which 'water' is so used as properly to be applied to any clear, tasteless, odorless, thirst-quenching liquid, regardless of its composition. It is not easy to say what it is about the prior *behavior* (specified in nonnormative terms) of a tribe that has experience only of, say, warm brackish water that determines whether they are (whether they know it or not) committed to apply the same term they use to the sort of cold, pure water that gushes from mountain springs they have never seen. But this sort of difficulty stems from the fact that many interpretations couched in normative vocabulary are compatible with any given specification of behavioral dispositions or regularities that is couched exclusively in nonnormative terms. The point being made here is rather that discursive practices (which must be specified in normative terms) and the concepts they make available are individuated in part by the causal commerce with objects of various kinds that they incorporate, and that for this reason the proprieties that are implicit in those practices may (according to an interpreter) outrun the explicit discriminative capacities of those being interpreted.
56. So to speak—for according to the idiom being recommended, there is nothing that is in principle <sup>s</sup>outside<sup>s</sup> discursive practice.
57. *PI*, sec. 95.
58. John McDowell has discussed this issue extensively and insightfully, beginning in "Singular Thought and the Extent of Inner Space," Chapter 5 in *Subject, Thought and Context*, ed. P. Pettit and J. McDowell (New York: Oxford University Press, 1986); in the manuscripts "Knowledge and the Internal" and "Scheme-Content Dualism, Experience, and Subjectivity"; and in his 1991 Locke lectures.
59. Tennyson, *The Idylls of the King: The Last Tournament*.

## 6. Substitution

1. If this sort of approach is possible, why is it not the one pursued here? Not because of inferentialism per se, but because of the dual collateral commitments: to understanding semantics in terms of pragmatics—that is, to conceiving the use of expressions as conferring their contents—and to treating asserting as the fundamental linguistic performance (the pragmatic priority of the propositional).
2. This issue is independent of the desirability of an account in nonsemantic terms

(though it is important not to confuse this with the demand for naturalistic—that is, nonnormative—terms) of what capacities the semantic capacities in some sense consist in the deployment of. It is one thing to claim (how could it be denied?) that causal interactions of various sorts with particular objects is a necessary condition of being able to represent empirical states of affairs; it is another to claim that some of these interactions ought to be understood as semantically primitive, in that what it is to represent such a state of affairs ought to be understood in terms of them.

3. *GL*, sec. 65 is one place the distinction is explicitly acknowledged.
4. *FPL*, p. 417.
5. In the standard situation, it has been stipulated which multivalues are designated, so the second assignment determines the first. It is important, however, to keep the two sorts of assignment conceptually distinct.
6. *FPL*, p. 361.
7. *Ibid.*, Appendix to Chapter 12 (on multivalued logic).
8. Thus the designatedness value *undesignated* need not be preserved by good inferences, and substitutions that turn undesignated into designated sentences are strictly irrelevant for assessments of validity of compound sentences.
9. *FPL*, chap. 12.
10. The expressive role of such propositional-attitude-ascribing locutions is explained in terms of the social articulation of substitution-inferential commitments in Chapter 8.
11. It is important that this formulation is in terms of deontic status and does not indicate how to translate the remarks in terms of social deontic attitudes. It does not indicate who is attributing the state and to whom it is attributed. It matters how one goes about eliminating this vagueness. That issue is addressed in Chapter 8, on ascribing propositional deontic attitudes. Notice also that this issue concerns local interdefinability of assertional and inferential commitments. To say that (commitment to) the goodness of inferences is not settled by (commitment to) the truth of the premises and conclusions is not yet to deny that inferential commitments as a whole might be fixed by assertional commitments as a whole, or vice versa. According to the picture presented in Chapter 3 of the sort of indissoluble conceptual package that assertional and inferential commitments and entitlements make, however, the intimate mutual presupposition of discursive and inferential commitments precludes such global reducibilities.
12. See the author's "Semantic Paradox of Material Implication," *Notre Dame Journal of Formal Logic* 22, no. 2 (April 1981): 129–132.
13. *FPL*, p. 401.
14. Strictly, what is referred to by a singular term is a particular. Not all particulars are *objects*; there are also events, processes, and so on. The present argument does not turn on the differences among these sorts of particulars, and it will often be more convenient simply to talk of objects, where in fact any sort of particular can be involved.
15. *Word and Object* (Cambridge: MIT Press, 1960), p. 96; see also p. 90.
16. For present purposes, the fact that the same adverb can modify multiplace predicates—and so should more generally be construed as a  $\{\langle T_1 \dots T_n \rangle \rightarrow S\} \rightarrow \{\langle T_1 \dots T_n \rangle \rightarrow S\}$ —can be ignored.
17. David Lewis's "General Semantics" (in *Semantics of Natural Language*, ed.

G. Harman and D. Davidson (Dordrecht: D. Reidel, 1972) is a masterful account of the ideas behind functional-categorial semantics.

18. *GL*, p. x, secs. 46, 60, 62.
19. Two works that endorse this order of explanation are D. Davidson, "Reality without Reference," and J. McDowell, "Physicalism and Primitive Denotation: Field on Tarski," both in *Reference, Truth, and Reality*, ed. Mark Platts (London: Routledge and Kegan Paul, 1980).
20. "Compound Thoughts," *Mind* 72 (1963): p. 1.
21. In a sense, of course, we do not know how many such sentences there are, even restricting ourselves to a basic vocabulary, since we do not have a syntactically adequate grammar for any natural language. But there are grammars that will generate *only* sentences of English. The difficult thing is getting one that will generate *all* of them, without generating all sorts of garbage as well.
22. For instance, Davidson emphasizes this point in his influential "Theories of Meaning and Learnable Languages," in *Proceedings of the 1964 International Congress of Logic, Methodology, and Philosophy of Science* (Amsterdam: North Holland Publishing, 1965).
23. Notice that the problem of projection such a strategy addresses concerns moving from *proprieties* governing the use of one set of sentences to proprieties governing the use of a superset. A quite different issue concerns the relation between the correct use even of the sentences in the initial subset and the actual occasions of use or dispositions of the community to use them. These puzzles must be sharply separated, for the first remains within the normative dimension, asking about the relation between two different sets of practically embodied norms, while the second asks about the relation between such norms and the nonnormative happenings that express them.
24. Type/token issues are suppressed for the purposes of this chapter. The complications they introduce are the topic of Chapter 7.
25. Corner quotes should be discerned as required here, so that '*p*' refers to the quote-name of the sentence the variable *p* stands for, not for the quote-name of the variable letter.
26. This requirement is not absolute. The author's "STSSD" shows how to make do just with substitutional relations among substituted-in expressions and how to do without antecedently distinguishable substituted-for expressions.
27. Strictly speaking, this is true only of what Dummett calls "complex" predicates, by contrast to "simple" ones, about which more below. But as Dummett points out in making the distinction, Frege "tacitly assimilated simple predicates to complex ones" (*FPL*, p. 30).
28. *Subject and Predicate in Logic and Grammar* (London: Methuen, 1974).
29. From Frege's mature point of view, this qualification does not need to be made: sentences *are* singular terms, and the frames *are* predicates. This is what motivates Frege's classification of sentences as singular terms. As will be pointed out below, this need not be the whole story about sentences, a fact that immunizes Frege somewhat from Dummett's scandalized response to this point. Qua subsentential expressions, sentences are singular terms; the thesis is innocent of the objectionable implications Dummett complains about (missing the special role of sentences as usable to make moves in the language game—as though Frege had no idea of force, and as though being a name of the True or the False did not

play a very special role for him) because sentences are not *essentially* subsentential expressions, and it is not as subsentential expressions that they have their special pragmatic position. I am grateful to John McDowell for pointing this out.

30. This point is distinct from, although related to, the distinction Dummett makes, in Chapter 2 of *FPL*, between simple and complex predicates. Dummett there points out (following Geach's discussion in "Quine on Classes and Properties," *Philosophical Review* 62 (1953): 409–412) that there is no simple *part* or subexpression common to "Rousseau admired Rousseau" and "Kant admired Kant" that is not also a part of "Kant admired Rousseau." Yet the first two share with each other a complex predicate that they do not share with the third. One of Frege's great discoveries was that one must be able to discern predicates in this sense (complex, or substitutionally derived ones) in order to appreciate the inferential role of sentences like "Anyone who admires someone admires himself." For one must appreciate the different patterns they instantiate in order to see that in the context of that quantificational claim, "Kant admired Rousseau" entails "Kant admired Kant." Thus the status of predicates as playing derived substitution-structural roles is what lies behind the second of Strawson's stigmata distinguishing predicates from singular terms: that they are subject to quantification. Concern with quantification, in particular with codifying the inferential role of quantificational claims, enforces the distinction between simple and complex predicates, between expressions that can be substituted for and those that are substitutional frames. But the need for this distinction is not, as Dummett claims (pp. 28, 30), simply a consequence of the presence of quantificational locutions in a language. Complex predicates must be discerned by anyone who has mastered the sort of pattern of inference that is typically made explicit by a quantificational expression, such as  $\{x\}\{y\}[Rxy \rightarrow Rxx]$ . Such inferential connections can be important already in a language, even though quantifiers have not yet been introduced to codify them explicitly as the contents of claims. Nontrivial work must be done (and "STSSD" shows that it can be done, and how) to turn the notion of predicate as equivalence class of substitutionally variant sentences, defined here, into the full-blooded notion of a cross-referenced predicate, as will be required for the introduction of quantifiers. Appendix I at the end of this chapter discusses some related points.
31. Truth-preserving and assertional commitment-preserving inferences (whether they are substitution inferences or not) include deductive inferences but define a wider category, for there is no implication in their case that the goodness of such inferences must be underwritten by their *form* (never mind their specifically *logical* form). Justification-preserving and, more generally, entitlement-preserving inferences include inductive inferences, but once again define a wider category.
32. It should not be thought that all goodnesses of inference must conform to the preservation model, in that there is a kind of status such that the inference is good iff the conclusion has the same status as the premises (any more than it should be thought that all good inferences have some sort of substitutional goodness). The notion of "transmission" of status is intended to indicate that the possession of a certain status by the premise (for instance, that *S* is assertively committed to it) guarantees or provides the reason for the possession of

that status by the conclusion. The remarks in the text apply to commitment-preserving inferences (the genus of which deductive inferences are a species), but it should be noted that they need not apply to *entitlement*-preserving inferences (the genus of which inductive inferences are a species). I am grateful to Ernie LePore for pointing this out.

33. The restriction to substitution inferences is required because one may, for instance, infer asymmetrically from the applicability of a singular term to the applicability of a predicate—from “The inventor of bifocals is Benjamin Franklin” to “The inventor of bifocals is an American.” These do not count as substitution inferences, even in the extended sense allowing replacement of frames, because they cross syntactic categorial boundaries.
34. Sortals, such as ‘dog’ and ‘mammal’, might seem to contradict this claim. For they are distinguished from predicates precisely in having associated with them not only criteria of application but also criteria of identity, and yet they can be materially involved in weakening inferences: “Wulf is a dog, so Wulf is a mammal.” But their criteria of identity apply not to substitutions materially involving the sortals themselves but to those materially involving the singular terms to which the sortals are applied.
35. A fuller account would address incompatibility relations, as well as strictly inferential ones. These are made explicit in quantified negations:  $(x)[Px \rightarrow \sim Qx]$ . A story along these lines is worked out in Mark Lance’s “Normative Inferential Vocabulary: The Explication of Social Linguistic Practice” (Ph.D. diss., University of Pittsburgh, 1988).
36. This following according to a general rule, albeit one that depends on the particular content nonlogical expressions are taken to have, is the source of the *modal* force or flavor that even material proprieties of inference exhibit. For it provides a sense in which a particular inference can be seen to be an instance of a valid type, even though what is quantified over is substitutional situations. This is a point about material inferences (here, the substitutional species of that genus) that can otherwise appear puzzling, since it is clear that the sort of inferences instanced in the text need not be good ones with respect to other possible worlds.
37. Of course what is at issue here is an inferentialist version of the distinction between extensional and nonextensional (or transparent and opaque) occurrences of, typically, singular terms, as discussed in Section II.
38. It need not be denied that occurrences whose significance is not governed in this way are semantically significant in a secondary sense, which can be explained only once the primary sense is understood. This is discussed further along.
39. These examples can represent the asymmetries only at the level of sentences. Singular terms *do not* behave asymmetrically, so real examples of asymmetrically behaving substituted-fors are not forthcoming. Probably the closest one can get in real grammar is sortals. Since they have associated with them criteria of identity for the singular terms they qualify, they are more termlike than predicates. Yet they do have proper inclusions, and a straightforward notion of inferential weakening applies to them, as to predicates. (The objection may now occur that these examples show that expressions like predicates, whose occurrences *do* have asymmetric significances, *can* occur embedded in inferentially inverting contexts, showing that something must be wrong in the analogous

argument to the conclusion that substituted-for expressions must have symmetric substitution-inferential significances. This legitimate worry is addressed further along, where the distinction between basic subsentential expressions, which can be substituted for, and derived subsentential expression-patterns [frames, of derived substitutional category], which can only be replaced [as outermost, hence never embedded], will be invoked.)

40. Recall that to take it that  $q$  is incompatible with  $p$  is to take it that commitment to  $q$  precludes entitlement to a commitment to  $p$ . In this way acknowledgments of material incompatibilities are implicit in the practices governing adopting attitudes (for instance, undertaking or attributing) toward the same pragmatic statuses of commitment and entitlement that inferences can be distinguished as preserving.
41. In conversation, Ken Manders has suggested the language of projective geometry as an example that is interesting in this connection. Sometimes ‘general points’ are appealed to, whose projective properties form a proper subset of the projective properties of some other point or points and so are asymmetrically inferentially related to each other the way sortals can be. The language in which projective properties are specified does *not* have negation or the conditional in it. The present argument explains the unobvious connection between these facts.
42. Indeed it could be argued that possession of this reflexive expressive capacity and all that goes with it makes so much difference that it provides a plausible place to draw the line between the linguistic and the nonlinguistic. The line between logical and prelogical languages is in any case important enough that researchers investigating what sorts of languages chimps and dolphins can be taught would be well advised to postpone trying to teach them an extra two hundred terms and predicates, and instead try to teach them to use conditionals and quantifiers.
43. Notice that this characterization of the conclusion could be accepted even by someone who was not persuaded by the expressive approach to understanding the demarcation of specifically logical vocabulary and so the function of logic.
44. Sentences displaying multiplace predicates in which the same term occupies several distinct argument places may seem to be an exception. A general technique for recognizing adicities in purely substitutional terms in spite of such cases is offered in “STSSD.” There also are predicates, such as ‘... carried the piano up the stairs’, that can be understood as taking variable numbers of arguments (‘Bruno’, ‘Bruno and Betty’, ‘Bruno, Betty, and Bill’ ...). The treatment of these is complex, but their existence does not affect the general point being made here about the discriminability of substituted-fors and substitution frames.
45. Frege, in the *Grundgesetze*, forbids substituted-in expressions that cannot be substituted for. That is, he insists that functions must have “complete” values. There is no technical reason for this restriction, and in the present context, no philosophical reason either.
46. I am grateful to Igual Kvart for raising this possibility.
47. This example is due to Bill Lycan.
48. Walter Edelberg has argued (“Intentional Identity” [Ph.D. diss., University of Pittsburgh, 1984]) that there is a further class of cases of “asymmetric identities”

involving singular terms. These are very interesting cases, but they are not relevant to the present point, since they essentially involve embeddings, indeed multiple embeddings, in propositional attitude constructions. Thus the occurrences of singular terms involved in the Edelberg examples are not primary semantic occurrences, as here defined. In any case, these cases are also sensitive to the particular sentence frames involved, and so also do not generalize in the way required for government by asymmetric SMSICs.

49. This requirement suffices to rule out cases corresponding to the ‘cat’ in ‘cattle’—though it would have to be more carefully stated, to acknowledge the fact that well-formedness *is* preserved in this context by *some* substitutions for ‘cat’, for instance ‘bat’ and ‘rat’.
50. “To begin with” because these equivalence classes will not without further conditions on the substitution relations respect cross-identifications of arguments. “STSSD” shows one way in which one might go on to get the fully individuated complex predicates needed to codify material inferences essentially involving the contents of predicates. See also Appendix I at the end of this chapter.
51. Defining a replacement relation on frames, derivative from substitution for basic expressions, involves defining an isomorphism from one equivalence class to another, preserving these sorts of substitution. Thus to replace ‘ $P\alpha\beta$ ’ by ‘ $Q\alpha\beta$ ’ requires defining a mapping  $h$  from { $s/s$  is in ‘ $P\alpha\beta$ ’} into { $s/s$  is in ‘ $Q\alpha\beta$ ’} that is an isomorphism with respect to the substitution relations among the sentences within those two sets. To say that  $s, s'$  are in the same set ‘ $P\alpha\beta$ ’ is to say that there is some substitution of terms for terms that turns the one into the other. Those substitution relations are indexed by sets of pairs of terms. For instance, the relation indexed by {⟨‘Carlyle’, ‘Ruskin’⟩, ⟨‘Hegel’, ‘Schopenhauer’⟩} turns “If Carlyle wrote *Sartor Resartus*, then Carlyle had an ambiguous relationship with Hegel” into “If Ruskin wrote *Sartor Resartus*, then Ruskin had an ambiguous relationship with Schopenhauer.” The isomorphism must be such that if  $s$  can be turned into  $s'$  in ‘ $P\alpha\beta$ ’ by a mapping Sub with an index of  $t$ , then  $h(s)$  can be turned into  $h(s')$  in ‘ $Q\alpha\beta$ ’ by the mapping Sub with the same index  $t$ . Only equipped with such a mapping can one replace “If  $\alpha$  wrote *Sartor Resartus*, then  $\alpha$  had an ambiguous relationship with  $\beta$ ” by “If  $\alpha$  wrote *Stones of Venice*, then  $\alpha$  was influenced by  $\beta$ .” No such isomorphism across frames (only the indexing of Sub relations within frames) is required to define substitution of one singular term for another. Appendix I expands on this topic.
52. “By and large” because terms can contain other terms. (Indeed, this condition can be used to diagnose the occurrence of terms in which other terms occur, as it is in “STSSD.”) Thus although both ‘Carlyle’s friend’ and ‘Carlyle’ are terms occurring in “John Stuart Mill was Carlyle’s friend,” substitution for either in general eliminates the occurrence of the other. (Only “in general” because of special cases such as substituting ‘Carlyle’s editor’ for ‘Carlyle’s friend’.)
53. This is the sense, explained above in Section II, subsection 2, according to which the fundamental sentential logical vocabulary must be ‘extensional’ in this sense.
54. Negated sentence frames must be discerned if the negation locution that is used to make explicit material incompatibility relations among sentences is to be able also to make explicit material substitution-incompatibility relations among sentence frames. An example of the latter would be what is made explicit by the

claim  $\{x\}[Px \rightarrow \neg Qx]$ , that is, that nothing that has property *P* has property *Q*. The sets of SMSICs that govern frames must include commitments regarding incompatibility as well as those regarding inferential relations among frames. Where the pragmatic significance of sentential inferential relations consists in preservation of some deontic status, whether commitment or entitlement, the pragmatic significance of sentential incompatibility relations consists in commitment to one content precluding entitlement to the other.

55. Though Appendix I shows how to define the second-order role of frames anyway.
56. Notice that this procedure is just a way of keeping track of the interactions between the two fundamental sorts of inverting contexts—negations and antecedents of conditionals. It is a technical device whose use for these purposes does not involve commitment to its being proper for *other* purposes to analyze  $p \rightarrow q$  as  $\neg p \vee q$ .
57. This objection is developed from a suggestion originally offered by Larry Sklar, to whom much thanks.
58. Or the singular terms can be individuated by the transformations. This is the route taken in “STSSD.” Appendix I offers some help with this point.
59. *Tractatus*, 5.62.
60. See the general discussion above at 3.1.3 of the significance of the distinction between associating semantic contents by stipulation and by conferral.

## 7. Anaphora

1. *GL*, sec. 60.
2. This can be only a preliminary discussion of the topic, for Frege does not explicitly take account of the *social* dimension of language use, which, it will be argued in Chapter 8, is essential for a successful account of the representational dimension of discursive practice. (As will emerge, however, he does implicitly acknowledge the social dimension in his concern with the epistemic fruitfulness or information content of identity claims.)
3. *GL*, sec. 84.
4. *Ibid.* On the use of ‘idea’, see for instance the footnote to sec. 58.
5. *Ibid.*, sec. 51.
6. *Ibid.*, sec. 57.
7. *Ibid.*, sec. 74n.
8. All phrases from *ibid.*, sec. 62.
9. See, for example, *ibid.*, sec. 56, 104.
10. *Ibid.*, sec. 74n. Frege uses ‘concept’ as the semantic correlate of predicates—once the distinction is firmly in place, in the 1890 essays, as what their referents are rather than as the senses they express. This usage derives from Kant and, partly due to Frege’s endorsement, still has considerable currency. It is not a harmless or philosophically neutral choice of terminology, however. John McDowell, in “De Re Senses,” *Philosophical Quarterly* 34, no. 136 (July 1984): 283–294, indicates some of the reasons why not. In the context of the present project, whatever is inferentially articulated is conceptually articulated, since the use of singular terms, like that of predicates, is determined by substitution inferences; singular terms are conceptually articulated every bit as much as predicates are. The discussion of anaphoric inheritance of the determinants of substitution-inferential commitments, in the latter part of this chapter, shows how even

deictic tokenings get to be inferentially, and so conceptually, articulated. The way in which such an approach overcomes the dualism of cause and concept about which McDowell complains is addressed below in 9.1.

11. *GL*, sec. 74n.
12. *Ibid.*, sec. 62.
13. *Ibid.*, sec. 84.
14. *Ibid.*, sec. 56.
15. *Ibid.*, sec. 46.
16. *Ibid.*, sec. 106.
17. If in addition the singular term occurs essentially in some *true* claim, then it not only *can* be a way an object is given to those who use the term, it also *is* a way an object is given to them. See the discussion of existential commitments in Section II of this chapter.
18. *GL*, sec. 62.
19. *Ibid.*, sec. 109.
20. Which at this stage in the development of Frege's terminology is equivalent to requiring that they have a determinate inferential role.
21. *GL*, sec. 76.
22. *Ibid.*, sec. 62 (section heading).
23. *Ibid.*, sec. 63.
24. *Ibid.*, sec. 104.
25. *Ibid.*, sec. 65.
26. *Ibid.*, sec. 107.
27. As required, for instance, in the passage quoted above from *ibid.*, sec. 84.
28. As required by the passage quoted above, from *ibid.*, sec. 62: "If we are to use a symbol *a* to signify [*bezeichnen*] an object, we must have a criterion for deciding in all cases whether *b* is the same as *a*."
29. Introducing such terms is accordingly introducing the objects they refer to as well. This is the reason for Frege's otherwise peculiar talk in this context (for example in the footnote to *GL*, sec. 74) about defining *objects* rather than just terms.
30. *Ibid.*, secs. 64, 65.
31. *Ibid.*, sec. 68. Frege uses extensions of concepts rather than collections of objects. Though this difference is important to him for various reasons, it can be ignored for present purposes. Note 34 below indicates some of the subtleties that are being ignored here.
32. *Ibid.*, secs. 63, 104.
33. *Ibid.*, sec. 65.
34. This is of course a tendentious description of what Frege says, one that specifies a lesson that is only implicit in his discussion. For the formulation offered conflates the second definition Frege considers (in the crucial substantive section of the *Grundlagen* that comprises secs. 62–69), which he explicitly rejects, with the third, which he endorses. The reason for the conflation is to avoid having to discuss idiosyncratic features of Frege's notion of concepts that are otherwise irrelevant to the points being extracted here from his treatment of objects. (Admittedly, disentangling these two strands of thought is a delicate matter, given how tightly bound up the two categories are for Frege.) The cardinal point of divergence from Frege's avowed position concerns the question of whether the sense of a numerical identity can be regarded as having been

settled, and so numerical expressions introduced as ways in which objects are given to us, when the truth-values of sortally heterogeneous identities (in the general case, those in which only one of the flanking terms is of the form  $fa$ ) have not been fixed, while those of the sortally homogeneous identities (in the general case, those of the form  $fa = fb$ ) have been. Frege indeed explicitly denies this; it is the reason he gives for rejecting his second definition and replacing it by his third: "This means does not provide for all cases. It will not, for instance, decide for us whether England is the same as the direction of the Earth's axis" (sec. 66). The justification for nonetheless taking the position pursued in the text as in a real sense implicit in Frege's discussion is that Frege never finds a satisfactory solution to the problem of fixing the sense of sortally heterogeneous identity claims—neither in the third and final definition endorsed in the *Grundlagen* nor in his subsequent writings, culminating in the crucial introduction of courses of values in sec. 10 of the *Grundgesetze*. The history of his attempts to come to grips with this issue, and the argument (requiring too many details to be rehearsed here) that those attempts ultimately fail, are presented in the author's "Frege's Technical Concepts," in *Frege Synthesized: Essays on the Philosophical and Foundational Work of G. Frege*, ed. L. Haaparanta and J. Hintikka, Synthese Library (Dordrecht: D. Reidel, 1986), pp. 253–295. The lesson extracted below follows from Frege's discussion using only the auxiliary hypothesis that, for the reasons offered in the essay just cited, the truth-values of the sortally heterogeneous identities *cannot* be settled by stipulation when terms referring to novel abstract objects are introduced.

35. See "Frege's Technical Concepts."
36. All involving the same predicate, that is, sentence-frame. This is a somewhat loose formulation of the situation he is ruling out, but it is a straightforward matter to use the vocabulary developed in Chapter 6 to make it precise.
37. *GL*, sec. 67.
38. *Ibid.*, sec. 76.
39. Like the 'cat' in 'cattle', to use Quine's example.
40. It was observed above that if all the antecedent terms have a complete set of identities associated with them, introducing a new term by stipulating one nontrivial identity relating it to a term in the antecedent vocabulary suffices to associate with it a complete set of identities. However, if their senses are in this way incomplete (not, it is being insisted, indefinite)—if each has associated only some nontrivial identities and nonidentities—then the new term acquires in this way also only an incomplete sense.
41. Sections III–V of this chapter will show that anaphoric structures essentially involve a corresponding holism.
42. "The Conditions of Thought," in *Le Cahier du Collège International de Philosophie* (Paris: Editions Osiris, 1989), pp. 165–171, reprinted in *The Mind of Donald Davidson*, ed. J. Brandl and W. Gombocz (*Grazer Philosophische Studien* 36 [1989]), pp. 193–200.
43. In *Knowledge and the Flow of Information* (Cambridge: MIT Press, 1981).
44. Since the two routes intersect in two places, it is necessary to think of arcs rather than lines as intersecting, and the term 'triangulation' may be less than happy.
45. See Appendix II to Chapter 6.
46. Intending to use an expression as a singular term is adopting a practical com-

mitment to doing what is necessary to make it appropriate for oneself and others to adopt such a stance toward one's performances. How to understand the scorekeeping attitudes such a stance consists in and the proprieties that govern adopting it are accordingly issues that must be addressed before such intentions can be made intelligible. That the order of explanation in this way dictates that semantic intentions not be appealed to as fundamental at this point is just one particular instance of the difference of explanatory strategy, insisted upon elsewhere, that divides the present approach from that of agent semantics.

47. *GL*, sec. 74n.
48. *Ibid.*
49. *Ibid.*, sec. 95.
50. Substitutional construals of quantification are defended against this and other objections, and shown to be not technically inferior to objectual construals, in J. Camp, "Truth and Substitution Quantifiers," *Nous* 9 (1975), and S. Kripke, "Is There a Problem about Substitutional Quantification?" in *Truth and Meaning*, ed. G. Evans and J. H. McDowell (Oxford: Clarendon, 1975).
51. Geach discusses some of the reasons for treating sortally restricted quantification as more fundamental than unrestricted quantification in *Reference and Generality* (Ithaca: Cornell University Press, 1962), for instance at pp. 150–151.
52. This claim, and other similar ones used as examples at various points in the text, are taken from David Wells's *Penguin Dictionary of Curious and Interesting Numbers* (1986).
53. *GL*, sec. 62.
54. *Ibid.*, sec. 22.
55. *Ibid.*, sec. 54. See also his criticism of appeals to 'unit' in place of substantive sortals, at secs. 25–26, 33, and 44.
56. *Ibid.*, sec. 54.
57. A. Gibbard, "Contingent Identity," *Journal of Philosophical Logic* 4 (May 1975): 187–221.
58. For this example and much other wisdom about sortals and sortally restricted singular terms, see A. Gupta, *The Logic of Common Nouns: An Essay in Quantified Modal Logic* (New Haven: Yale University Press, 1980).
59. Geach initiated discussion of this topic. See for instance his "Ontological Relativity and Relative Identity," in *Logic and Ontology* ed. M. K. Munitz (New York: New York University Press, 1973).
60. For the origin of this phrase, see below at 8.5.1.
61. Strictly, it is an abbreviation for one. But so is 100·365·24·60·60. Just as for some purposes the difference between true successor numerals and their decimal abbreviations can be ignored, so for others can the difference between their decimal and their more complex arithmetic abbreviations. This is a matter of what assumptions one is willing to take for granted about various operations in treating expressions as canonical, and accordingly varies with the context of inquiry. The commitments involved in treating the power-set successor sequence as forming canonical designators of transfinite cardinals, and so the utility of such an attitude, appeared differently before the demonstration of the independence of the continuum hypothesis than they did afterward.
62. This account of physical existence in terms of coordinates ultimately mappable onto egocentric space is meant to piggyback on the sort of story Gareth Evans

tells about the relations between public and egocentric specifications in his discussion of demonstratives in Chapter 6 of his *Varieties of Reference*, ed. John McDowell (New York: Oxford University Press, 1982).

63. The use of the expression 'I', which anchors these trajectories, is discussed below in 8.5.2.
64. Strictly, specifications of *regions* are required, but this complexity can safely be suppressed here. Thus the point-specifications offered here should be understood as standing in for the much more complex coordinate specifications of space-time regions.
65. Hard cases, such as beams of light and noises, will not be dealt with here—only paradigms of physical existents. These others often need to be individuated by further physical-dynamic variables. See the discussions in Gibbard, "Contingent Identity," and Gupta, *The Logic of Common Nouns*.
66. Lakatos's fascinating case studies in *Proofs and Refutations*, ed. J. Worrall and E. Zahar (Cambridge: Cambridge University Press, 1976), indicate how various candidates for canonical designators of polyhedra suffered this fate.
67. This account leaves entirely open the question of whether and in what sense various *properties* might be said to exist. One could take a Fregean line and restrict the notion of existence to objects, while denying that properties can be picked out by singular terms. If the introduction of such terms is admitted, then the problem becomes one of determining an appropriate class of canonical designators.
68. This is a consequence of what Gareth Evans calls the "Generality Constraint" (*Varieties of Reference*, sec. 4.3).
69. Schröder, in an early misguided attempt at greater rigor, includes in his algebraic axioms, along with commutativity and associativity of his operators, the stipulation that wherever in his text he uses a symbol such as '×', it is to be understood to have the same meaning. It should be clear that this principle ought to be accorded quite a different status from the others; for instance, commutativity makes sense apart from associativity, but what sense do the statements of those conditions make independently of Schröder's terminological stipulation?
70. Although holding the inheritance relation constant, if the SMSICs governing the second were different, those governing the first would have had to be different; but this is a different counterfactual. The latter resembles the claim that if Oswald *didn't* shoot Kennedy, then someone else *did* (since someone shot him), while the former resembles the claim that if Oswald *hadn't* shot Kennedy, then someone else *would* have.
71. This is not to deny that in some cases (for instance where syntactically definite descriptions are used as anaphoric dependents) what one takes the recurrence class to be may depend on what substitutional commitments one acknowledges.
72. Although the examples considered here have been in the grammatical category of singular terms, it should be clear that the account of anaphora as inheritance of substitutional commitments is not so restricted but applies to sentences, predicates, sortals, and so on, as discussed in Chapter 5.
73. The Appendix to this chapter mentions some of the otherwise important issues about anaphora that are for this reason *not* discussed here.
74. Indeed, the anaphoric antecedent may be a merely *possible*, rather than an *actual*, tokening. This possibility was adverted to in Chapter 5 in connection

with the way modal operators combine with anaphoric antecedents formed using the '... refers ...' operator. The status of merely possible tokenings is discussed further at 8.4.2–3.

75. John Perry, "The Problem of the Essential Indexical," *Nous* 13 (1979): 3–21; David Lewis, "Attitudes *De Dicto* and *De Se*," *Philosophical Review* 88 (1979): 513–543; David Kaplan, "Demonstratives," in *Themes from Kaplan*, ed. J. Almog et al. (New York: Oxford University Press, 1989). Some of these issues are discussed below at 8.5.2–4.
76. Evans, *Varieties of Reference*, especially chap. 6.
77. For the sake of definiteness, the discussion here addresses the demonstrative use of singular terms. Similar remarks apply to demonstratives in other grammatical categories, with suitable adjustments for the differences in what sort of substitutional commitments govern the use of those expressions.
78. Insisting that demonstratives or indexicals put us in cognitive touch with the world only insofar as they can be hooked up to the rest of our conceptual apparatus in this way is the point of the discussion of "Sense Certainty" with which Hegel opens the *Phenomenology*.
79. This is to say that expressions of the form /The *K* *S* demonstrated while saying ⟨t⟩/ should be understood by analogy to the anaphorically indirect definite descriptions analyzed above in 5.4. They are not descriptions but anaphorically indirect demonstrations. Such expressions pass the iteration test proposed there for a suitably broad notion of demonstration—one that allows pointing at an image in a mirror, photograph, or video screen to count as a demonstration of the object imaged.
80. In a certain sense, however, it is not the case that *all* anaphoric token-recurrence structures exhibit modal rigidity. The telling exceptions might be called 'impure' or 'descriptive anaphoric structures'. Thus anaphoric dependents that are indirect definite descriptions do *not* behave rigidly (as is pointed out above in 5.4.5). Leibniz might not have been so-called: it is possible that the one referred to as 'Leibniz' not be Leibniz. So 'the one referred to as Leibniz', though in fact anaphorically dependent on a tokening of 'Leibniz', need not be coreferential with it. These cases are complicated, but they can be accounted for by assembling the raw materials presented here. The important thing is to see why they do not falsify the account offered in the text, for in the counterfactual situations being envisaged, various expressions *would* belong to *different* token-recurrence structures than they in fact do. Such cases are parasitic on the basic ones discussed here, and a full treatment of their intricacies would have to appeal to the phenomena described here in the base-level discussion of pronouns, descriptions, and rigidity. A complementary phenomenon is the anaphora that underlies the use of *reflexive* constructions. Here counterfactual mutation of the recurrence structures is forbidden; we cannot coherently specify situations in which Leibniz would not be *himself*.
81. For instance in "On the Logic of Demonstratives," *Journal of Philosophical Logic* 8 (1979): 81–98, reprinted in *Propositions and Attitudes* ed. N. Salmon and S. Soames (New York: Oxford University Press, 1988); and "Dthat," in *Syntax and Semantics*, ed. P. Cole (New York: Academic Press, 1978), reprinted in *Demonstratives*, ed. P. Yourgrau (New York: Oxford University Press, 1990).

82. So long as only pure or nondescriptive anaphoric links are involved.
83. "Naming and Necessity," in *Semantics of Natural Language*, ed. G. Harman and D. Davidson (Dordrecht: D. Reidel, 1972).
84. This suggestion was made by Chastain already in the late 1960s, at the outset of the contemporary discussion of causal-historical theories of reference (in his "Reference and Context," in *Language, Mind, and Knowledge*, Minnesota Studies in the Philosophy of Science, vol. 7, ed. Keith Gunderson [Minneapolis: University of Minnesota Press, 1975], pp. 194–269), but it has not been pursued.
85. A good introduction to the recent linguistics literature on anaphora is Joseph Aoun's *Grammar of Anaphora* (Cambridge: MIT Press, 1986). A very useful philosophical discussion of anaphora that (like the current approach) does not depend on a fundamental distinction between intrasentential and discourse anaphora is J. Hintikka and J. Kulas's *Anaphora and Definite Descriptions* (Dordrecht: D. Reidel, 1985).
86. Interpretation in this sense should be sharply distinguished from Davidsonian interpretation, which is allegedly the same within languages and across languages. Interpretation in that sense requires explanatory hypothesis formation and inferences starting from the *noises* another utters. Against this it has rightly been objected that sharing a language involves being able to "hear another's meanings" rather than another's noises. At issue here is the subcapacities involved in being able to do that. This is discussed further at 8.2.1–2.
87. Or one could conjoin with each of  $p$  and  $q$  a set of auxiliary hypotheses of the form  $(p \text{ or } q) \rightarrow r$ . Clearly versions of this strategy will apply to theories  $T_1$  and  $T_2$ .
88. Frege had made the same point (distinguishing by their lack of inferential consequences cognitively trivial identities from the substantive ones that give us cognitive access to an object) already in the *Grundlagen*, in the passage from sec. 67 quoted above in Section I, Subsection 5.
89. Both quotations are from *Logic*, by Immanuel Kant, trans. R. S. Hartman and W. Schwarz (New York: Dover Publications, 1974), pp. 45 and 118 respectively.
90. I think that putting this word in is a slip on Kant's part (it is not a problem with the translation—the original does say that nothing further [*weiter*] is then known). The *Logik* consists of published lecture notes, and Kant did not revise or edit it as carefully as he did some of his other works. On Frege's view, as rehearsed in Section I, and also according to Kant's own official view, this sentence should end: ". . . I know nothing of him at all."
91. As long as the expressions occurring in the tautology also have nontautologous occurrences (that is, also occur in claims that are rich in inferential consequences), they have a sense—both for Kant and for Frege.
92. For present purposes it is possible to ignore the involvement of a claim in empirical entries and practical exits—perception and action—which add noninferential (but still norm-governed) circumstances and consequences of application and hence contribute to its content.
93. See for instance his *Against Method: Outline of an Anarchistic Theory of Knowledge* (London: Humanities Press, 1975).
94. Addressed in Israel Sheffler's *Science and Subjectivity* (Indianapolis: Bobbs-Merrill, 1967).

95. The difference between committive and permissive inference is being suppressed here.
96. Reprinted as Chapter 14 of David Lewis's *Philosophical Papers*, vol. 1 (New York: Oxford University Press, 1983); originally appeared in *Language, Mind, and Knowledge*, Minnesota Studies in the Philosophy of Science, vol. 7, ed. Keith Gunderson (Minneapolis: University of Minnesota Press, 1975), pp. 3–35.
97. Locus classicus is "Concepts As Involving Laws, and Inconceivable without Them," in *PPPW*, pp. 87–124.
98. For that reason it cannot be stated as a series of theses about what Fregean senses really are, and how they ought to be understood to be related to their referents. Compare the more orthodox multileveled schemes that Lewis considers in "Tensions," reprinted as Chapter 14 in his *Philosophical Papers*; originally in *Semantics and Philosophy*, ed. M. K. Munitz (New York: New York University Press, 1974).
99. So in spite of the absence of shared intensions, we are able to understand Rutherford, to extract information from his claims, and to agree or disagree with him by using the same interpretive strategies we use to cope with each other—by finding out what the other is talking about, how the other is representing things as being. Concrete scorekeeping practices do the duty of abstract intension-functions. The question of what it is for one rather than another set of finely individuated contents to be conferred on expressions and performances by their use is then traded for the question of what it is for a community to engage in one rather than another set of scorekeeping practices.
100. More precisely: it *cannot* arise for the case of intrasentential anaphora, for it can occur even in the intrapersonal case—for instance if there is sufficient lapse of time for the individual in question to have altered commitments.
101. K. Donnellan, "Reference and Definite Descriptions," *Philosophical Review* 77 (1966): 281–304; S. Kripke, "Speaker Reference and Semantic Reference," in *Contemporary Perspectives in the Philosophy of Language*, ed. P. A. French, T. E. Uehling, Jr., and H. K. Wettstein (Minneapolis: University of Minnesota Press, 1977), pp. 6–27.
102. Gareth Evans, "Pronouns, Quantifiers, and Relative Clauses," part 1, *Canadian Journal of Philosophy* 7 (1977): 467–536; reprinted in Evans, *The Collected Papers* (Oxford: Clarendon Press, 1985).
103. Stephen Neale, *Descriptions* (Cambridge: MIT Press, 1990), esp. pp. 170–183, whose account is followed in many respects in the discussion below.
104. *Ibid.*, p. 171.
105. Neale draws a corresponding conclusion at *ibid.*, pp. 182–183.
106. Evans, "Pronouns." Neale offers this formulation of the notion of c-command: "A phrase  $\alpha$  c-commands a phrase  $\beta$  if and only if the first branching node dominating  $\alpha$  also dominates  $\beta$  (and neither  $\alpha$  nor  $\beta$  dominates the other)" (*Descriptions*, p. 173).
107. In Geach, *Reference and Generality*, p. 117; and in his "Referring Expressions Again," in *Logic Matters* (Berkeley: University of California Press, 1972).
108. L. Karttunen, "Definite Descriptions and Crossing Coreference," *Foundations of Language* 7 (1971): 157–182.
109. Neale, *Descriptions*, p. 197.

## 8. Ascribing Propositional Attitudes

1. Jefferson is reported to have said about a reported sighting by a Yale geologist that it is easier to believe that Yankee professors lie than that stones fall from the sky.
2. See for instance E. Sosa, "Propositional Attitudes *De Dicto* and *De Re*," *Journal of Philosophy* 67 (1970): 883–896.
3. The account offered in what follows can be understood as adapting this standard analysis of the two readings (in terms of scope) to the methodological setting provided by deontic scorekeeping accounts of linguistic-social practice. The different repertoires of discursive commitments associated with the one undertaking a commitment and the one ascribing it provide the different contexts of evaluation.
4. The claim expressed by this sentence, or its French equivalent (though Voltaire spoke excellent English).
5. "Quantifiers and Propositional Attitudes," *Journal of Philosophy* 53, no. 5 (1956): 177–187; reprinted in Quine's *Ways of Paradox* (New York: Random House, 1966), pp. 183–194, and in *Reference and Modality*, ed. L. Linsky (New York: Oxford University Press, 1971).
6. Dennett argues for this way of thinking of the distinction in "Beyond Belief," in *Thought and Object*, ed. A. Woodfield (New York: Oxford University Press, 1982). Those who are inclined to think instead of *de dicto* and *de re* as two kinds of belief should be given pause by the fact that on this line it would seem there must be not just two sorts but a potentially infinite number of kinds of belief. For ascriptional operators iterate, so that *S* may believe of *t* that *S'* believes that  $\Phi(it)$ , or believe that *S'* believes of *t* that  $\Phi(it)$ , and these "mixed" ascriptions can themselves be embedded in turn in further ascriptions of either the *de re* or *de dicto* sort. These kinds of iterated ascriptions are discussed in more detail in the Appendix to this chapter.
7. Notice that 'that individual' here is an anaphorically dependent expression. The need to use anaphora in expressing this relation is reflected in the ascription-structural anaphora that connects an antecedent in the scope of the 'of' to a dependent in the scope of the 'that' in regimented *de re* ascriptions. This connection is a manifestation of the deep point that *de re* ascriptions make explicit the sort of *communication* that has already been picked out (in Chapter 7) as implicitly secured by interpersonal anaphora. This point is discussed further below.
8. From a technical point of view, as will become clear in the next section, this proposal just transposes into the current idiom a suggestion that Kaplan offers and then rejects early on in "On Quantifying In," in *Words and Objections: Essays on the Work of W. V. Quine*, ed. D. Davidson and J. Hintikka (Dordrecht: D. Reidel, 1969), also in *Reference and Modality*, ed. L. Linsky (New York: Oxford University Press, 1971), pp. 112–144. The immense influence of this essay, which secured the (from the point of view of this work misleading, indeed disastrous) emphasis on epistemically strong *de re* ascriptions as the phenomenon of most philosophical interest, is largely responsible for what is here taken to be the real philosophical significance of *de re* ascriptions having been overlooked.

9. Davidson, for instance, is criticized on these grounds in Ian Hacking's "Parody of Conversation," in *Philosophy of Donald Davidson: A Perspective on Inquiries into Truth and Interpretation*, ed. E. LePore (Oxford: Basil Blackwell, 1986), pp. 447–458, and in Stephen Mulhall's "Davidson on Interpretation and Understanding," *Philosophical Quarterly* 37, no. 148, (July 1987): 319–322.
10. *PI*, 201.
11. Only some claims are rules, for rules are a kind of normative claim (discussed in Chapter 4).
12. The phenomenology is not really to the point. The claim here is that consciousness in an important sense *consists in* the capacity to keep score, make substitutions, and so on. So on pain of an infinite regress, these must be things that can be done without conscious deliberation or rehearsal.
13. Iteration of ascription-forming operators produces further varieties as combinations of these basic ones, since a *de re* ascription can be embedded in a *de dicto* ascription ("S' believes that S believes of t that Φ(it)") and so on. The way in which the account of the basic forms is to be extended to these cases is discussed in the Appendix to this chapter.
14. Assuming that it is appropriate to treat intentions (or at least the theoretical version of them: practical commitments, commitments to act) as having *propositional* contents. This is not obviously correct. The primary locution for expressing and ascribing intentions in English is intending to do something rather than intending *that* something be true. The relations between them are complex, and pursuing them here would require more detail than is appropriate given the broad-brush strokes in which the rest of the treatment of intention is painted. Roughly, the idea is this. 'Intend' is a verb whose complement is restricted to agentives. It can be presented in a regimented language as taking exclusively propositional complements provided these are restricted to propositions whose outermost operator is an agentive-explicating operator, of which the paradigm is Belnap and Perloff's STIT. (See M. Perloff and N. Belnap, "Seeing To It That: A Canonical Form for Agentives," *Theoria* 54 [1988]: 175–199; reprinted, with corrections, in *Knowledge, Representation, and Defeasible Reasoning*, Studies in Cognitive Systems, vol. 5, ed. H. E. Kyburg, Jr., R. T. Loui, and G. N. Carlson [Dordrecht: Kluwer Academic Press, 1990], pp. 167–190.) The canonical form of an ascription of intention is then something like:

*i*: S intends that *a*(S) STIT(*p*),

where */a*(S)*/i* is an ascription-structural anaphoric dependent whose antecedent is */S/i*. Indexicals and quasi-indexicals play a special role in such ascriptions, as discussed below in the section on strong *de re* ascriptions. The point being made in the text, which does not depend on these subtleties, emerges more clearly if they are suppressed in favor of a parallel treatment of ascriptions of doxastic and practical commitments as alike taking propositional content-specifying complements.

15. Requiring that the noninferential elicitation of the performance be the exercise of such a reliable differential responsive disposition (discussed for the case of perception in the first section of Chapter 4) obviates the difficulty otherwise raised by Davidson's nervous mountaineer. In offering this conceptual placeholder in lieu of Davidson's more general "caused in the right way" condition,

the theory-sketch presented here is more specific. In the absence of a discussion of the difficult problem of individuating skills and dispositions and sorting particular performances according to which are being exercised, this suggestion presents only the form of an account.

16. Compare the discussion of speaker-referring uses of definite descriptions—which pick out an object by a description that is not true of it.
17. This tendency is widespread. One relatively sophisticated representative of it is Colin McGinn's "Structure of Content," in *Thought and Object*, ed. A. Woodfield (New York: Oxford University Press, 1982).
18. In fact it is harder than generally acknowledged to tell a coherent story at all about states that are narrow in this sense. The canonical approach, through twin-earth examples, involves a relativity to admissible descriptive vocabulary whose significance is sometimes not understood. Narrow states are what would be shared by individuals as much alike as they could be, embedded in different total environments. But a difference in context is a difference in what is true of the individuals. They can be made out as twins only by restricting the vocabulary that is permitted to be used in distinguishing them. And there is no way to motivate privileging a vocabulary that will accomplish what is desired, without begging the question—the twins are not, for instance "atom for atom identical" when the vocabulary is restricted only to purely *physical* terms, for "being twenty feet away from 100 kilograms of H<sub>2</sub>O," rather than XYZ, is a perfectly good physical predicate. Nor will it do to talk in terms of what the individuals can discriminate, for what they can respond differentially to depends on what counts as distinct responses, and the same problem arises for motivating restrictions on the vocabulary in which one specifies those responses. The challenge is to specify a vocabulary meeting the two constraints, first, that twins in different environments are indistinguishable by descriptions in that vocabulary (the language of physics or of purely physical properties of behavior will *not* do) and, second, that the underdetermination of semantic properties of their states by descriptions in that restricted vocabulary shows something interesting. The point is not that these constraints are impossible to satisfy but that doing so requires a more delicate touch than is often appreciated.
19. A worry one might have about this part of the program is that the account of the distinction between *de re* and *de dicto* in terms of social differences of inferential perspective corresponding to different repertoires of discursive commitments is tailored specifically to expressions of propositional attitudes, formed using some analog of a "believes that" sentential operator. But other sentential operators, most notably those expressing tense and alethic modality, also exhibit a distinction between *de re* and *de dicto* readings—a distinction that can usefully be construed as a difference in the scope of those operators. Since tensed constructions, necessity, and possibility are not among the locutions officially reconstructed as logical (explicative of discursive practice) in this work, there is a methodological excuse for not addressing this worry in detail. A word or two about how the present account might be extended to tense and modal operators is nonetheless in order. Each of these cases is like that of propositional attitudes in involving the adoption of different perspectives, specifying the content of a claim from the point of view of different repertoires of commitments. In the case of tense, these are commitments that have been or

will be adopted, rather than those adopted by different contemporary individuals. In the modal case, the repertoires are associated not with different interlocutors but with different possible situations. These can be construed as sets of facts—that is, true claims ('claims' in the sense not of claimings but of what is claimed). Modulo such differences in how the different repertoires of commitments that provide the different perspectives are conceived, the social-perspectival account of the distinction between *de re* and *de dicto* content-specifications offered here can be applied to the case of tense and modality, as well as to that of propositional attitudes, which provides the paradigm.

20. Of course it was claimed above that some of the components of propositional-attitude-ascribing constructions—paradigmatically the intentional 'of' or 'about' whose scope indicates *de re* uses of content-specifying vocabulary—play a *semantically* explicitating expressive role, making representational substitutional commitments explicit that would otherwise remain implicit in the interpersonal interpretive scorekeeping practices constitutive of communication. Here, as is the case with interpersonal anaphora and other phenomena involving hybrid deontic attitudes, it is impossible to keep semantic and pragmatic considerations distinct.
21. One need not have produced (nor have been disposed to produce) a performance explicitly acknowledging a commitment in order nonetheless to have undertaken it (a scorekeeper need not take one to have produced or been disposed to produce such a performance in order to attribute the corresponding commitment); some commitments are acquired consequentially in virtue of others one *is* prepared to acknowledge, and others may be acquired by default simply in virtue of one's *not* being prepared to disavow them. Yet such assertional performances are essential to discursive practice in the sense that the ideas of doxastic commitments undertaken consequentially and by default are unintelligible in abstraction from practices that include the possibility of avowing those commitments overtly.
22. Notice that even though one can find out about another's commitment-acknowledging performance by testimony rather than by perceiving it, doing so still requires perceiving the assertional performance that conveys the testimony.
23. This way of putting things is chosen to indicate that it is essential to ascriptional tropes that the one to whom a commitment is ascribed *may* be someone other than the ascriber. Apart from this possibility, the role of ascriptions as making deferrals explicit, for instance, is unintelligible. In the context of such an interpersonal practice, however, nothing rules out self-ascriptions. Section V (Subsection 2) discusses some aspects of self-ascriptions, in connection with the use of the word 'I', which makes explicit self-ascriptions when the self-attributions they express have the significance of *acknowledgments* of commitments. For not all attributions of a commitment to oneself are acknowledgments of it—they must, as it were, be attributions to oneself *as* oneself. So self-attribution cannot in general be identified with acknowledgment. Neither can it be identified with undertaking a commitment, for just as one can undertake a commitment without acknowledging it, one can undertake it without attributing it to oneself.
24. Quine's corner quotes are adopted here as a terminological convenience for expressing generalizations involving quotation. Thus '*p*' is to function as a

variable ranging over the results of applying quotation marks to the sentences that the variable  $p$  ranges over. (This fussy convention is needed because the other order of precedence in applying the two operations is standard for ordinary quotation marks: “ $p$ ” is a quote-name of a letter of the alphabet.)

25. Davidson's piece was originally published in *Words and Objections*, ed. D. Davidson and J. Hintikka (Dordrecht: D. Reidel, 1969), pp. 158–174; reprinted in Davidson's *Inquiries into Truth and Interpretation* (New York: Oxford University Press, 1984), pp. 93–108. McDowell's essay is in *Reference, Truth, and Reality*, ed. Mark Platts (London: Routledge and Kegan Paul, 1980), pp. 206–237.
26. For some applications (see the next note), the account requires talk of merely *possible* tokenings. McDowell goes to some trouble to show that such an appeal should not be metaphysically or ontologically worrisome; possible tokenings can just be thought of as ordered triples of a lexical-syntactic sentence type, a speaker, and a time (and of course further indices, for instance specifying a possible world, could be added as needed).
27. “Would be” because the displayed sentence tokening is not used assertively and so is not a “saying” in the sense being reconstructed here. This formulation accordingly invokes a virtual tokening, of the same type as the displayed tokening in the ‘that’ clause, and in largely the same circumstances of evaluation as the ascribing tokening. This way of expressing the relation presents another avenue for avoiding the “subtle flaw” Davidson finds in more quotational treatments of indirect discourse, related to but distinct from the different strategies Davidson and McDowell put forward.
28. Which of course is not to say that it would be *successful*, for that would depend on whether the one to whom the claim is ascribed is entitled to it, and whether that entitlement is heritable by the ascriber.
29. “De Re Senses,” *Philosophical Quarterly* 34, no. 136 (July 1984): 283–294.
30. Davidson does not consider the possibility of a variant that would assimilate the relation between ‘that’ and the sentence tokening that follows it to *anaphoric*, rather than *demonstrative*, constructions. Such a course is recommended by the fact that the relation involved is, like anaphora and unlike demonstration in general, intralinguistic. It may be noticed also that where saying that has the sense considered here, of saying *claimingly*, ‘that’ can be replaced by the properly prosentential ‘that is true’ in Davidson’s account—Galileo said that is true: the earth moves. The upshot of the discussion here, though, is that neither of these heroic expedients is required by the essential insights of Davidson’s approach.
31. Recall Field’s approach, mentioned in Chapter 3, according to which  $S$  believes that  $p$  is analyzed into two parts: (i)  $S$  believes\*  $\sigma$  (some sentence), and (ii)  $\sigma$  means that  $p$ .
32. Of course the sentential/propositional division is not exhaustive, and to mention these challenges is not to suggest that they cannot be met. A prominent approach, which Stalnaker calls “pragmatic,” looks to intentional action to connect propositional contents with attitudes. In some forms, this approach appeals to decision-theoretic practical rationality in the individual case, and to conventions in the communal linguistic case. A primary source is David Lewis’s “Languages and Language,” in *Language, Mind, and Knowledge*, Minnesota Studies in the Philosophy of Science, vol. 7, ed. Keith Gunderson (Minneapolis:

University of Minnesota Press, 1975), pp. 3–35, reprinted in Lewis's *Philosophical Papers*, vol. 1 (New York: Oxford University Press, 1983.)

33. Section V looks in more detail at communication that should be understood in terms of asymmetric constellations of tokenings bound together by anaphoric commitments. Section VI discusses in more general terms how propositional and conceptual contents should be thought of according to this model.
34. Recall the discussion above in 7.5 of the importance of anaphora in securing a common referent and so making communication possible.
35. Tyler Burge, "Belief *De Re*," *Journal of Philosophy* 74 (1977): 341, talking about a phenomenon previously noted by Castañeda and Loar.
36. Although for simplicity singular terms are used in all the examples here, it should be kept in mind that expressions of *any* grammatical category can be exported into the scope of the *de re* 'of'—not only predicates and sortals, but even whole sentences, as in "Senator McCarthy believed of the first sentence of the Communist Manifesto that it was true," in which the ascription-structural anaphoric dependent is a prosentence.
37. "Quantifiers and Propositional Attitudes."
38. The fact that what is reconstructed as corresponding to belief is the deontic status of doxastic commitment, construed as a normative social status conferred by linguistic scorekeeping practices, represents a difference from the traditional approach that can be largely ignored for the purposes of making the present point.
39. Sellars proposes a denotational view of what is expressed by *de re* belief ascriptions, in "Some Problems about Belief," in *Words and Objections*, ed. D. Davidson and J. Hintikka (Dordrecht: D. Reidel, 1969).
40. "On Quantifying In," first sentence of sec. 8.
41. The idea of *canonical designators* employed above in 7.2 in connection with the explication of existential commitments grows out of this idea of Kaplan's.
42. Sosa, "Propositional Attitudes *De Dicto* and *De Re*," p. 890.
43. Burge "Belief *De Re*"; John Perry, "The Problem of the Essential Indexical," *Nous* 13 (1979): 3–21, reprinted in *Propositions and Attitudes*, ed. N. Salmon and S. Soames (New York: Oxford University Press, 1988); David Lewis, "Attitudes *De Dicto* and *De Se*," *Philosophical Review* 88 (1979): 513–543; David Kaplan, "Dthat," in *Syntax and Semantics*, ed. P. Cole, (New York: Academic Press, 1978), pp. 221–243, reprinted in *Demonstratives*, ed. P. Yourgrau (New York: Oxford University Press, 1990), pp. 11–33; see also D. Kaplan, "Thoughts on Demonstratives," in *ibid.*, pp. 34–49.
44. The essays in N. Salmon and S. Soames, eds., *Propositions and Attitudes*, present a good overview of this idea.
45. Dennett calls what is expressed by *de re* ascriptions on the denotational view "weak aboutness" and distinguishes it from "strong aboutness" in "Beyond Belief," p. 67.
46. One thinker who defends the not uncommon view that *de re* beliefs can be made sense of antecedently to any sort of *de dicto* beliefs is Fred Dretske, in *Knowledge and the Flow of Information* (Cambridge: MIT Press, 1981), esp. pt. 3.
47. Particularly helpful on this point is McDowell's criticism of Burge's formulations of the doctrine, in "De Re Senses." How one should think of the conceptual in this connection is discussed further in the Conclusion of this work.

48. In "The Problem of the Essential Indexical" and "Frege on Demonstratives," *Philosophical Review* 86 (1977): 474–497. The earlier discussions are in Hector-Neri Castañeda's "'He': A Study in the Logic of Self-Consciousness," *Ratio* 8 (1966): 130–157, and "Indicators and Quasi-Indicators," *American Philosophical Quarterly* 4 (1967): 85–100, discussed below.
49. For prior intentions. In the case of intentions in action, the acknowledgment of a practical commitment that could be expressed by using 'I' (or 'now') is the nonlinguistic performance itself—fleeing the bear or heading for the meeting. See above at 4.4–6.
50. In *Mind and Language*, ed. S. Guttenplan (New York: Oxford University Press, 1975), reprinted in *Demonstratives*, ed. P. Yourgrau (New York: Oxford University Press, 1990), and in Anscombe's *Collected Philosophical Papers*, (Minneapolis: University of Minnesota Press, 1981).
51. Anscombe rejects the suggestion that 'I' is a pronoun, though not one that behaves just like any other, as "no good, because 'pronoun' is just a rag-bag category; one might as well say 'It is the word that it is.'" ("The First Person," p. 143; this and subsequent page references are to the Yourgrau collection). But this overstates the case, given that, as she acknowledges "I' functions syntactically like a name" (p. 138)—that is, it plays the substitutional role characteristic of singular terms, including figuring in the identity claims that make the substitutional commitments that govern their use explicit. The challenge is, as she says, just to specify its meaning, how it *is* used.
52. *Ibid.*, p. 151.
53. *Ibid.*, p. 137.
54. *Ibid.*, p. 138.
55. This discussion of Anscombe is much indebted to McDowell's lectures on her essay, but it should not be inferred that he would agree with what is said about it here.
56. See Gareth Evans's discussion, in Chapter 6 of his *Varieties of Reference*, ed. J. McDowell (New York: Oxford University Press, 1980).
57. McDowell, "De Re Senses," p. 105.
58. Hector-Neri Castañeda, "'He"'; "Indicators and Quasi-Indicators"; and "The Logic of Self-Knowledge," *Nous* 1 (1967): 9–22.
59. In "The First Person," p. 136.
60. It may be objected that in that case no definite belief has been ascribed because it has not been specified which painting is the subject of the belief. According to the anaphoric analysis of '*this*' offered below, the object has in fact been specified, though perhaps not in a form recoverable by the audience of the ascription. This can be remedied within the regimentation of ascriptions, however, as is discussed below.
61. The ascription is in the ascriber's language, though the tokening (perhaps merely virtual) that is reported need not be.
62. If one language is at issue—otherwise a tokening of the corresponding type in the language of the target of the ascription.
63. Note that doing without the quasi-indexicals would require a special convention concerning the understanding of 'as' clauses in the case where the ascription is in a different language than the believer would use to express the belief.
64. This formulation is incomplete, as appears below. The one responsible for a *de*

*re* ascription of a strong *de re* attitude also undertakes an *existential* commitment governing the use of the term *t*, in virtue of commitment to the identity. This point is discussed below.

65. For instance, it should now be clear how to extend the ‘A’ language discussed above so as to incorporate an anaphoric ‘U’ (corresponding as ‘you’, in some of its uses, does to ‘I’) that is used only as a dependent on some tokening of ‘A’.
66. Bertrand Russell, “The Philosophy of Logical Atomism,” reprinted in *Logic and Knowledge*, ed. R. C. Marsh (London: George Allen and Unwin, 1956), pp. 177–281.
67. David Kaplan introduces this notion; see for instance “Demonstratives,” draft no. 2 (Mimeo graph, UCLA Department of Philosophy, 1977). A general discussion can be found in Nathan Salmon, *Reference and Essence* (Princeton: Princeton University Press, 1981). Salmon puts the claim this way: “The contribution made by an ordinary proper name, demonstrative, or other simple singular term to securing the information content of, or the proposition expressed by, declarative sentences (with respect to a given possible context of use) in which the term occurs (outside the scope of nonextensional operators such as quotation marks) is just the referent of the term, or the bearer of the name (with respect to that context of use)” (“Reflexivity,” reprinted in *Propositions and Attitudes*, ed. Nathan Salmon and Scott Soames [New York: Oxford University Press, 1988], p. 241). See also the essays by Mark Richard and Scott Soames in *ibid.*
68. John McDowell, “Singular Thought and the Extent of Inner Space,” in *Subject, Thought, and Context*, ed. P. Pettit and J. McDowell (New York: Oxford University Press, 1986), pp. 137–168. See also McDowell, “De Re Senses.”
69. One might be inclined to say that a properly produced tokening of ‘this’ always succeeds in picking out *something*, but this is not the case, for reasons that Anscombe admirably explains: “It used to be thought that a singular demonstrative, ‘this’ or ‘that’, if used correctly, could not lack a referent. But this is not so, as comes out if we consider the requirement for an answer to ‘this what?’ Someone comes with a box and says ‘This is all that is left of poor Jones’. The answer to ‘this what?’ is ‘this parcel of ashes’; but unknown to the speaker, the box is empty. What ‘this’ has to have, if used correctly, is something that it latches on to (as I will put it): in the present example it is the box . . . Thus I may ask ‘What’s that figure standing in front of the rock, a man or a post?’ and there may be no such object at all; but there is an appearance, a stain perhaps, or other marking of the rock face, which my ‘that’ latches on to. The reference and what ‘this’ latches on to may coincide . . . But they do not have to coincide, and the referent is the object of which the predicate is predicated where ‘this’ or ‘that’ is a subject” (“The First Person,” p. 143).
70. This is an additional condition on the use of expressions of the form *S believes of<sub>strong</sub> t that Φ(it)*, beyond those indicated a few paragraphs back.
71. See McDowell’s discussion of this sort of case in “Singular Thought and the Extent of Inner Space.”
72. McDowell, following Evans, argues persuasively that he was not, in “De Re Senses” and “Singular Thought and the Extent of Inner Space.” He would restrict the transparency claim to individuative omniscience concerning senses grasped *together*, in one thought. This does not exclude false beliefs of identity or nonidentity, but McDowell would insist that the man who does not know that Hesperus is Phosphorus is not confused about which object his Hesperus-

- beliefs are about: he knows that full well—they are about Hesperus. The effect of this view is achieved here as well, but the interpretation of demands for transparency is different.
73. "Eines des folgenreichsten," in the first paragraph of "ÜSB." Recall the discussion below at 7.5.2.
  74. Although as usual the discussion here is limited to singular terms, expressions in other grammatical categories can be used in the strong *de re* way—sortals and predicates prime among them. The account offered here generalizes straightforwardly to these cases.
  75. Thus in accord with the phenomenalist methodology endorsed here, what matters is not who the name-user would treat as an *expert* with authority over the use of the name (one who knows who or what the name refers to), or even how it is determined who *ought* to exercise such authority, but rather what it is for a scorekeeper to *attribute* such authority to someone.
  76. Saul Kripke, "A Puzzle about Belief," in *Meaning and Use*, ed. A. Margalit (Dordrecht: D. Reidel, 1979), pp. 239–283, reprinted in *Propositions and Attitudes*, ed. N. Salmon and S. Soames (New York: Oxford University Press, 1988), pp. 102–148 (page references are to this version).
  77. *Ibid.*, p. 130.
  78. *Ibid.*, p. 122.
  79. *Ibid.*, p. 132.
  80. *Ibid.*, p. 134.
  81. *Ibid.*, pp. 112–113.
  82. *Ibid.*, p. 113.
  83. This issue is masked rhetorically by the structure of the piece. The question would arise evidently and urgently if the 'Paderewski' case were discussed immediately after the disclaimer about ambiguity in the statement of the disquotational principle. In fact the two are separated by a long detour through the (ultimately irrelevant) bilingual cases, in which translation is also involved.
  84. Kripke, "A Puzzle about Belief," p. 111.
  85. *Ibid.*, p. 112.
  86. *Ibid.*, p. 111.
  87. *Ibid.*
  88. Any sort of belief, whether strong *de re* or not, can be ascribed in either the *de dicto* or the *de re* style. The *de re* form of ascription of strong *de re* beliefs is derived from the *de dicto* form of ascription of such beliefs in just the same way that *de re* specifications of the contents of ascriptions are formed from *de dicto* ones in the case of ordinary beliefs. (Which is not to say that on a case-by-case basis, the ascriber must be in a position to *produce* an underlying *de dicto* ascription in order to be entitled to the corresponding *de re* specification.)
  89. Kripke, "A Puzzle about Belief," p. 111.
  90. This social dimension may remain implicit, as when an interpreter looks at some nonscorekeeping system from the outside and takes or treats it as having representationally contentful states—that is treats it as having states that express a doxastic point of view that can in principle be contrasted and combined with our own in the way expressed by *de re* ascriptions. It is in this spirit that one can interpret the illumination of a region of the alarm panel as representing the front door as being open. See Section II, Subsection 4 above.
  91. The Dummettian phrase "circumstances and consequences of application" used

in Chapter 2 has now been superseded by the full scorekeeping account of the force or pragmatic significance characteristic of various speech acts—paradigmatically assertion—that was offered in the following chapter. There the notion of *appropriateness* of circumstances and consequences of application is further articulated in terms of the deontic distinction between commitment and entitlement. A finer-grained account of the relation between circumstances and consequences is then offered in terms of the three notions of permissive, or entitlement-preserving, inference; commissive, or commitment-preserving, inference; and incompatibility, linking entitlement and commitment. These further developments do not make a difference to the point being pursued in the text here, however, so the original phrase can continue to be used as a sort of shorthand.

92. Recall the distinction offered above in 1.2–3 between *simple* regularity theories and those that invoke regularities concerning the practical *assessment* of performances by sanctions. The gerrymandering argument shows that only in the latter form can social regularity theories make sense of the notion of incorrect performances, even for individuals.
93. This is, attributing it to oneself *as* oneself; see Section V, Subsection 2.
94. "Thought and Talk," in *Inquiries into Truth and Interpretation* (New York: Oxford University Press, 1984), pp. 169–170.
95. Indeed, in the same essay Davidson implicitly connects this point about how the notion of beliefs as having objective truth conditions arises out of the contrast of doxastic perspective between interpreter and interpreted with what is expressed by ascriptions specifying the content of the attributed belief in the *de dicto* and *de re* styles. He says "the general and not very informative reason" why the attribution of any propositionally contentful intentional states (what he calls "thoughts") implicitly appeals to a situation in which *speech acts* are interpreted by another is that "without speech we cannot make the fine distinctions between thoughts that are essential to the explanations we can sometimes confidently supply." Those distinctions turn out to be just the ones expressed in the regimentation employed here by *de dicto* and weak *de re* ascriptional locutions: "Our manner of attributing attitudes ensures that all the expressive power of language can be used to make such distinctions. One can believe that Scott is not the author of *Waverly* while not doubting that Scott is Scott; one can want to be the discoverer of a creature with a heart without wanting to be the discoverer of a creature with a kidney; one can intend to bite into the apple in the hand without intending to bite into the only apple with a worm in it; and so forth. The intensionality we make so much of in the attribution of thoughts is very hard to make much of when speech is not present. The dog, we say, knows that its master is home. But does it know that Mr. Smith (who is his master), or that the president of the bank (who is that same master), is home?" (*ibid.*, p. 163). These are, of course, just *de dicto* and *de re* specifications of the same belief. The dog knows of the president of the bank that he is home, he just does not know *that* the president is home. It was pointed out in Section III that one wants to appeal to the belief that his master is home to explain why the dog is so happy, and to its being a belief *of* the president of the bank (whether the dog knows that or not) in order to explain why one result of the dog's happiness is that he slobbers on the president of the bank. We want to be able to offer

intentional explanations of both sorts of things the dog does, and this requires content specifications that depend for their intelligibility on our mastery of the scorekeeping practices involved in interpreting the *utterances* of others.

96. Recall that 'claims' as reconstructed here is equivalent to 'is committed to the claim'; it does not require the performance of an overt speech act to that effect.
97. Notice that one reflection of the modal character of incompatibility entailments is that there is no requirement that a content that defeats a conditional claiming such an entailment (by being incompatible with the consequent but not with the antecedent) be *true*—or one the one assessing the conditional endorses. If in the definition

$$p \rightarrow q \text{ iff } \{x: x/q\} \text{ is a subset of } \{x: x/p\}$$

the range of  $x$  were restricted to those endorsed by the scorekeeper assessing the conditional, then conditionals could be nontrivially assessed only by those who endorsed something incompatible with their consequents, and hence someone who is precluded from detaching from those conditionals. The only restrictions on the range of  $x$  are those imposed by the order in which logically complex expressions must be introduced (so that if  $p$  and  $q$  are logically atomic, for instance, conditionals and negated claims are not yet available in assessing the conditional).

98. For present purposes, incompatibility may be assumed to be a symmetric relation; though that condition can be relaxed, doing so does not materially affect the demonstrations offered here.
99. For instance, if  $p^\alpha$  is  $\{p\}|S^\alpha$  claims that  $p$ , then  $p^\alpha$  will be incompatible with  $\sim(S^\alpha$  claims that  $p^\alpha)$  because the latter negation is incompatible with a substitution instance of the former quantificational claim. The recipe below will produce a  $q^\alpha$  that defeats the conditional (i) for any *nonascriptional*  $p^\alpha$ .
100. One might quibble here that whether or not the restricted quantifier over believers includes ' $\exists x D x$ ' as a substituend is part of what is at issue in the case. The claim that  $(S)|S$  claims that  $p^\alpha$ ) entails  $(\exists x D x$  claims that  $p^\alpha)$  does turn on this issue, but the claim that these are not incompatible (which is all the proof appeals to) patently does not.
101. These issues are explored further in the author's "Truth and Assertibility," *Journal of Philosophy*, 73, no. 6 (March 1976): 137–149.
102. Nor, of course, have I undertaken incompatible commitments by these ascriptions, but that would be true even if I were ascribing to  $S$  commitment to incompatible claims. It remains true that incompatible claims are not attributed to  $S$ , even if I indicate that  $S$  attributes to me the use of the first-person pronoun:

$$S \text{ claims } [\text{he}^{(I)} \text{ does not claim that } p] \text{ and } S \text{ claims } [p]$$

are still not ascriptions to  $S$  of incompatible commitments.

103. Of course the original problem *does* arise if what I attribute to  $S$  is ascribed rather by my endorsement of:

$$S \text{ claims that he}^{(I)} \text{ does not claim that } p, \text{ and } S \text{ claims that } p.$$

But given the account offered earlier of the use of 'I' and of quasi-indexicals anaphorically dependent on it, it is clear these ascriptions do not involve his attitudes toward claims with the same contents as mine. This can be seen as

well in the fact that *S* cannot vindicate his entitlement to the claim expressed in his mouth by [I do not claim that *p*] by deferring ascriptionally to my assertion of [I do not claim that *p*.]

104. W. V. O. Quine, "Quantifiers and Propositional Attitudes"; McGinn, "The Structure of Content."

## 9. Conclusion

1. *Principia*, First Scholium to the Definitions, p. 11 in Motte's translation of 1729 (Berkeley: University of California Press, 1934, 1962). These words actually appear as the antecedent of a conditional, but since Newton immediately goes on to endorse its conclusion, it is clear that he endorses this claim. (Of course, for him, as for others of his time, this dictum applies only to words, not to the contents of the concepts they express.)
2. Cf. *Critique of Pure Reason*, A51/B75. What Kant actually says is that thoughts without content are empty, and intuitions without concepts are blind.
3. At least as far as the doctrines of the first two Critiques are concerned.
4. See John McDowell's "De Re Senses" (*Philosophical Quarterly* 34, no. 136 [July 1984]: 283–294), in which this strand of thought is discerned in Tyler Burge's "Belief De Re" (*Journal of Philosophy* 74 [1977]: 338–362), which presents very clearly views that are endorsed by many other thinkers.
5. Kant distinguishes the first two ways of characterizing the distinction between intuitions and concepts as "logical," in contrast to the "metaphysical" way of characterizing that distinction, which appeals to receptivity and spontaneity. (See for instance Section V of the Introduction to the *Logic*.)
6. As the tradition shows, use of this term is fraught with danger because of what Sellars (in "EPM") calls "the notorious 'ing'/'ed' ambiguity."
7. "On the Very Idea of a Conceptual Scheme," in *Inquiries into Truth and Interpretation* (New York: Oxford University Press, 1984), pp. 183–198.
8. This is the problem with which Hegel opens the Introduction to the *Phenomenology*.
9. For instance in "Concept and Object," in *Translations from the Philosophical Writings of Gottlob Frege*, ed. P. T. Geach, and M. Black (Oxford: Basil Blackwell, 1970), pp. 42–55. For many purposes it is more helpful to consult Frege's notion of *Sinn* than his notion of *Begriff* in order to learn about what is discussed here as conceptual articulation.
10. The point being made concerns contemporary ways of thinking about concepts that Kant inspired, rather than the details of his view. Nonetheless, particularly in the context of the present project, it would be unfair to accuse Kant of confusion on this point. As "double affection" readings of Kant make clear, a major interpretive challenge facing the readers of the first Critique is to make sense of talk of spontaneity and receptivity at the noumenal or transcendental level, even though the causal relations one naturally turns to in such a case are conceived as products of those two faculties, and so remain at the phenomenal level. The general picture presented here about the lessons to be learned from Kant suggests that the answer is not far to seek: spontaneity and receptivity are normative notions for Kant, not causal ones. Roughly, talk of spontaneity is talk about responsibility for, or authority over, something. Talk of receptivity is talk

about responsibility *to*, or being subject to the authority *of*, something. These notions accordingly line up with:

—The ‘I think’ that accompanies all representations as a purely *formal* feature—not part of what is represented or the content expressed by the representing, but of the form of the representing as such. The root notion is the taking of responsibility *for* a claim or judgment. It is because judgments are the minimal unit of responsibility or accountability that they are the minimal unit of experience for rational beings. This is also why spontaneity—that is, the understanding, the faculty of judgment—traces back to the Transcendental Unity of Apperception (the locus of coresponsibility for claims), which in turn underwrites the categories.

—The ‘object = X’ that implicitly accompanies all judgments of experience, as a purely formal feature of the representing. This is the dimension along which such judgments are responsible for their correctness *to* something represented—which is part of its being a representing at all: its representational purport.

That both of these purely formal features are necessary for representation at all is what makes the refutation of idealism work. Notice that in this Kantian dualism, the formal two-sides-of-one-coin relation of responsibility *for* and responsibility *to* as essential to discursive representation, *is* carried over into (indeed is essential to) the scheme presented in this work.

11. One modern version of Kant’s contrast between spontaneity and receptivity takes the form of a contrast between the repeatable, purely qualitative or descriptive terms whose deployment is entirely within the “cognitive world” of the thinker, on the one hand, and the unrepeatable, contextually determined indexical expressions whose significance outruns what is within the grasp of the thinker and is the ultimate source of constraint on thought, on the other. Here are two representative passages: “A *de re* belief is a belief whose correct ascription places the believer in an appropriate nonconceptual, contextual relation to objects the belief is about”; and “A sufficient condition for a belief context to be *de re* . . . is for it to contain an indexical expression used deictically” (Burge, “Belief *De Re*,” pp. 346, 347).

Part of the responsibility for this line of thought lies in a common misreading of the nature of Kripke’s project in “Naming and Necessity” (in *Semantics of Natural Language*, ed. G. Harman and D. Davidson [Dordrecht: D. Reidel, 1972]). While Kripke was addressing exclusively the semantics and epistemology of proper-name usage, people often derive consequences for semantics and epistemology generally. His distinction between descriptive and causal-historical approaches to proper names works well enough in the context of his actual project. But disaster results if one takes him as having shown, or even suggested, that there are two general styles of theory about how language or mind encompasses the world—namely descriptive and causal-historical or contextual ones. For that distinction patently does not extend to predicates. No one was ever a descriptivist about predicate denotation, because of the regress such a view would obviously involve. Everyone, even the ones who wanted to appeal to predicates via descriptions to account for proper names, always thought that at least the basic predicates achieved their reference in a different way, one that involved contact with the properties. Burge, for instance, says nothing about

how he conceives our relation to the semantic correlates of the predicates appealed to in the descriptions that underwrite *de dicto* (purely conceptual) attitudes for him. It is hard to see how his story can be extended to include them, for he seems to be working with a picture that allows only descriptive or demonstrative contact with the world, and he cannot treat predicates in either way without either falling into the regress or conceding that there cannot be attitudes that are *de dicto* in his sense.

12. Only “roughly” because at the next level of analysis, the work done by the undifferentiated notion of propriety (appealed to in talk of “circumstances of *appropriate* application” and “*appropriate* consequences of application”) is taken over by a theoretical idiom that distinguishes the deontic statuses of commitment and entitlement. Then the scorekeeping significance that must (in context) be determined by conceptual content is articulated according to commitment-preserving inferences, entitlement-preserving ones, and incompatibility inferences (which involve both deontic statuses). See Chapter 3.
13. Notice that this definition says *what* it is to take one tokening to be anaphorically dependent on another, not *when* it is appropriate to do so. Thus it does not purport to answer the linguist’s fundamental question about anaphora—namely what it is that determines what is the proper antecedent of any given anaphorically dependent expression. It purports to answer only the philosopher’s fundamental question about anaphora—namely how to understand the relationship at issue in the linguist’s question.
14. According to the theory expounded in Chapter 5, this eight-word phrase is itself a complex, anaphorically indirect definite description, which is anaphorically dependent on the original tokening of ‘that’.
15. In the classical cases. In fact most of what appear syntactically as definite descriptions are not semantically definite descriptions, but function instead as anaphoric dependents.
16. ‘Claim’ in the sense of what is claimed, not the *claiming* of it. See 5.5 above.
17. To show why this *must* be the case, why the conceptual articulation of sentences must take a form analyzable in terms of the substitution-inferential roles characteristic of singular terms and predicates, is the burden of Chapter 6.
18. So it is here that the notion of *freedom* first finds application. The notions of will and choice are not available independently in order to make sense of this notion of freedom. Rather they are themselves to be understood in terms of the practical attitude of acknowledging norms as binding (acting according to a conception of a rule).
19. Though, as explained in Chapter 1, nondiscursive creatures can treat each other as bound by norms that are not inferentially articulated in the way constitutive of concepts.
20. Of course (as a helpful reader put it) it is <sup>s</sup>norms all the way down<sup>s</sup> only if one digs in the right direction. For if what is to be explained is, say, the fact that the planets have elliptical orbits (the truth of what is claimed, rather than what we are doing in claiming it), the explanation does not need to appeal to anything normative. For that would have been a fact (what is claimed would have been true), even if norm-instituting, claim-making creatures had never evolved. Being entitled to say both of these sorts of things is a prime criterion of adequacy of the present account. See 1.5.3; 5.5.3–4; 8.6.3–5.
21. According to the account offered in Chapter 7, making explicit the sense in

which deontic statuses or discursive norms exist requires specifying the use of a class of *canonical designators* corresponding to this realm—expressions that can play for normative objects the role played for physical objects by spatiotemporally definite descriptions, for natural numbers by successor numerals, and so on. Although such a story is not filled in here, the raw materials for it are provided by the account of scorekeeping with deontic attitudes (in Chapter 3), of the way in which such attitudes can be made explicit using normative vocabulary (in Chapter 4), and of the use of singular terms and the formation of definite descriptions (in Chapters 6 and 7).

22. As a helpful reader points out, the terms in which one says that the umpire's taking the pitch to be a strike caused a fight are not purely causal, according to the theory put forward here. A more careful statement would be that some event that constituted the umpire's taking the pitch to be a strike caused a fight. The "constituted" here is normative talk about the pragmatic or scorekeeping significance of a performance. (Compare the way in which moving a pen in certain ways can constitute acknowledging a commitment to pay the bank a certain sum every month.)
23. Chapter 3 (Section IV) discusses how scorekeeping proprieties are codified in contents in the context of a theory of speech acts.
24. "Thought and Talk," in *Inquiries into Truth and Interpretation* (New York: Oxford University Press, 1984), p. 156.
25. This sense in which simple intentionality is derivative from the original intentionality attributable to language users should not be confused with the sense in which the intentionality of linguistic performances and expressions is derivative from that of the community of language users. See 1.6.2.
26. As indicated below, it should be clear by this point in the exposition that this is compatible with the community being ignorant or mistaken about what their practice has made their words mean (what concepts those words have come in practice to express).
27. Recall that it was applying this stance to internal intentional interpretation that makes possible the reconstruction of the hybrid deontic attitude of attributing knowledge, compounded in social-perspectival terms out of the *attribution*, to the one taken thereby as a knower, of commitment (belief) and entitlement (justification), and the *undertaking*, by the one attributing knowledge, of commitment to that same content—since undertaking a doxastic commitment of the sort acknowledged by asserting is taking-*true*. Again, it permits the epistemological via media between justificatory internalism and reliabilist externalism about entitlements—namely, a perspectival reliabilism that turns on an account of the inferences endorsed by scorekeepers who *attribute* reliability (and hence entitlement; see 4.3.3).
28. Giving and asking for reasons in discursive games of this sort belongs in this regard in a box with baseball, rather than with purely formal games such as chess. Baseball cannot be played except with a ball of a specified size and composition, on a field of a specified size and composition. By contrast, chess is an empirically and practically *hollow* or abstract game; it can be played with pieces of any size and composition, on a board of any size and composition, so long as the pieces are distinguishable and interpreted as standing and moving in appropriate formal relations to each other. Solid discursive practices incorporate nonlinguistic things in them (are *corporeal*). In the same way, the practice in

which a performance can have the significance of hitting a home run incorporates objects.

29. For instance in "The Conditions of Thought," in *Le Cahier du Collège International de Philosophie*, (Paris: Editions Osiris, 1989), pp. 165–171, reprinted in *The Mind of Donald Davidson*, ed. J. Brandl and W. Gombocz (*Grazer Philosophische Studien* 36 [1989]), pp. 193–200; "Meaning, Truth, and Evidence," in *Perspectives on Quine*, ed. R. Barrett and R. Gibson (Oxford: Basil Blackwell, 1990), pp. 68–79; and "The Myth of the Subjective," in *Bewusstsein, Sprache, und die Kunst*, ed. M. Benedikt and R. Burger (Edition S Verlag der Österreichischen Staatsdruckerei, 1988), pp. 45–54, reprinted in *Relativism: Interpretation and Confrontation*, ed. M. Krausz (Notre Dame, Ind.: University of Notre Dame Press, 1989).
30. "Two Dogmas of Empiricism," Essay 2 in *From a Logical Point of View* (Cambridge: Harvard University Press, 1953).
31. "Concepts As Involving Laws, and Inconceivable without Them," in *PPPW*, pp. 87–124. More recently, Dretske has taken a similar line, in his *Knowledge and the Flow of Information* (Cambridge: MIT Press, 1981).
32. Cf. Quine's remark in "Two Dogmas": "Meaning is what essence becomes when it is divorced from the object of reference and wedded to the word."
33. This notion is introduced at 8.1.1 and again at 8.1.3. Iteration of propositional attitude ascriptions is discussed in the Appendix to Chapter 8.
34. This use of 'should' has been explained (in 4.2.8) as making explicit the attribution of a practical discursive commitment, and so the endorsement or undertaking of commitment (on the part of the one employing 'should' or to whom its use is attributed) of a pattern of practical inference. 'Should' in its "all-in" sense, as opposed to its *prima facie* sense, is the third-person, or attributing, form whose first-person or acknowledging, form is 'shall'.
35. In the terms of the *Phenomenology*, the phenomenal attitudes whose development is being considered have been developed to the point where they coincide with the phenomenological view we have adopted all along in considering them. Hegel's alarming term for this sort of explicit interpretive equilibrium is "Absolute Knowledge."
36. The social 'we' is constructed out of perspectival *I-thou* scorekeeping or interpretive relations. At this point no distinction of kind remains separating the internal *I-thou* relations between the scorekeeping attributor of a commitment and the one to whom it is attributed, on the one hand, and those between the external attributor of scorekeeping practices and those to whom they are attributed, on the other.
37. Notice that to say this is to invert the Davidsonian order of explanation, which discusses conversation within a language as a special case to be modeled on interpretation across languages.

---

# Index

---

< > type designation, 314  
// token designation, 314  
[ ] corner quotes, 688n25, 704–705n24  
` ` scare quotes, 545–547, 588–590, 672n19

'about', xvii, 138, 500–503, 520, 547–548, 590. *See also* aboutness  
aboutness, 6, 70, 279, 280, 306, 333, 336, 514, 518, 528, 547–548, 568, 583, 622, 647. *See also* representation  
and *de re* ascriptions, 391, 500, 502, 503, 547–548, 584  
and objectivity, 138, 530, 649, 672n18  
strong and weak, 584–586, 706n45  
abstraction, 419–422, 445, 614, 695n34  
abstract objects. *See* objects, abstract  
acknowledgment, 31–32, 35, 37, 43, 50–52, 202, 262–266, 525, 541, 542, 552–553, 629, 630, 633, 680n61. *See also* attributing, deontic attitudes; undertaking  
and attribution, xiv, xx, 55, 554, 646  
causal efficacy of, 259–262, 271, 596  
of commitments, 259, 262–266, 271, 596, 704n21  
explicit/implicit, 32, 624, 629, 639  
and undertaking, 194, 554  
action(s), xv, xx, 7, 8, 16, 69, 131, 142, 167, 230, 239, 243–245, 254–256, 290, 332, 507, 527–529, 614, 645. *See also* causes; deontic scorekeeping; making-true; normative vocabulary; practical

commitments; practical reasoning; reliable differential responsive dispositions  
and cognition, 5, 200, 276, 289, 618, 665n33  
as language exit transition, 5, 221, 233–235, 294  
and linguistic practice, 155, 232  
and perception, 233–238, 261, 335, 336  
and practical reasoning, 231, 245–249, 255–256, 678n29  
and rationality, 230–233, 244–245  
and reasons, 7, 56, 171, 194, 203, 243–245, 249–256, 262  
and trying, 290–291, 295, 527–529  
adicity, 369, 385, 394, 408  
agentives, 556, 702n14  
agent semantics. *See* semantics, agent  
akrasia, 269–271. *See also* practical reasoning  
'A' language, 555–559, 708n65. *See also* 'I'  
anaphora, xvi, 285, 301–303, 306, 311–331, 438, 456, 462, 467, 471, 473, 486–494, 535–536, 559–567, 574, 580, 583, 585, 589, 622, 623, 683n35, 684n38, 698n80, 700n100, 705n30, 714n13.

**anaphora (*continued*)**

- See also inference, substitution, and anaphora; substitution  
 ascription-structural, 516, 609–611, 701n7  
 in ascriptions, 536, 539, 577–578  
 as asymmetric token-recurrence structures, 455, 457, 467, 473, 490–491  
 and communication, 458, 459, 485–486, 503  
 and coreference, 306, 314, 485–486, 588  
 and deontic scorekeeping, 432, 460, 488, 494, 592, 633  
 expressive role of, 458, 467, 473, 588  
 inheritance of, 283, 303–304, 325, 391, 472–473  
 and inheritance of substitutional commitments, 198, 283, 325, 432, 455, 472, 473, 539  
 presupposed by deixis, 458, 462–468, 473, 511, 573, 585, 621  
 and quantification, 321, 491, 493  
 and recurrence, 309–310, 432, 453, 464, 490–492, 498, 621, 685n40  
 and reference, 306, 307, 322  
 and rigidity, 468–472, 698n80  
 anaphorically indirect definite descriptions, 305–307, 310–323, 464, 512, 547, 588, 589, 714n14  
 anaphoric antecedents, 301–305, 313–315, 450, 456, 459, 460, 486, 491–494  
 anaphoric chains, 307–309, 313–315, 319, 460, 462, 469, 470, 491, 493–494, 581–582  
 anaphoric commitments, 456, 472, 495, 583, 650, 706n33. *See also* substitution; token-recurrence structures  
 anaphoric dependents, 305–307, 309–331, 456, 458, 467, 470, 494, 574, 579–583  
 anaphoric initiators, 308–310, 458, 459, 462, 573–574  
 animals, 7, 33, 47, 83, 629, 691n42. *See also* intentional stance, simple intention states of, 150, 152, 155  
 Anscombe, G. E. M., 255–256, 558, 707n51, 708n69  
 Aoun, Joseph, 699n85  
 ascription(s), xx, 502, 506, 516, 530–531, 533–535, 537–539, 541–544, 547–548, 550, 566, 575–579, 581, 588–590, 595–596, 605, 609–613, 629, 703n19, 704n23, 711n102. *See also* *de dicto* as-

- criptions; *de re* ascriptions; explicit; pragmatics; representation  
 as attribution and undertaking of commitments, 504–505, 531–533, 544, 640  
 direct discourse, 533–535, 542  
 of essentially indexical beliefs, 559–567  
 expressive role of, 503–504, 543, 588–589, 591, 608, 641, 679n42  
 indirect discourse, 534–535, 537, 539–540, 542  
 iterated, 545, 549, 608–613, 701n6, 702n13  
 paratactic theory of, 535–538, 577  
 of propositional attitudes, 279, 453, 485, 498–501, 530  
 regimentation of, 112, 501–503, 505–506, 516, 535, 540, 542–544, 546, 609, 613, 629  
 ascriptionally indirect definite descriptions, 547, 565, 590  
 assertibility, 121, 124, 299, 604  
 assertion(s), xiv, xxi, xxiii, 82, 156–159, 167–175, 179, 186, 190–191, 200–203, 278, 329, 335, 531, 639, 667n67, 675n45. *See also* deontic scorekeeping; doxastic commitments; giving and asking for reasons; linguistic practice; propositional contents; sentence(s); taking-true  
 authority/responsibility, 171–175, 179, 229, 234, 384, 532  
 bare, 228  
 and belief, 153–154, 194, 200, 229, 230  
 and commitment, 142, 179, 191, 200, 230  
 and communication, 156, 175–176, 221 and entitlement, 179, 180, 191, 200, 221, 229  
 and inference, 158, 167–168, 171, 339, 367, 687n11  
 and linguistic practice, 156, 158, 167–168, 173–176, 190, 200, 203, 221, 358, 623, 628–629, 637, 686n1  
 and pragmatic significance, 157, 168, 190, 200, 234, 358, 637  
 and truth, 199–204, 226, 229, 232  
 assertional commitments. *See* doxastic commitments  
 assessment, 18, 33, 37–42, 52, 64–66, 78, 236, 516, 526, 528, 614, 658n45, 659n47, 660n52. *See also* norms appraisal/deliberation, 18–19, 287

- of correctness, 29, 32, 52–55, 63, 647  
and sanctions, 34–36, 42–45
- attitudes. *See* deontic attitudes; normative attitudes
- attributing, 196–197, 563, 598, 629, 630, 640, 642, 644, 645. *See also* deontic attitudes
- and acknowledging, xiv, xx, 55, 554, 646
- self-, 680n61
- and undertaking, 62, 161–163, 554, 596–598, 677n11
- Austin, J. L., 209, 288
- authority, 10–11, 50–52, 160, 163, 175, 212, 217, 226, 234, 238, 242, 677n24.  
*See also* inheritance of entitlements; justification; responsibility
- of noninferential reports, 214–221, 223–227, 234
- of norms, 50–52
- and responsibility, xii, 162–163, 165, 242, 673n24
- of testimony, 39–41, 175, 205, 234, 709n75
- autonomy, 20–21, 50–52, 171, 216, 265, 277, 293, 295, 465, 631
- auxiliary hypotheses, 92, 121–122, 475, 477, 488, 490, 517, 541, 596, 633, 635, 646. *See also* perspectives
- and communication, 139, 480–481, 511, 632
- and inferential significance, 139, 475, 478, 488, 504, 506, 509, 587, 633, 635
- and social perspectives, 519–520, 598, 608
- avowals, 194–196, 540, 577. *See also* assertion(s); belief(s)
- Bach-Peters sentences, 493–494
- Baier, Kurt, 172
- baptism, 580
- bare assertions. *See* assertion(s), bare
- barn facades, 209–212
- Bedeutung, 80–81, 111, 352
- behavior, 8. *See also* action(s); speech acts
- belief(s), xv, 5, 13–14, 17–18, 56, 142, 146, 152–157, 194–196, 201–202, 215, 228–230, 232–233, 250, 290–291, 515, 552–559, 573–579, 594–595, 656n18. *See also* assertions; doxastic commitments; intentionality; interpretation
- avowals of, 146, 195–196, 227–229, 540, 577
- de dicto/de re* [*see de dicto ascriptions; de re ascriptions*]
- essentially indexical, 551–552, 559–567
- mere, 229
- notional/relational, 547–548, 550
- object-dependent, 58, 551, 567–568, 570, 578, 583
- strong *de re*, 547–552, 555, 561–562, 566, 567, 581, 582, 585, 708n70, 709n88
- two senses of, 195, 507 (*see also* commitment(s), acknowledged/consequential)
- belief-forming mechanisms, 207–208, 210, 217
- 'believes', 116, 228, 261, 352, 498. *See also* belief(s)
- Belnap, Nuel D. Jr., 125, 300–301, 702n14
- Bennett, Jonathan F., 147, 668n85
- Boghossian, Paul, 326–327
- Boyd, Richard, 481
- Bradley, F. H., 375
- Brentano, Franz, 68, 70–71
- Burge, Tyler, 551, 566, 713n11
- Camp, Joseph, 300–301
- canonical designators, 112, 440, 442–449, 550, 569, 696n61, 706n41, 715n21.  
*See also* singular terms
- and existential commitments, 442–447
- Carnap, Rudolph, 96, 123, 215
- Carroll, Lewis, 22, 100–101, 206
- Castañeda, Hector-Neri, 263, 552, 563–564, 566
- categories, 335, 361, 362, 370–372, 404, 406
- causes, 11–14, 209–211, 259–262, 271, 427, 428, 429, 617, 620, 626, 687n2. *See also* action(s); norms; perception
- c-command, 700n106
- censure, 34, 37, 43
- certainty, 9–11
- challenges, 178, 192–193, 238
- Chastain, Charles, 307–309, 311, 456, 491, 683n32, 699n84
- Chomsky, Noam, 365–366
- circumstances and consequences of application, xiii, 18, 28, 40, 51, 89, 98, 117–131, 136, 159, 162, 182, 243, 331–332,

- circumstances and consequences of application (*continued*)  
 372, 383, 419, 421, 432, 433, 482, 541,  
 600, 618, 631, 640, 668n82, 674n41,  
 709n91, 714n12. *See also* concepts; inferentialism  
 and inference, 118–120, 126, 129, 131,  
 206, 225  
 'claims', 539–542, 711n96, 714n16. *See also* 'says'  
 claims, xv, 96, 112, 141, 153–157, 167,  
 229, 276, 291, 330, 453, 536, 586, 601,  
 619, 625, 641, 676n3. *See also* assertion(s)  
 and facts, 327–331, 622, 624, 625,  
 704n19  
 classification, 33, 35, 79, 85–86, 88–90,  
 666n35  
 and concepts, 85–89, 614–618, 623  
 codification, 107–108, 145, 267, 385, 397,  
 402, 403  
 cognition, 5, 200, 276, 289, 363, 614, 618,  
 665n33 *See also* deontic scorekeeping;  
 doxastic commitments; knowledge; sapience  
 commands, 51, 64–66, 172, 270  
 commitment(s), 8, 55, 74, 157–161, 163–  
 166, 168–172, 200, 201, 203, 209, 233,  
 250, 275, 474, 481, 583, 627, 649, 650,  
 674n43, 680n56. *See also* anaphoric  
 commitments; deontic attitudes; deontic  
 scorekeeping; deontic status(es);  
 doxastic commitments; existential  
 commitments; expressive commitments;  
 inferential commitments; practical  
 commitments; recurrence  
 commitments; simple material substitution-inferential  
 commitments (SMSICs); substitutional commitments  
 acknowledged/consequential, 193–197,  
 259–262, 270, 596, 633, 646, 704n21  
 attributing, 161–163, 205, 596–597, 646  
 and entitlement, xvi, 55, 159–161, 165,  
 245, 606, 675n44  
 making explicit, 334, 402, 505  
 undertaking, 161–163, 205, 596–597  
 communication, 473–475, 477–483, 485,  
 528, 566–567, 588, 635, 636, 644–647,  
 716n37. *See also* information; linguistic  
 practice  
 and anaphora, 282, 458, 459, 477–482,  
 485–486, 503, 528  
 and assertion, 156, 174–176, 221, 357  
 and auxiliary hypotheses, 139, 480–481,  
 511  
 and contents, xxiii, 158, 169–170, 474,  
 478–481, 513–517, 562, 590, 633–  
 636  
 and deontic scorekeeping, 156, 158, 169–  
 170, 282, 474, 478, 480–481, 485, 496,  
 588, 635  
 and *de re* ascriptions, 508, 513–517, 519,  
 547, 701n7  
 and inheritance of entitlement, 175–176,  
 204, 205, 217, 226  
 and perspectival character of content,  
 485, 509–510, 635–636, 647  
 and strong *de re* beliefs, 561–562, 566  
 community, 594, 600, 630. *See also*  
*I-thou/I-we* sociality; we  
 and explicit discursive scorekeeping  
 stance, 54, 639, 642–643  
 membership, 4, 39–40, 67, 643, 659n51,  
 660n52  
 and norms, 37–43, 67, 599–600, 639, 649  
 components, 298–299, 338, 343, 367, 376  
 compositionality/decompositionality, 351,  
 354–358, 361, 367  
 computers, 7  
 concepts, 8, 10–11, 18, 46, 79, 85–86, 89–  
 91, 109, 113, 125–130, 207–208, 293,  
 484, 614, 616–620, 622–624, 634–636,  
 694n31. *See also* circumstances and  
 consequences of application; content(s);  
 grasping; inferentialism  
 conceptual change, 125, 127, 375, 402  
 conceptual schemes, 616, 618, 645  
 criticism of, 124, 126  
 dualistic conceptions of, 614–618  
 empirical, 119–120, 432, 458, 473  
 form/matter, 616–619, 622  
 inferential conception of, 87–91, 331,  
 551, 614–623, 633–635  
 and objectivity, 37, 109, 124, 126, 592–  
 597, 614, 616, 624, 633, 636  
 practical, 9–10, 293, 616–618, 620, 622,  
 623  
 sharing, 587, 617, 620, 623, 631–636  
 theoretical, 9–10, 91, 616–617, 622  
 conceptual content(s), 4, 13, 107–108, 133,  
 225, 282, 426, 475, 582–583, 591, 636.  
*See also* conferring content; content(s);  
 inferentialism  
 and discursive practice, 13–14, 107–108,

- 453, 485, 562, 590, 591, 614, 636, 645, 686n55  
in Frege, 95–96, 281, 348, 355  
inferential conception of, 95–96, 98, 102–106, 115, 130–132, 478, 485  
objectivity of, 63, 497, 529, 552–554, 593–594, 601–607, 630, 636  
perspectival character of, 485, 586–592  
and representation, 54, 140, 280, 583, 601–607, 630  
conditionals, 22, 103–104, 109, 112, 121–122, 247–248, 267, 352, 382, 393, 395, 396, 404, 414, 474, 498, 587, 666n48, 691n42, 697n70. *See also explicit; logical vocabulary*  
and embedding, 298–299  
expressive role of, xix, 108–110, 126–127, 383, 385, 395, 401–403, 498  
and inference, 98, 100, 101, 351, 352, 380, 381, 474, 602–603, 640, 693n56  
two-valued, 111–116, 344, 353, 530, 666n48  
conferring content, xv, 61, 64, 77, 91, 107–108, 115–117, 133, 137, 140, 145–147, 149, 151, 156, 261, 263, 592, 642, 663n89. *See also content(s); practice(s)*  
conceptual, 167, 169, 625, 627, 630, 632, 646, 649  
propositional, 7, 61, 63, 77, 141–142, 153, 159, 221, 275, 277, 284, 623, 638, 645  
conformism, 34  
conjunction, 115, 118, 352  
connectives, 117–130, 345, 347  
consciousness/self-consciousness, 85–87, 559, 614, 643, 644. *See also classification; sapience; sentience*  
consequences of application. *See circumstances and consequences of application*  
conservativeness, 123–125, 127, 129, 668n82  
constraints, 331–333, 529, 614, 617, 620, 621, 713n11, 714n11  
content(s), 67–71, 127, 137, 141–142, 145–147, 151, 153, 157, 167–175, 186–191, 199, 212, 330, 339, 343, 346–350, 352, 354, 359, 364, 368, 370–372, 374, 386, 401, 481, 516, 562, 605. *See also conceptual content(s); conferring content; deontic scorekeeping; explicit; propositional content(s); representational contents; semantics*
- and circumstances and consequences of application, 118, 129  
and discursive practice, 6, 12, 77–79, 199, 334, 359  
empirical (cognitive), 6, 212, 221, 225, 234, 295  
and force, 111, 186–190, 298–299  
freestanding, 338–340, 345, 348, 353, 354, 356, 358, 359  
inferential conception of, 89–90, 94, 130–132, 134–136, 186–190, 336, 342, 345, 347–348, 351, 353, 354, 374, 482, 485  
ingredient, 122, 338–340, 342, 345, 355, 356, 358, 359  
narrow/wide, 526, 703n18  
objective, 18, 136–139, 151, 153, 156, 595–596  
practical, xv, 221, 234, 295  
and representation, 70–72, 130–132, 135–138, 151, 153, 156, 517  
content-specifying expressions, 524, 646–647  
in ascriptions, 504–505, 588, 591, 608–611  
*de dicto/de re*, 513–514, 516–517, 535, 543–544, 547–548, 584–585  
and social perspectives, 520, 595–596  
contexts, 316, 318, 342, 344, 347, 398–399, 483, 617. *See also deixis; deontic scorekeeping*  
extensional/intensional, 279, 281, 392  
heterogeneous/homogeneous, 344–346, 350, 356  
inferentially inverting, 381–382, 393, 394, 397, 403, 693n56  
sentential, 343–344, 346  
contract, 49–51, 242, 656n10  
conventions, 232, 233, 670n3  
coreference, 306, 308–309, 314, 318, 468, 485–486  
corporeality, 332, 631–632. *See also solidity of practices*  
correctness, 17–18, 32, 207–208, 278, 280, 291, 594–595, 614, 627–628, 632, 637, 666n41, 672n18. *See also norms; pragmatics, normative; normative vocabulary*  
assessment of, 9–11, 13–14, 29, 32, 52–55, 63, 647

- correctness (*continued*)  
 and practice, 22, 25, 32, 625, 628  
 and regularities, 27–28, 62, 207–208,  
 212  
 and rules, 8, 18–19, 24, 64–66, 628  
 and taking as correct, 25, 32–34, 37, 52–  
 53, 63, 291, 626  
 correspondence, 291, 326, 330–333  
 cotypicality, 221, 308–310, 319, 469, 533,  
 535–536, 579, 622  
 counterfactuals, 484, 574, 634–635, 648,  
 697n70  
 counting, 438, 444  
 criteria of identity, 416, 437, 438
- Davidson, Donald, 97, 150–152, 231, 232,  
 246–247, 255–256, 426, 523, 524, 614,  
 616, 622, 710n95  
 on ascriptions, 535–539, 550, 566, 575,  
 577  
 on interpretation, 15, 155, 262–264, 412,  
 599, 629, 659n50, 670n6, 678n37,  
 699n86, 716n37  
 on practical reasoning, 230, 246–248,  
 253–254, 259, 269, 337, 663n84  
*de dicto* ascriptions, 485, 490, 502–508,  
 526, 529–534, 539–542, 566, 589, 601,  
 646, 703n19, 709n88, 710n95. *See also*  
*also* ascription(s); belief(s); commit-  
 ment(s); *de re* ascriptions  
 and intentional explanation, 522–  
 527  
 and 'that', 506, 598, 608
- deductive inferences. *See* inferences, com-  
 mitment-preserving
- default-and-challenge structure of entitle-  
 ment. *See* entitlement, default-and-  
 challenge structure of
- defeasors, 191, 602–605
- deference, 204, 234, 453. *See also* author-  
 ity, of testimony; communication; in-  
 heritance of entitlements,  
 interpersonal, intracontent
- deferrals, 192–193, 196, 212, 531–532,  
 534, 536, 704n23
- definite descriptions, 298, 316, 386, 387,  
 415, 420, 431, 433, 439, 441, 456, 460,  
 464, 468, 469, 486, 617, 714n15. *See also*  
*also* anaphorically indirect definite de-  
 scriptions; ascriptionally indirect  
 definite descriptions; singular terms
- and anaphora, 305, 308–309, 314, 458,  
 459, 464, 472
- attributive/referential use of, 488, 579,  
 581  
 existence and uniqueness conditions on,  
 415, 434–435, 471  
 and indefinite descriptions, 307–310, 459  
 definitization transformations, 309, 456,  
 459, 491–494  
 deflationism, 325–327, 329  
 deixis, 198, 282, 432, 453, 464, 468, 620.  
*See also* noninferential reports  
 and anaphora, 306, 456, 458, 462, 464–  
 468, 473, 511, 573, 585, 621  
 deliberation, 158, 287, 290  
 demarcation, 3–11, 46, 50–51, 87, 114,  
 135, 200, 644, 645. *See also* rational-  
 ity; sapience; we
- demonstratives, 132, 282, 319, 439, 459–  
 464, 466, 510–513, 550, 563, 582, 617,  
 621, 698n79, 705n30  
 and anaphora, 304, 462, 467, 469, 472,  
 511, 621
- Dennett, Daniel C., 15, 55–62, 73, 99–100,  
 195, 230, 547–548, 550
- denotation, 317, 318, 364, 547–548, 566,  
 706nn39, 45. *See also* reference; repre-  
 sentation
- 'denotes', 318. *See also* denotation
- deontic attitudes, 137, 221, 271, 290, 339,  
 497, 595–597, 599–607, 612–613, 637,  
 645, 649, 677n11. *See also* acknowl-  
 edgment; attributing; deontic score-  
 keeping; deontic status(es); intentional  
 states; normative attitudes; undertak-  
 ing
- and deontic statuses, 165–166, 357, 598–  
 601, 623, 648, 687n11
- hybrid, 187, 202, 220, 228, 297, 521–  
 522, 525, 528–529, 545, 556, 568–569,  
 586, 704n20
- deontic scorekeeping, xiv, xvii, 141–143,  
 166–167, 181–187, 192–193, 202–203,  
 233–234, 242, 260, 264–266, 278, 282,  
 325, 330, 349, 436, 443, 461, 470, 480–  
 481, 495–496, 503, 516, 555–560, 574,  
 584–585, 591, 605, 608, 624, 636, 639–  
 642, 645, 648–649, 696n46, 700n99,  
 714n12. *See also* deontic attitudes;  
 deontic status(es); entitlement;  
*I-thou/I-we* sociality; incompati-  
 bility; linguistic practice; prag-  
 matics
- and anaphora, 432, 460, 470, 487–488,  
 494

- and assertion, 157–159, 167–175, 190–191, 200–203  
and commitments, 157–166, 178–180, 193–196, 237, 243, 263  
and communication, 156, 174–176, 282, 478, 633  
and content, 141–142, 145–147, 167–175, 186–191  
and entitlement, 159–166, 176–180  
and interpretation, 475, 508, 630, 644–645  
model, 166, 168–172, 180–186, 190–191  
and objectivity, 324, 529, 601–607, 627–628, 636  
as perspectival, 185, 332, 488, 590, 602, 604, 627, 649  
and practical reasoning, 230–233, 244, 256, 625, 640  
and pragmatic significance, 142, 167–168, 182–190, 262, 284, 710n91  
and speech acts, 142, 182–193  
*deontic status(es)*, 55, 165–166, 189–190, 201–205, 226, 237, 275, 290, 595–596, 636–637, 649. *See also commitment(s); deontic attitudes; deontic scorekeeping; entitlement; intentional states; normative status(es); norms and deontic attitudes*, 161–162, 165–166, 334, 357, 593, 598–601, 623, 648, 687n11  
and objectivity, 197, 201, 599–607  
*de re* ascriptions, 391, 499–508, 511–515, 525, 542–548, 565, 566, 584, 589, 598, 601, 646, 648, 701n7, 703n19, 709n88, 710n95. *See also ascription(s); de dicto ascriptions; information; representation*  
denotational sense of, 547–548, 566, 706nn39, 45  
expressive role of, 138, 502, 512, 516, 522, 525  
and intentional explanation, 522–527  
and objectivity, 528, 595, 598, 600–601  
and ‘of’, 506  
of propositional attitudes, 187, 279–281, 485, 490  
and representation, xvii, 279, 499–503, 517, 519, 520, 586  
strong, 503, 530, 566, 569–570, 581  
weak, 503, 513, 529–530, 547–552, 566  
Descartes, René, 6, 9–11, 31, 74, 93–94, 279, 614, 623, 655n2, 671n8, 681n5  
designatedness, 340–350, 352, 353, 356–358, 687n8. *See also multivalues; substitution*  
designation, 69, 84, 85, 665n32, 696n61.  
*See also canonical designators; proper names; rigidity; semantics; singular terms*  
desires, 56, 58, 240, 246, 250, 256  
Devitt, Michael, 481  
Dewey, John, 289, 299  
direct discourse ascriptions. *See ascription(s), direct discourse*  
direct reference. *See reference, direct*  
disavowals, 192–193, 670n6  
discursive attitudes. *See deontic attitudes*  
discursive commitments. *See commitment(s)*  
discursive practice. *See linguistic practice*  
discursive scorekeeping. *See deontic scorekeeping*  
discursive statuses. *See deontic status(es)*  
disjunction, 115, 429–430, 434–435, 438, 441–442  
dispositions, 28–29, 35, 42, 45–46, 208, 625, 628, 629, 636, 638. *See also causes; regularities; reliable differential responsive dispositions*  
disquotation, 300–301, 303, 323, 575, 577, 579, 581  
distinction  
acknowledging/attributing, 55, 193–197, 259–262, 270, 554, 596, 633, 646, 649  
action/perception, 7–8, 119–120, 209–211, 233–238, 261, 335, 336  
analytic/synthetic, 145, 345, 358, 484  
anaphoric/causal-historical approaches  
to proper names, 308–309, 458, 459, 470–471, 572–574, 579–583, 585–586  
asserting/inferred, 158, 167–168, 171, 339, 347–348, 350, 351, 353, 355, 358, 359, 367  
asymmetric/symmetric, 372, 376–381, 384–385, 388, 391–395, 403, 455, 457, 490, 499, 564, 619–622  
attitude/status, 33, 37, 161–162, 165–166, 194, 197, 261, 290, 334, 357, 497, 593, 595–596, 598–601, 623, 648  
attributing/undertaking, 61, 62, 161–166, 196, 506–507, 525, 554, 596–601, 608  
authority/responsibility, xii, 161–165, 171, 174, 179, 229, 238, 242, 532

- distinction (*continued*)  
 Cartesian certainty/Kantian necessity, 9–11, 30, 636  
 causal-functional/normative models of intentionality, 9, 15, 22, 30, 56–57, 60, 234, 270  
 causal/normative, 12–15, 27–30, 33, 45–50, 93, 160, 196, 234, 617, 621, 625, 626  
 circumstances/consequences of application, xiii, 18, 28, 40, 51, 89, 98, 117–131, 136, 159, 162, 182, 243, 331–332, 372, 383, 419, 421, 432, 433, 482, 541, 600, 618, 631, 640  
 claimed/claiming, 327–330, 333, 488, 595, 606, 622–625  
 cognitive/practical, xv, 6, 9–10, 119–120, 212, 221, 225, 234, 293, 295, 432, 458, 473, 616–623  
 commitment/entitlement, xiv, 55, 159–161, 165–168, 179, 237, 238, 245, 252, 259–260, 541, 606, 649  
 complex/simple predicates, 371, 406–409, 434, 436  
 conferral/stipulation, 115–117, 145–147  
 conferring/instituting, 46–50, 52–55, 107–108, 115–117, 137, 145–147, 165–167, 626–628, 638, 646  
 correct/taken-correct, 29, 32, 52–55, 63, 595, 597, 647  
*de dicto/de re* ascriptions, 499–508, 511–515, 529–534, 539–548  
 deference/inference, 175–177, 192–193, 196, 204, 212, 218, 234, 241–242, 453, 531–536  
 derivative/original intentionality, 60, 143, 171, 629–644  
 designatedness values/multivalues, 340–350, 356–358  
 distinction/dualism, 614–624, 626  
 doing/saying, 30, 62, 77, 108–110, 115, 135, 639–641  
 doxastic/practical commitments, 171, 233, 236, 238–243, 271  
 entries/exits, language, 142, 221, 233–235, 258, 271, 335–336, 528, 632  
 epistemically strong/weak *de re* ascriptions, 503, 513, 529–530, 547–552, 566, 569–570, 581  
 explicit/implicit, xviii, xx, 18–26, 77–78, 85–86, 107–110, 121–122, 126–128, 147, 149, 247–249, 262–264, 432–433, 451–452, 498–499, 503–506, 512–513, 518–519, 531–532, 601–607, 612–613, 641, 649–650  
 facts/norms, 58, 137, 331, 623–626  
 force/content, 111, 186–190, 298–300, 322, 327–328  
 formal/material inferences, 97–102, 104–105, 117, 133–136, 345, 383, 616, 619, 622  
 formal/philosophical semantics, 143–145, 199  
 freestanding/ingredient content, 122, 338–340, 342, 345, 348, 353–359  
 inferentialism/representationalism, 6, 31–32, 92–94, 135–136, 205, 283, 285, 334–338, 360  
 intentions in action/prior intentions, 256–259, 558  
 intracontent/intercontent, intrapersonal/interpersonal, 169–170, 175–176, 179, 226, 241–242  
*I-thou/I-we* sociality, 39, 62, 508, 522, 526, 590, 593, 598–607  
 knowing-how/knowing-that, 23, 25–26, 101, 110, 135–136, 591, 641  
 linguistic/pragmatic approaches to intentionality, 16, 22, 76, 148–150, 152, 229, 230, 631  
 'looks/'is' talk, 292–297, 681n5  
 making-true/taking-true, 5, 8, 13, 46, 233, 236, 277, 287–291, 297–299, 521  
 natural/normative, xiii, xv, 12, 31, 35, 63, 149, 208, 289, 299–300, 624  
 objective/subjective, 52–55, 197–198, 212, 526, 592–602, 604, 609  
 practices/rules, 20–26, 32, 45, 55, 62, 64–66, 91, 99–100, 110, 509, 625  
 purported/successful representation, 6–7, 70–75, 89–90, 360  
 regularism/regulism, 18–29, 32, 36–42, 46, 62–63, 110, 208, 594  
 reported/reporting tokenings in ascriptions, 535, 537, 541, 566  
 sapience/sentience, 4–8, 87, 88, 231, 275–277, 520, 559, 591, 644  
 'shall/'should', 245, 258–259, 261, 263–264, 267–271, 553  
 true/taking-true, 287–292, 296–299, 322  
 weak/strong/hyper-inferentialism, 131–132

- doing, 4, 80, 87, 91. *See also* explicit; knowing-how/knowing-that; pragmatics  
and saying, 30, 62, 77, 108–110, 115,  
135, 639–641, 658n40
- donkey sentences, 490–493
- Donnellan, Keith, 488
- doxastic commitments, 142, 157–159, 167,  
178–180, 200, 228, 238, 266, 276, 344,  
346–348, 351, 472, 520, 679n40,  
687n11. *See also* ascription(s); assertion(s); deontic scorekeeping; entitlement  
and assertion, 142, 194, 230  
and belief, 157, 196, 201, 228  
and practical commitments, 171, 233,  
236, 238–243, 271
- Dretske, Fred I., 428, 430
- 'dthat', 469, 573
- dualism, 614–624, 626
- Dummett, Michael, 16, 17, 98–99, 116–  
118, 127, 144, 159, 187, 232–233, 296,  
329, 347–348, 350, 352, 359  
on assertion and belief, 153–155, 200  
on circumstances and consequences of  
application, 117–131, 136, 162, 182  
on complex/simple predicates, 406,  
688n27, 689n30  
on content, 339–340, 343, 345, 349, 358  
on inference/truth, 96–97, 111
- Edelberg, Walter, 691n48
- embedding, 298–300, 322, 338, 359, 381,  
605, 609, 680n60. *See also* content(s);  
force; logic, multivalued  
test, 298–300, 604
- empiricism, 10, 85–86, 89–90, 614. *See*  
*also* representationalism
- "Empiricism and the Philosophy of  
Mind," 89–90, 215, 293, 465
- endorsement, 208, 293–294. *See also* asser-  
tion(s); commitment(s); undertaking
- Enlightenment, 6, 10, 47–49, 92, 93,  
660n53
- entailment. *See* incompatibility; infer-  
ence(s); commitment-preserving; infer-  
ence(s); entitlement-preserving
- enthymemes, 98, 101, 206, 218, 635. *See*  
*also* material inferences  
in practical reasoning, 246–248, 252
- entitlement, 159–161, 176–179, 206, 208,  
212–215, 221, 226–227, 238, 239, 334,  
681n2. *See also* commitment(s); deontic  
attitudes; deontic scorekeeping; in-  
compatibility; inheritance of  
entitlements  
and commitment, xiv, 55, 142, 159–161,  
238, 245, 606, 675n44
- default-and-challenge structure of, 176–  
179, 184, 204, 221, 226, 238–239, 242
- and justification, 174, 204–206, 218, 532–  
533
- and reliability, 167, 206–211
- entitlement-preserving inferences. *See in-*  
*ference(s)*; entitlement-preserving  
entries and exits. *See language(s),*  
*entry/exit transitions*
- epistemology, 6, 93, 201, 203, 207–208, 216
- equivalence classes, 281, 342, 345, 348,  
351, 375, 392, 394, 400, 405, 408, 420–  
422, 450, 453, 454, 486, 619
- error, 31–32, 240, 295, 603, 606. *See also*  
ignorance
- evaluative judgments, 267
- Evans, Gareth, 460, 491, 567, 582, 696n62,  
708n72
- existential commitments, 71, 304, 415,  
416, 436, 440–449, 569, 694n17,  
697n67, 706n41  
and canonical designators, 320, 442–444,  
448
- negative existential claims, 319–320,  
448
- and substitutional commitments, 431,  
434, 440, 441, 445, 447
- exit transitions. *See language(s), entry/exit*  
*transitions*
- experts, 39–41, 600, 660n52, 709n75
- explanatory strategy, 45, 55–56, 84, 136,  
154, 199–201, 403, 716n37
- assertion and judgment, 151, 199, 200,  
202, 221, 232
- believing and claiming, 153–156
- bottom-up, 337–338, 340–341, 357, 358,  
364
- inferentialism/representationalism, xvi,  
6, 69, 93, 94, 97, 135–136, 149, 334–  
338, 495–496, 500, 503, 519, 584,  
667n70, 669nn90, 92
- intention and convention, 232,  
696n46
- linguistic practice and rational agency,  
155, 232

- explanatory strategy (*continued*)  
 and material inference, 132–133  
 naturalistic and normative, 149  
 normative phenomenalist, 25, 625, 628,  
 636–637
- pragmatist/regulist, xiii, 25–26, 101, 112–  
 113, 135, 149, 205, 657n31
- semantic, 149, 199, 362, 495–496
- sentential/subsentential expressions, 82–  
 83
- top-down, 337–338, 354, 358, 364
- explicating locutions. *See* logical vocabulary
- explicit (making explicit), xviii, xx, 18–26,  
 32, 46, 62, 77, 85–86, 101, 112, 126–  
 128, 140, 261, 279, 330, 395, 415, 418,  
 432–433, 442, 498, 505, 520, 530, 563,  
 586, 592, 599, 604–605, 624, 629, 639,  
 640, 644, 649–650. *See also* assertion(s); expressive rationality
- ascriptions, 228, 282, 498–499, 503–506,  
 530, 543, 588, 593, 612–613, 616, 629,  
 640, 679n42, 680n61
- deferrals, 226, 531–532, 534, 704n23
- de re* ascriptions, 280, 391, 508, 512–  
 513, 515, 522, 584, 586, 595
- inferential commitments, xix, 104, 106,  
 108–110, 231, 247–248, 396, 667n58
- logical vocabulary, 107–108, 116, 319,  
 374, 382, 383, 399, 402, 418, 419, 498–  
 499, 601–607, 619, 639
- norms, 18–23, 130, 247–249, 270, 271,  
 625, 639, 714n21
- ourselves, 275, 587, 641, 642, 650 (*see also* we)
- practical commitments, 247–249, 259,  
 262–264, 266, 267, 269, 403, 641
- pragmatics, 116, 121–122, 498–499,  
 650
- propositional contents, 77–78, 113, 135,  
 228, 401, 485, 586, 649
- reliability inference, 218–219, 221
- representation, 138, 280, 431, 608,  
 665n31
- semantics, xx, 116, 121–122, 137–138,  
 650
- substitutional commitments, 115, 319,  
 417, 467, 512–513
- token recurrence, 451–452
- exportation, 502, 516, 598, 610–611
- expressions, 75–77, 105–106, 384, 392,  
 393, 399, 403, 591, 650. *See also* sen-  
 tence(s); singular terms; subsentential  
 expressions; tokenings
- expressive commitments, 516, 545, 586,  
 588, 589, 608. *See also* *de re* ascrip-  
 tions; explicit
- expressive completeness (equilibrium),  
 111–116, 138, 613, 641–643, 650
- expressive deduction, xxiii, 401, 403. *See*  
*also* singular terms
- expressive development, 642
- expressive rationality, 105–111, 116, 125,  
 130–132, 642
- expressive role, 228, 245–246, 310, 330,  
 396, 414, 458, 473, 474, 486, 498–499,  
 541, 590. *See also* anaphora, expres-  
 sive role of explicit
- of ascriptions, 502, 504, 505, 522, 529–  
 530, 533, 591, 613
- of conditionals, 108–110, 126–127, 474
- of 'I', 554–559, 566
- of identity locutions, 115, 382, 476
- of language, 342, 352, 377
- of logical vocabulary, 95–96, 107–108,  
 110, 113, 114, 125, 359, 382, 383, 385,  
 530, 619, 629, 635, 641, 644, 650
- of normative vocabulary, 245–252, 261,  
 267, 271, 625
- of 'of', 138, 704n20
- of representational locutions, 138, 284–  
 285, 330, 499, 502, 505, 522, 529–530,  
 584
- of semantic vocabulary, 285, 311, 325, 414
- of token recurrence, 310, 453, 454, 458,  
 473, 474, 486, 590
- of 'true', 278, 284–285, 324, 326–333, 568
- expressivism, 92–93, 682n14
- extensionality, 344, 350, 352, 359, 392,  
 484–485, 668n72, 690n37. *See also*  
 logic, multivalued; substitution
- extensions, 109, 484–485, 681n9, 694n31.  
*See also* intensions
- externalism. *See* semantic externalism;  
 justification
- facades. *See* barn facades
- facts, xxi, 76, 84, 245, 324, 328, 331, 333,  
 622, 631–632. *See also* objectivity; ob-  
 jects; representation
- and norms, 137, 331, 623–626, 625
- semantic, 76, 326–329, 331, 333
- as true claims, 327–330, 333, 488, 595,  
 606, 622, 624, 625, 704n19

- failure, 258–259, 295  
 Feyerabend, Paul K., 480–481  
 Field, Harry H., 154, 156, 481  
 force, 5, 12, 17, 56, 82, 111, 186–190, 288, 297–299, 322, 339, 343–344, 367, 604, 661n65, 680n60. *See also* assertion(s); content(s); pragmatics; pragmatic significance; speech acts  
 formal inferences, 97–102, 104–105, 107–108, 133–136, 340, 351, 383, 619, 635. *See also* conditionals; material inferences; validity  
 formalism, 97–102, 110, 112, 135, 635  
 formal semantics. *See* semantics, formal/philosophical  
 foundationalism, 90, 177, 204, 216, 221, 681n5. *See also* ‘Empiricism and the Philosophy of Mind’; ‘looks’ talk; non-inferential reports  
 frames. *See* substitution frames  
 freedom, 32, 50–51, 662n77, 714n18  
 Frege, Gottlob, 11–13, 23, 80–82, 97, 112, 139, 200, 279, 354, 363, 433, 439, 443, 449, 475–477, 688n29, 693n2. *See also* tactile Fregeanism  
 on concepts, 72, 84, 282, 316, 352, 354, 355, 617, 693n10, 694n34  
 on content, 14, 72, 94–97, 107–108, 123, 281, 288, 298, 340, 344, 345, 347–348, 356, 475, 476  
 and expressive role of logic, 97, 107–111, 113  
 and inferentialism, 11–12, 80, 82, 94–97, 117, 281, 351  
 on objects, 279, 355, 360, 365, 366, 413–415, 419, 421–424, 435, 444, 448  
 on recognition judgments, 417, 418, 442  
 on substitution, 81, 105, 138, 281–282, 346, 367, 369, 414, 422–424, 436–438  
 fruitfulness, 476–477  
 functionalism, 16, 147–149, 159, 160, 196. *See also* incompatibility  
 Gauthier, David, 49  
 ‘gavagai’, 409–412, 429, 430  
 Geach, Peter T., 298, 492–493, 696n51  
 Geach-Frege test. *See* embedding, test  
 Gentzen, Gerhard, 125  
 gerrymandering, 28–29, 36, 41, 62, 208–212, 214, 645, 647, 710n92. *See also* regularism; reliability, reliabilism  
 Gettier, Edmund, 675n1  
 Gibbard, Allen, 682n14  
 giving and asking for reasons, xiv, 20, 46, 54, 136, 200, 209, 212, 230, 232, 275, 278, 330, 403, 449, 453, 496, 498, 590, 628, 637, 648, 715n28. *See also* communication; inferentialism; linguistic practice; reasons  
 and action, 156, 158, 194, 215, 233, 243, 244, 248, 263, 630  
 and assertion, 79, 89, 117, 139, 141, 158, 159, 167, 173, 205, 221, 229, 233, 449, 520, 593, 601–607, 624, 629, 641  
 Goldman, Alvin I., 209–211  
 ‘good’, 289, 298  
 grammar, 304, 361, 404, 406, 688n21  
 grasping, 9, 120, 355, 583, 635, 636. *See also* concepts; tactile Fregeanism; understanding  
 Grice, H. Paul, 146  
 Grotius, Hugo, 18–19  
 Grover, Dorothy, 300–301, 321–322  
 harmony, 124–130  
 Haugeland, John, 34, 36, 37  
 ‘he’, 312  
 Hegel, G. W. F., 50, 85–86, 92–93, 663n1, 669n93, 677n26, 698n78, 716n35  
 Heidegger, Martin, 661n64, 666n35  
 Hempel, Carl G., 679n44  
 ‘here’, 463  
 hierarchy, 36, 51, 160, 216, 242, 362, 404, 408  
 assertional/inferential, 350–353, 355, 356, 359  
 substitutional, 351, 353, 358  
 Hintikka, J., 681n7, 699n85  
 Hobbes, Thomas, 49, 51  
 holism, 89, 92, 426, 477–481, 587  
 Hume, David, 10–11  
 ‘I’, 439, 537, 552–566, 704n23, 707n51, 711n103  
 idealism, linguistic, 331  
 identity, xix, 314, 319, 372, 383, 398, 416, 439, 498, 530, 571, 573, 589, 695n34. *See also* singular terms; substitution claims, 112, 315–317, 324, 418–422, 424, 432, 441, 443, 444, 468, 476, 477, 489 locutions, 115, 372, 374, 382, 416–419, 451  
 ignorance, 240, 602, 605. *See also* error imperatives. *See* commands

- implicit. *See explicit; norms; practice(s)*
- incommensurability, 480–481, 483
- incompatibility, 189–190, 225, 332, 429, 560, 602, 634–635, 672n19, 690n35, 691n40, 711n102. *See also deontic scorekeeping; entitlement; functionalism*
- asymmetric/symmetric, 673n23, 674n39, 711n98
- and entailment, 160, 382, 602–603
- and inference, 12, 89, 115, 132, 169, 178, 190–191, 196, 201, 674n43, 710n91, 714n12
- material, 92, 160, 169, 384
- and negation, xix, 115, 436
- and practical commitments, 237, 253, 259, 269–271
- incorrigibility, 292–295. *See also 'looks' talk; trying*
- indefinite descriptions. *See definite descriptions*
- indeterminacy of translation, 409–412
- indexicals, 132, 282, 303, 309, 462, 535, 550–553, 558, 560–561, 565, 566, 577–578, 617, 633. *See also deixis; demonstratives; quasi-indexicals; singular terms; tokenings*
- and anaphora, 309, 460, 473, 585
- indirect descriptions. *See anaphorically indirect definite descriptions*
- indirect discourse ascriptions. *See ascriptions, indirect discourse*
- induction, 168, 189–190. *See also inference(s); entitlement-preserving*
- inference(s), 5, 12, 87, 90–91, 97–102, 104–108, 131–134, 139, 189–190, 206, 214–221, 225, 228, 260, 340, 347–348, 351, 372, 377, 383, 385–386, 392, 400, 402, 472, 619, 627, 634–635, 673n28, 689n32. *See also concepts; content(s); deontic scorekeeping; formal inferences; giving and asking for reasons; inferentialism; linguistic practice; material inferences; representation*
- and assertion, 91, 95, 158, 167–168, 190, 194, 218, 266, 367, 687n11
- commitment-preserving, 168, 189–190, 200, 237, 238, 344, 541, 553, 673n27, 679n40, 689n31, 710n91, 714n12
- committive (*see inference(s); commitment-preserving*)
- and conditionals, 91, 98, 100, 101, 106, 666n48
- and content, 89–92, 95–96, 102–103, 112, 144, 190, 354, 601–607, 618–623, 634–635
- entitlement-preserving, 168, 200, 220, 237, 238, 541, 673n27, 675n44, 678n38, 679n44, 681n2, 710n91, 714n12 (*see also induction*)
- four kinds of, 189–190
- and incompatibility, 89, 132, 189–190, 710n91, 714n12
- inferential strengthening/weakening, 379–382
- permissive (*see inference(s); entitlement-preserving*)
- reliability, 189–190, 215–218, 221, 228
- and representation, xvi, 93–94, 136, 391, 665n31
- social dimension of, 54, 91–93, 138, 158, 197, 358, 518, 519, 593, 601, 605
- substitution, 370–374, 410, 430, 621–622, 689n31, 690n33
- and truth, xvii, 5–6, 96–97, 104–105, 107–108, 277, 689n31
- inference, substitution, and anaphora (ISA), xvi, 198, 281–283, 391, 449–450, 457, 467, 472–473, 495, 621–623, 649, 650. *See also content(s); deontic scorekeeping; inferentialism; semantics*
- inferential articulation, 168–172, 275, 414, 430, 431, 466
- of discursive practice, 79, 91, 132, 142, 157, 186–190, 198, 225, 233–234, 237, 277, 284, 430, 449
- social dimension of, 167, 477–478, 586
- inferential commitments, 116–117, 248, 347, 351, 357, 451, 454, 506, 586–587, 640, 650, 687n11. *See also conditionals*
- inferential involvements, one-way, 371–372, 377, 386, 388, 392. *See also singular terms; substitution*
- inferentialism, xxi, xxii, 93–94, 104, 107–110, 117–132, 135, 137, 200, 214–221, 281, 334–338, 413, 429, 475, 495, 608, 620, 669n90, 686n1, 690n37, 690n37. *See also concepts; Dummett, Michael; Frege, Gottlob; representationalism; Sellars, Wilfrid; semantics*
- and objectivity, 109, 134, 137, 354, 478, 622–623, 633–636
- and representation, xxi, xxii, 92–94, 132, 205, 283, 285, 334–338
- inferentially inverting contexts. *See contexts, inferentially inverting*

- inferential polarity, 381, 387, 398, 399. *See also* contexts, inferentially inverting inferential roles, 89–90, 96, 105–106, 114, 281, 349, 413, 429, 618, 620, 636, 667n58. *See also* content(s); grasping; inference(s); inferentialism inferential significance, 475, 478, 480–481, 483, 633, 635 information, 474, 510, 514, 517, 546. *See also* communication; *de re* ascriptions 'ing/'ed', 330 inheritance, 168, 169, 283, 306, 460, 472. *See also* anaphora, inheritance of substitutional commitments, 132, 283, 454–455, 472–473, 499, 564, 581, 583, 621–622 inheritance of entitlements, 168–171, 175–176, 179, 193, 204, 212, 217, 218, 239, 242, 249, 276. *See also* authority; deference; entitlement; default-and-challenge structure of; justification; testimony interpersonal, intracontent/intrapersonal, intercontent, 169–170, 175–176, 179, 205, 218, 221, 226, 241–242 instituting, 165–167, 202, 284. *See also* practice(s); pragmatism of norms, xiii, 46–50, 52–55, 137, 140, 498, 626–628, 646 of statuses by attitudes, 61, 64, 115, 133–134, 142, 161–162, 169, 593, 597, 623–624, 630, 638 intellectualism, 20–22, 32, 77, 110, 135, 231, 669n92. *See also* Kant, Immanuel, platonism; regulism; rules intensions, 482–485. *See also* extensions intention(s), 8, 13–14, 58, 146–147, 193, 232, 233, 239, 253–259, 261–267, 523–526, 670n3, 681n1, 696n46, 702n14. *See also* action(s); reliable differential responsive dispositions in action, 256–259, 670n6, 678n34, 707n49 prior, 256–259, 558, 670n6, 678n34, 707n49 pure, 256–257 and reasons, 255, 261, 678n37 intentional explanation, 56, 57, 268–269, 364, 521, 524, 711n95. *See also* action(s); interpretation and *de dicto/de re* ascriptions, 522–526 as normative, 15–18, 195, 268–269 intentional interpretation. *See* interpretation intentionality, xv, 6–7, 15, 59, 61, 70, 99, 148–149, 155, 415, 416, 631. *See also* belief(s); deontic scorekeeping; interpretation; propositional contents; stance(s) analogical/relational linguistic approaches to, 16, 150–152 derivative/original, 60, 61, 143, 171, 629–644, 671n8, 715n25 descriptive/normative conceptions of, 9, 15, 22, 56–57, 60, 671n8 discursive, 7, 8, 24, 61–62, 67–70, 142, 631, 649 linguistic/pragmatic theories of, 16, 22, 76, 148–150, 152, 230, 631 and representation, 67–70, 336, 547–548 simple, 59, 171, 631 intentional stance, 55–62, 67, 629–631, 636–639, 642, 715n27 simple, 629–630, 639, 642, 643 intentional states, 16–18, 75, 118, 133, 147, 157, 196, 270. *See also* normative attitudes; normative status(es) normative significance of, 8, 13–16, 23, 27–29, 46, 55–57, 62, 67 intentional systems, 59, 60, 61, 629–630, 642 interlocutors, 559–560. *See also* deontic scorekeeping internalism, 215, 219, 221. *See also* reliability, reliabilism; semantic externalism interpretation, 66, 74, 139, 152, 510–513, 628–632, 637, 638, 642, 644–648, 699n86. *See also* deontic scorekeeping; *de re* ascriptions and communication, 513, 588, 645, 646, 670n6 and demonstratives, 510–513 and deontic scorekeeping, 136, 475, 508, 636, 644–645 external and internal, 645–648, 715n27, 716n37 intentional, 55–62, 83, 84, 158, 232, 715n27 and original intentionality, 632, 640 and personal pronouns, 510–513 as substitution, 20–21, 65, 353, 508–513, 591 and understanding, 508, 517, 658n39, 660n56 and Wittgenstein, 20–22, 61, 62, 509–513 interpretive equilibrium, 641–644

- ISA. *See* inference, substitution, and anaphora
- iteration, 295, 313–316, 319, 701n6, 702n13. *See also* anaphora
- I-thou/I-we* sociality, 39, 62, 508, 522, 526, 590, 593, 598–607, 659n50, 716n36. *See also* deontic scorekeeping; norms; objectivity; perspectives; we
- James, William, 287–288
- JTB account of knowledge. *See* knowledge, JTB account of
- judgment, 7, 12, 84–86, 95, 200, 363, 614. *See also* assertion(s); Frege, Gottlob; Kant, Immanuel
- justification, 11–14, 22, 90, 167, 201–202, 204, 217, 221, 228, 294, 515, 532–533. *See also* assertion(s); inference(s); knowledge and entitlement, 174, 204–206, 234 externalism and internalism about, 219, 221 and reliability, 207–208
- justified-true-belief account of knowledge. *See* knowledge, JTB account of
- Kant, Immanuel, 9–11, 14, 47, 58, 80, 92–94, 102, 200, 230, 337, 475–477, 614, 617, 625, 655n1, 662n93, 665n29, 669n93, 712n10, 713n10 on acting according to conceptions of rules, 30–33, 35, 41–42, 45, 50, 65 and classificatory conception of concepts, 85–86, 614–617 on intentionality as normative, 7–11, 23, 29–30, 289 on norms as rules, 8, 18–19, 27, 30, 32, 52, 200, 206, 623–624, 655n1, 656n19, 657n31, 712n5 on primacy of judgment, 79–80, 95, 136, 362–363, 614, 619, 643, 658n41 three dualisms of, 18, 614–618, 622, 661n65 on will, 32, 50–52, 270, 271, 665n33
- Kaplan, David, 469, 547–548, 550, 561–562, 566, 617
- knowing-how/knowing-that, 23, 25–26, 135–136, 591, 641, 658n40. *See also* explicit; pragmatism
- knowledge, 74, 177, 200, 202, 204, 209, 210, 221, 228, 297, 645, 676nn3, 7, 715n27. *See also* assertion(s); observational knowledge and assertion, 199–204 and doubt, 177, 209 as hybrid deontic status, 201–205, 213–215, 220, 297
- JTB account of, 201, 207, 228, 297, 515, 675n1 and reliability, 209–211, 219–220 and truth, 202, 204, 221, 297 and understanding, 90, 209, 213–215
- Kremer, Michael, 668n82, 673n20
- Kremer, Philip, 674n42
- Kripke, Saul A., 28, 37, 209, 322, 468–471, 483, 488, 575–579, 582, 603, 656n10, 660n52, 713n11
- Kulas, J., 699n85
- Kvart, Igal, 691n46
- Lakatos, Imre, 697n66
- Lance, Mark, 668n73, 673n23, 674n42, 690n35
- language(s), 24, 146–147, 232, 342, 352, 365, 377, 403, 411, 500, 650. *See also* assertion(s); deontic scorekeeping; explicit; giving and asking for reasons; inference(s); linguistic practice; norms; we and *de re* ascriptions, 499–503 entry/exit transitions, 335–336, 528, 632. *See also* action(s), as language exit transition; perception, as language entry transition games, 91, 172, 179 and mind, xv, xxiii natural, 145, 411, 499, 504, 520, 688n21 use (see pragmatics)
- Leibniz, Gottfried Wilhelm, 10, 93, 337, 614
- LePore, Ernest, 690n32
- Lewis, C. I., 616
- Lewis, David K., 147, 180–183, 185, 187, 232, 404, 460, 483, 550, 670n3
- licenses, 161, 163–165, 196
- Lindenbaum algebras, 342, 345
- linguistic practice, xi, 16, 141, 155–156, 167–168, 172, 232, 275–277, 331–332, 360, 496, 586, 623, 628, 630–632. *See also* assertion(s); deontic scorekeeping; explicit; giving and asking for reasons; inference(s); norms

- and assertion, 141, 167–168, 172, 173, 276, 367, 586, 623, 628–629, 686n1 inferential articulation of, 141–143, 156–158, 167–168, 183, 279–281, 431, 449, 608, 614, 637, 649
- Locke, John, 146, 614
- logic, 12, 96–100, 108–111, 110, 135, 231, 340, 347–348, 353, 384, 435. *See also* explicit; logical vocabulary
- expressive approach to, xix, 108–111, 117, 125, 131, 650
- multivalued, 340–346, 358, 359
- logical vocabulary, xviii, xx, 76, 101, 103–105, 112, 114, 116, 123–125, 127, 135, 266, 346, 381–384, 394, 414, 498, 530, 559, 566–567, 601–607, 639–643, 667n67, 691n42. *See also* ascription(s); conditionals; deontic scorekeeping; explicit; identity; negation; normative vocabulary; quantification
- expressive role of, 13, 95–96, 109, 113, 116–117, 131, 335, 350, 359, 382–383, 385, 392, 402–403, 530, 619, 629, 635, 641, 643–644, 650
- and singular terms, 393, 395–397, 401
- 'looks' talk, 292–297, 681n5. *See also* foundationalism
- Lycan, William, 691n47
- making-true, 5, 8, 13, 46, 233, 236, 277, 521, 681n1. *See also* action(s); phenomenism, normative; practical commitments; taking-true
- Manders, Kenneth, 691n41
- maps, 65, 74, 518, 665n31. *See also* representation
- material inferences, 97–102, 104–105, 125–137, 189, 206, 218, 345, 349, 359, 373, 374, 383, 402, 619, 666n48, 690n36.
- See also* conditionals; content(s); enthymemes; explicit; formalism; incompatibility; inference(s); inferentialism; Sellars, Wilfrid
- and conceptual contents, 98, 102–108, 618, 623, 634–635
- and formal inferences, 97–102, 104–105, 117, 135–136, 345, 383, 666n41, 672n19
- McDowell, John H., 29, 54, 318, 535–539, 561–562, 566, 567, 577, 582, 603, 655n2, 658n39, 660n59, 686n58, 689n29, 693n10, 705n26, 707n55, 708n72
- McGinn, Colin, 703n17
- meaning, xii–xiii, 4, 13–14, 60, 62, 73, 88–89, 121–124, 126, 146, 478, 649, 716n32. *See also* content(s); inference, substitution, and anaphora; inferentialism; norms; pragmatics; semantics; understanding
- Meinong, A. Alexius, 71
- methodology, 229–233, 526, 592. *See also* explanatory strategy
- explicit/implicit, xiv, 109, 587
- phenomenalism, 597, 636–637
- substitutional, 81, 95–96, 104–105, 138, 281
- mind, xv, xxiii, 650. *See also* explicit; inference(s); intentionality; language(s); rationality; sapience; sentience; we
- mistakes, 21, 27–28, 31, 52, 54, 258
- modality, 13, 105–106, 318, 690n36, 703n19
- model theory, 667n58
- Moore, G. E., 299
- multivalued logic. *See* logic, multivalued multivalues, 340–346, 348–353, 358. *See also* designatedness; logic, multivalued; substitution
- 'must', 12, 14, 30
- name-bearer model, 69, 84, 352, 359. *See also* designation; nominalism; proper names
- names. *See* proper names
- naturalism, xiii, xv, 10–15, 31, 42–46, 289, 299–300, 460–461. *See also* causes; dispositions; norms; regularism; regularities; reliability
- natural kind terms. *See* sortals
- Neale, Stephen, 491–494
- necessity, 9–11, 30, 624. *See also* Kant, Immanuel
- negation, xix, 92, 115, 319, 381–382, 393, 436, 498, 530, 692n54. *See also* explicit; incompatibility; logical vocabulary
- expressive role of, 401, 498
- negative existential claims. *See* existential commitments, negative existential claims
- nominalism, 84, 665n32. *See also* designation
- noninferential reports, 64–66, 88–91, 207, 209–210, 212, 214–221, 225, 236, 293, 335–336, 458, 473, 531–532, 620,

noninferential reports (*continued*)

- 674nn38, 40. *See also* content(s), empirical (cognitive); foundationalism; inferentialism; observational knowledge; perception; reliable differential responsive dispositions  
 authority of, 122, 212, 215–217, 221, 223–227, 234  
 as not autonomous, 216, 465–466  
 and circumstances and consequences of application, 119–120, 221, 226  
 enabling/defeating conditions, 226–227  
 Sellars on, 215–217  
 and solidity of linguistic practice, 631–632  
 normative attitudes, 32–37, 39, 42, 45, 47–50, 52, 54, 64, 626, 627, 639. *See also* assessment; deontic attitudes; normative status(es); norms; pragmatics, normative; sanctions  
 normative phenomenism. *See* phenomenism, normative  
 normative pragmatics. *See* pragmatics, normative  
 normative significance, 48–50, 52, 54, 167, 656n17. *See also* instituting; norms; pragmatic significance  
 of intentional states, 8, 13–16, 55–57, 62, 67  
 normative stances, all the way down, 638  
 normative status(es), 16–18, 33, 37, 39–41, 47, 64, 627–628, 676n4. *See also* deontic status(es); instituting; norms; practice(s); pragmatics, normative  
 normative vocabulary, xiii–xv, xviii, 47, 116, 233, 246–249, 267, 624, 625, 637, 640. *See also* explicit; norms; practical reasoning  
 expressive role of, 250–252, 267, 625  
 norms, xx, 7–18, 20–21, 30, 35, 37–42, 44–46, 49–52, 55–62, 200, 226, 636–639, 641, 648, 649, 656n18, 661n69. *See also* deontic scorekeeping; explicit; instituting; practice(s); pragmatics  
 all the way down, 44, 625, 638, 649, 660n59, 714n20  
 and causes, 27–29, 33, 45–50, 625, 626, 661n65, 687n2  
 conceptual, 12–14, 46, 53–55, 624, 636, 637

and facts, 623–626

implicit in practices, xiv, 20–22, 25, 26, 29–30, 41, 45, 46, 54, 55, 62, 623, 626, 627, 631, 648, 649

objective, xxiii, 52–55, 63, 253, 631  
 origin of, 626–628

pragmatism about, 21–23, 32

regularism/regulism about, 18–23, 26–29, 41, 46, 99, 648  
 as rules, 7–11, 18–23, 32, 200, 624, 656n19

novel terms, 420, 421, 423, 435, 442

'now', 463, 559

numbers/numerals, 437, 442–444, 449

'object' as pseudosortal, 438

object-dependence, 567–573. *See also* beliefs, strong *de re*; existential commitments; singular thoughts

object-directedness, 415, 416

objectivity, xvii, 78, 136–138, 253, 498, 529, 592–597, 602–607, 649, 672n18.

*See also* I-thou/I-we sociality; norms; perspectives; representation

of conceptual norms, xxiii, 53–55, 63, 497, 529, 593–594, 599–607, 631, 633, 636

and deontic attitudes/statuses, 599–601, 627

and *de re* ascriptions, 598, 600–601

as form of contents, 597, 600–601

and intersubjectivity, 599–607

perspectival character, 52–55, 197–198, 592–602, 604, 609

and representation, 140, 280, 530, 609, 672n18, 677n11

objectivity proofs, 601–607

objects, 292, 333, 360, 403, 438, 571, 649, 694n29, 713n10. *See also* expressive deduction

abstract, 84, 421, 422, 449, 695n34

and facts, 84, 331, 333, 622

picking out, 375, 413–417, 423, 425, 430–432, 444, 451, 462

recognizing as the same again, 416–419, 424, 425, 467

representation of, 136, 280, 333, 337, 617, 665n29

and singular terms, xxi, 69–70, 414–416, 687n14

and solidity of practice, 631–632

why are there?, xxii, 404

- observational knowledge, 209, 216–226, 297. *See also* noninferential reports
- observation reports. *See* noninferential reports
- occurrence, 342, 373–375, 389, 454, 465. *See also* recurrence; substitution
- primary, 374–375, 378–379, 381, 389, 392, 394, 395, 397, 400, 421
- 'of', 68, 138, 501–503, 505–506, 516, 520, 543, 546, 565, 569–570, 585–586, 590, 598, 608, 704n20. *See also de re ascriptions*
- ofness. *See de re ascriptions; 'of'*
- 'one', 438
- ontological commitments. *See* existential commitments
- opacity, 571, 574–575, 583
- 'ought', 5, 31, 56, 61, 252–254, 270, 271, 289, 290. *See also* intentional explanation; practical reasoning
- paradoxes, semantic, 321–322
- paratactic theory of ascriptions. *See* ascriptions, paratactic theory of
- parrots, 88, 122
- particularity, 620, 623, 687n14. *See also* deixis; inference; substitution, and anaphora; objects
- paycheck sentences, 490–492
- Peirce, C. S., 289
- perception, xv, 7–8, 119–120, 122, 131, 142, 209–211, 261, 276, 332, 556, 618, 704n22. *See also* content(s); empirical (cognitive); deontic scorekeeping; noninferential reports; observational knowledge; reliability; reliable differential responsive dispositions
- and action, 233–238
- as language entry transition, 221, 233–235
- performances. *See* action(s); speech acts
- Perloff, Michael, 702n14
- permissive inferences. *See* inference(s); entitlement-preserving
- Perry, John, 460, 550, 552–553, 558, 561, 566
- perspectives, 37, 62, 197, 503, 591, 594–595, 598, 608, 630, 635, 710n95, 715n27. *See also* deontic scorekeeping; *de re* ascriptions; *I-thou/I-we* sociality; objectivity; representation
- and ascriptions, 508, 549, 584, 590, 591, 613
- and attributing/undertaking, 61, 508, 597, 598, 649, 677n11
- and communication, 139, 635, 647
- and contents, 139–140, 485, 517, 529, 586–597, 601, 635, 636
- and knowledge, 202, 205, 297, 715n27
- privileged, 599–600, 604, 606
- and semantic externalism, 633, 647
- phenomenalism, 291–297, 327, 624, 631, 682n14, 709n75. *See also* deontic scorekeeping; 'looks' talk; norms; objectivity; stance(s)
- about deontic statuses, 334, 339, 637
- about norms, 25, 280, 627
- about truth, 287, 291, 292, 296–297, 322
- generic, 295–296
- normative, 627, 636, 637, 644
- and pragmatism, 296, 322
- subjective, 292–297
- philosophical semantics. *See* semantics, formal/philosophical
- physicalism, 47. *See also* causes; norms; regularism
- Plato, 201
- platonism, 20–22, 110, 231. *See also* intellectualism
- points of view. *See* perspectives
- practical commitments, 233, 237, 244, 245, 253–256, 262–266, 276, 679n40. *See also* action(s); content(s); practical; deontic scorekeeping; norms; practical reasoning
- causal efficacy of, 259–262, 271, 596
- and doxastic commitments, 233, 236, 238–243, 271
- entitlement to, 238, 253–256, 265–266
- inferential articulation of, 233–234, 237
- and intentions, 193, 256–259, 525, 702n14
- practical reasoning, 7, 83, 100, 116, 158, 233, 240, 244, 246–253, 268–269, 287, 290, 521, 537, 640. *See also* action(s); deontic scorekeeping; explicit; intentional explanation; material inferences; norms
- and action, 8, 231, 336, 507
- and doxastic and practical commitments, 237, 243, 246, 253, 254

- practical reasoning (*continued*)  
 and entitlement-preserving inferences, 249–250  
 and normative vocabulary, 233, 247–249, 625, 640  
 patterns of, 233, 245–252, 271  
 and ‘shall’/‘should’, 245, 258–259, 261, 263–264, 267–271, 553, 680nn56, 60  
 and theoretical reasoning, 520–522  
 and will, 233, 270
- practice(s), xi, 7, 22, 25, 41, 52–55, 61, 62, 65, 83, 91–93, 130, 159, 165–168, 205, 233, 339, 346, 627–628, 638, 660n51. *See also* conferring content; deontic scorekeeping; explicit; giving and asking for reasons; instituting; linguistic practice; norms; regularism; regulism and contents, 64, 77–79, 145–147, 497 idealized, 128, 168  
 norms implicit in, 25, 29–30, 55, 625, 628, 648, 649  
 and regularities, 26, 625, 638  
 and rules, 20–22, 32, 66, 91, 99–100, 110 solidity of, 332, 528, 631–632, 686n55, 715n28
- pragmatic priority of the propositional, 79–81, 83, 95, 337. *See also* assertion(s); Frege, Gottlob; Kant, Immanuel; Wittgenstein, Ludwig
- pragmatics, 133, 140, 159, 187, 212, 334, 474, 592, 624, 681n6. *See also* circumstances and consequences of application; deontic scorekeeping; force; pragmatic significance; speech acts normative, xiii, 132–134, 140, 199, 623, 649  
 and semantics, xiii–xv, xvii, xxii, 64, 68, 83–85, 91, 132–134, 140, 143, 363, 496, 498, 649, 686n1, 704n20
- pragmatic significance, 118, 142, 157, 163, 168, 173, 182–186, 262, 339, 343, 345, 359, 386, 392, 424, 450, 710n91. *See also* content(s); deontic scorekeeping; instituting; norms; practice(s) of assertions, 157, 168, 190, 200, 234, 358, 637
- pragmatism, xii, 21–23, 76, 101, 110, 120, 123, 132, 137, 143, 205, 289, 296, 300, 322, 496, 591–592, 686n1. *See also* norms; practice(s); regulism; Wittgenstein, Ludwig about norms, 21–23, 42, 55, 509
- classical, 123, 285–291, 296, 303, 322 and truth, 285–286, 288–291, 297
- predicates, xv, 320, 369, 378, 380, 391–395, 398, 400, 401, 406, 410, 411, 617, 620, 622, 623. *See also* singular terms; substitution
- complex/simple, 371, 406, 407, 409, 434, 436
- preferences, 240, 246, 248, 678n29. *See also* practical reasoning
- prima facie reasons. *See* reasons, prima facie
- Prior, Arthur N., 125
- prior intentions. *See* intention(s), prior privilege, 28–29, 36, 43, 63, 104–105, 117, 292–293, 634–635. *See also* perspectives, privileged
- pro-attitudes, 246–247, 267. *See also* normative vocabulary; practical reasoning
- proform-forming operators, 283, 313–315, 318, 325
- projection, 366, 395, 396, 399, 636, 647, 688n23
- promises, 163–165, 172, 180, 262–266, 289, 672n20. *See also* practical commitments
- pronoun-forming operators, 305, 306, 322
- pronouns, 132, 301, 305, 308–309, 311–331, 450, 456, 458, 460, 469, 486, 491, 510–513, 582, 707n51. *See also* anaphora; ‘refers’; singular terms
- proof theory, 667n58
- proper names, 308–309, 420, 439, 469, 470, 549, 573–579, 581–583. *See also* singular terms; substitution; token-recurrence structures
- anaphoric account of, 308–309, 458, 459, 470, 572–574, 579–583, 585–586
- causal-historical theories of, 470–471, 580, 713n11
- properties, 9, 12, 47, 52, 324, 697n67. *See also* proprieties
- propositional, pragmatic priority of. *See* pragmatic priority of the propositional
- propositional-attitude-ascribing locutions. *See* ascription(s), of propositional attitudes
- propositional attitudes, ascriptions of. *See* ascription(s), of propositional attitudes
- propositional contents, xiv–xv, 5, 8, 11, 75, 77–78, 83, 135, 141, 168, 203, 209, 221, 230, 275, 329, 333, 335, 339, 413,

- 495, 520, 567–568, 624, 640, 649, 665n32, 702n14, 705n32. *See also* assertion(s); content(s); explicit; giving and asking for reasons; linguistic practice; objectivity; objects; representation and assertion, xxiii, 12–14, 157, 187, 221, 278, 335, 586, 604  
 conferring of, 63, 275, 277, 279, 284, 623, 645  
*de re* specifications of, 513–517, 520  
 and inference, 6, 91–92, 101, 104, 200, 209, 277, 281, 337–339, 413, 449, 495  
 perspectival character of, 485, 497, 630, 636  
 representational dimension of, 75, 84, 153, 333, 414, 474, 485, 508, 519–520, 583, 649  
 and truth conditions, 6, 277, 278, 326, 329  
 proprieties, 9, 12, 25, 47, 52, 63, 83, 159, 411, 628, 636, 676n4, 677n11, 688n23.  
*See also* norms; practice(s); properties  
 prosentence-forming operators, 303–305, 327  
 prosentences, 283, 301–306  
 prosentential theory of truth. *See* truth, prosentential theory of  
 prosortals and pseudosortals, 438–439  
 psychologism, 11–12, 23  
 Pufendorf, Samuel, 18–19, 46–50, 160, 662n75  
 punishment, 34–36, 40, 42, 43, 63, 179, 662n75. *See also* sanctions  
 purport. *See* representation, purport/up-take; singular referential purport  
 Putnam, Hilary, 299, 481  
 “Puzzle about Belief,” 573–579  
 quantification, 300–301, 303–304, 307, 321, 323, 382, 435–437, 439, 440, 491–494, 530, 696n51, 711n100. *See also* explicit; logical vocabulary; predicates; substitution  
 expressive role of, 383, 396, 399, 434–436, 498, 696n50  
 quasi-indexicals, 559–567, 569–570, 581, 588, 590, 707n63, 711n103. *See also* ascription(s); communication; explicit  
 quasi-indicators, 563–564, 566  
 queries, 172, 192–193. *See also* deontic scorekeeping; speech acts  
 Quine, W. V. O., 92, 129, 223, 296, 342, 360–361, 409–412, 477–482, 547–548, 550, 587, 634, 704n24  
 quotation  
 <> type designation, 314  
 // token designation, 314  
 ↗ corner quotes, 688n25, 704–705n24  
 and direct/indirect discourse ascriptions, 529–534, 537  
 s<sup>s</sup> scare quotes, 545–547, 588–590, 672n19  
 ralugnis mrets, 390–391  
 rational agency. *See* action(s); practical reasoning  
 rationalism, 10, 85–87, 92–94, 102, 103, 614. *See also* inferentialism  
 rationality, 5, 6, 15, 17–18, 50–51, 56–57, 62, 99, 108–110, 127–128, 131, 135, 183, 195, 196, 229–233, 244–245, 252–253, 271, 364, 643. *See also* explicit; giving and asking for reasons; intentionality; interpretation; linguistic practice  
 expressive, 105–106, 116, 125, 130–132  
 instrumental, 106, 108–110, 131, 231  
 and will, 32, 50–51, 661n69, 665n33  
 reasons, xxiii, 8, 11–12, 116, 141, 171, 200, 203, 253, 275, 276, 383, 402, 449, 474, 590, 618, 649. *See also* giving and asking for reasons; inference(s); inferentialism; norms; practical reasoning; rationality; theoretical reasoning  
 and action, 56, 171, 230, 237, 240, 244, 245, 249–253, 261, 678n37  
 and causes, 234, 259–260, 271, 661n65  
 force of, 5, 17, 56, 354  
 prima facie, 249–253  
 primary, 240, 246, 259, 262, 663n84  
 recognition, 4, 31–32, 67, 640. *See also* we  
 recognition judgments, 112, 417–419, 424, 432, 444. *See also* Frege, Gottlob; identity, claims; singular terms; substitution  
 recurrence, 455, 457–459, 463, 465, 468.  
*See also* anaphora; occurrence; substitution; token-recurrence structures; tokenings  
 recurrence commitments, 456–457, 472  
 redundancy, 288–289, 291, 299–300, 303–304, 316, 322, 329. *See also* truth

- reference, 84, 305–327, 355, 356, 360, 467, 471, 477–482, 550, 703n16. *See also* explicit; objects; ‘refers’; representation; singular terms; substitution; truth  
 and anaphora, 304, 306, 307, 325  
 direct, 464, 550, 567–568, 571–574, 582  
 and inference, xvi, 136, 391  
 as a relation, 306, 323, 325–326  
 reference classes, 210–212. *See also* reliability  
 referential purport. *See* representation, and purport/uptake; singular referential purport  
 ‘refers’, xvii, 116, 137, 138, 279, 283–285, 305–306, 312–322, 325, 464. *See also* anaphora; reference  
 and anaphora, 305–306, 323  
 and representation, 336, 499  
 and ‘true’, 283–285, 305–307  
 reflection, Socratic, 105–106, 127–128, 130  
 reflexives, 563, 573, 698n80  
 regress, 22, 31–32, 36, 61, 74, 77, 177–178, 205, 206, 221, 451–452, 519, 600–601, 646, 647. *See also* pragmatism; regulism; Wittgenstein, Ludwig  
 of rules, 20–30, 36, 45–46, 61–66, 206, 509, 657n31  
 regularism, 26–29, 32, 36–42, 46, 62, 208, 594, 658n35. *See also* norms; regularities  
 regularities, 11, 28–29, 37–42, 624, 629, 659n47, 660n52, 710n92. *See also* regularism  
 and correctness, 27–28, 62, 207–208, 212  
 and norms, 26–29, 45–46, 99, 628, 648  
 and practices, 625, 638  
 privileged, 28–29, 36, 63  
 and reliability, 207–209, 211  
 regulism, 18–29, 32, 62–63, 110, 594. *See also* intellectualism; norms; pragmatism; rules  
 Reichenbach, Hans, 683n34  
 reinforcement, 34–35, 37, 42–43, 659n45. *See also* dispositions; sanctions  
 reliability, 207–214, 226, 331, 427, 532–533, 556, 646, 681n2. *See also* barn facades  
 attribution of, 217–221  
 and authority, 215–217, 219–220  
 and entitlements, 167, 207–208, 210–211  
 and inference, 189–190, 215–218, 221, 228  
 and regularity, 207–209, 211  
 reliabilism, 121, 206–209, 219–220, 228, 715n27  
 and success, 521, 527–528  
 reliable differential responsive dispositions, 5, 33, 35, 42, 89, 119, 122, 162, 210, 214–221, 223–225, 263, 269–271, 427, 430, 473, 518, 531, 555, 556, 618, 658n45. *See also* perception  
 and action, 234, 262, 269–270, 524–525, 702n15  
 and inferentialism, 86–88, 91, 119–120, 156, 429, 622  
 and noninferential reports, 7–8, 209, 223–224, 235, 261, 293, 465  
 repeatables, 282, 449–452, 454, 592, 623. *See also* token-recurrence structures  
 replacement, 369, 371, 394, 396, 406–409, 692n51. *See also* substitution frames  
 reports. *See* noninferential reports  
 representation, xvi, 7, 10–11, 17, 31–32, 69–72, 74, 77–79, 88–90, 116, 136, 330–333, 337, 391, 414, 495, 502, 518, 519, 609, 617, 649. *See also* aboutness; inferentialism; objectivity; propositional contents; representationalism; semantics  
 and content, 6, 54, 69, 70–72, 78, 79, 84, 333, 474, 497, 508, 519–520, 590, 593–595, 601–607, 609, 649  
 and deontic scorekeeping, xvii, 187, 324, 584  
 and *de re* ascriptions, 516, 547–548, 552, 586  
 and discursive practice, 279–282  
 expression of, 75–77, 335–336, 518–519, 608, 665n32, 709n90  
 and inference, 93–94, 518–519, 665n31  
 and intentionality, 67–70, 279  
 and objectivity, 54, 75, 78, 140, 151, 153, 280, 630, 677n11  
 and perspectives, 496, 497, 529  
 and purport/uptake, 6–7, 62, 70–75, 77, 78, 89–90, 138, 489  
 and success, 70–75  
 representational contents, 6, 69, 74, 75, 84, 135–140, 497, 517, 520, 528, 609. *See also* propositional contents

- representationalism, xxii, 6, 31–32, 84, 92–94, 97, 334–338, 669n90. *See also* empiricism; inferentialism
- representational locutions, 279, 284, 306, 496, 500, 519, 584. *See also* *de re* ascriptions; explicit; ‘refers’
- ‘represents’, 70–72, 517. *See also* representation
- responsibility, 7, 10–11, 17–18, 163, 170, 172, 173, 178, 179, 254, 516, 624, 712–713n10. *See also* assertion(s); justification; norms
- and authority, xii, 162–163, 165, 242, 673n24
- rewards, 34–36, 40, 42, 63. *See also* sanctions
- rigidity, 14, 468–472, 486, 573, 698n80. *See also* anaphora
- Rosenberg, Jay F., 147
- Rousseau, Jean-Jacques, 51, 662n70
- Royce, Josiah, 375
- rules, 8, 10–11, 20–22, 25, 31–32, 64–66, 91, 130, 616, 658n40. *See also* intellectualism; norms; pragmatism
- conceptions of, 9, 30–32, 65, 624
- and correctness, 18–19, 23, 24, 26, 64–66
- elimination/introduction, 101, 117–130
- and norms, 7–11, 15, 18–23, 62, 624, 656n19
- and practices, 20–22, 91, 99–100, 110
- regress of, 20–27, 29–30, 36, 45–46, 62–66
- and regulism, 18–23
- Russell, Bertrand, 307, 415, 436, 437, 567–568
- Ryle, Gilbert, 23
- Salmon, Nathan, 708n67
- samesaying, 536–537, 539, 562, 566
- sanctions, 34–37, 42–46, 63, 162, 164, 179, 659n45
- external/internal, 44–46, 162–165, 178–180, 660n54, 662n75
- and normative attitudes, 35–36, 42, 45, 166
- and reinforcement, 34–35, 42–43
- sapience, 4–8, 87, 88, 231, 275–277, 520, 559, 591, 644. *See also* linguistic practice; rationality; sentience; we
- saying, 18–19, 64–66, 107–108, 336, 535–538, 629, 647, 649. *See also* assertion(s); explicit; ‘says’
- and doing, 30, 62, 77, 108–110, 115, 135, 639–641, 658n40
- ‘we’, 3–5, 7, 643–645
- ‘says’, 529–539, 542. *See also* saying
- Scheffler, Israel, 481
- score. *See* deontic scorekeeping
- scorekeepers. *See* deontic scorekeeping
- Searle, John R., 60, 70, 147, 256–257
- seeming. *See* ‘looks’ talk
- self-consciousness. *See* consciousness/self-consciousness
- Sellars, Wilfrid, 16, 105–107, 130, 155, 209, 234, 258, 306, 660n51, 679n39, 706n39
- on concepts, 10, 89–91, 93, 94, 96, 117, 484, 618, 634
- on giving and asking for reasons, 139, 167, 263
- on ‘looks’ talk, 293–294
- on material inferences, 102–103, 105–106, 109, 618
- on noninferential reports, 122, 214–221, 228, 465–466
- on regress of rules, 23–26, 184
- selves, 554, 555, 559–560. *See also* ‘I’ as coresponsibility classes, 559–560
- semantic externalism, 219, 332, 631–633, 645–649. *See also* *de re* ascriptions
- semantic facts. *See* facts, semantic
- semantic primitives, 79, 133–134, 136, 283, 285, 681n6, 687n2
- semantics, 140, 144, 187, 360, 363, 370–372, 592, 687n2. *See also* content(s); inferentialism
- agent, 147, 149, 151, 671n8, 696n46
- formal/philosophical, 143–145, 199
- inferential, xiii, 93–94, 132–134, 334–338, 360, 623, 649
- and pragmatics, xiii–xv, xvii, xxii, 64, 68, 83–85, 91, 132–134, 363, 498, 592, 649, 686n1, 704n20
- semantic vocabulary, xvi, 82, 145, 279, 280, 283–285, 325. *See also* explicit
- traditional, 116, 285, 323, 325–327, 330, 432, 450
- semiotics, 80
- senses, 340, 343, 355, 356, 536–537, 570, 571, 574–575, 579, 582–583, 681n9. *See also* content(s); tactile Fregeanism
- sentence(s), 5, 69–70, 81, 82, 84, 118, 146, 173, 288, 329, 338, 347, 363, 364, 366, 377, 385, 391–395, 409–412, 426, 453,

- sentence(s) (*continued*)  
 533, 539, 669–670n1, 688n29. *See also* deontic scorekeeping; explicit; giving and asking for reasons; inference(s); linguistic practice; pragmatic priority of the propositional; propositional contents  
 and assertion, 82, 85, 141, 156, 157, 168, 199, 277, 367, 586  
 and beliefs, 146–147  
 compound, 199, 340–342, 345, 347–349, 351, 352, 355, 358, 395, 397, 398, 402, 403  
 frames, 377, 384, 386, 393, 394, 403, 405, 407  
 freestanding, 199, 298–300, 392, 399  
 nominalizations, 303–306, 516  
 novel, 362–367, 375, 414  
 sentience, 5, 7, 275–277. *See also* sapience  
 'shall', 258–259, 261, 267–271, 680nn56, 60, 716n34. *See also* practical reasoning  
 and practical commitments, 245, 261, 263–264, 269, 271, 553  
 sharing concepts, 562, 590, 631–636. *See also* communication  
 'should', 267–271, 716n34. *See also* practical reasoning  
 and practical commitments, 269, 553  
 significance. *See* normative significance; pragmatic significance  
 simple material substitution-inferential commitments (SMSICs), 372–376, 378, 382–385, 389–392, 397, 398, 400, 410, 424, 432, 434, 444, 450, 454, 455. *See also* identity; singular terms; substitution; substitutional commitments asymmetric/symmetric, 375, 377, 379–382, 386, 388, 393–396  
 and singular terms, 424, 432, 454  
 and substitution inferences, 374–376  
 singular referential purport, 308, 361, 415, 416, 418, 422, 425, 426, 432, 442, 571. *See also* objects; representation; singular terms  
 success of, 431, 433, 441, 443, 444  
 singular terms, 308, 323, 360–362, 370, 376, 385–391, 398–401, 404, 409–411, 413–419, 422, 426, 432–434, 443, 453, 454, 539, 560, 687n14, 694n17, 695n40  
 in compound sentences, 395–397  
 and content, 198, 396, 425, 453, 617, 623  
 dual substitutional characterization of, 385, 390, 397, 400  
 and logical vocabulary, 379–382, 393, 395  
 maximal/minimal substitutional requirement on, 419–426  
 and objects, xxi, 69–70, 288, 307, 326, 414–416, 578, 687n14  
 primary occurrences of, 342, 374–375, 378–379, 381, 389, 392, 394, 395  
 semantic/syntactic substitutional role of, 363, 376, 378, 380, 384–385  
 and sentences, 81, 82, 669n1, 688n29  
 substitution-inferential significance, 370–372, 376, 378, 400, 403  
 and substitution-structural roles, xv, 324, 369, 371–373, 378, 384–390, 401, 406, 424, 454, 619–620, 622  
 what are?, 361, 367, 399, 400  
 why are there any?, 376, 378–381, 384, 385, 399–401, 403, 414  
 singular thoughts, 567–574, 583  
 SIS. *See* substitution-inferential significance  
 SMSICs. *See* simple material substitution-inferential commitments  
 social dimension, 496, 497, 520, 590, 601–607  
 of inference, 138, 519–520, 586, 593, 605  
 social perspectives. *See* perspectives  
 social practices. *See* practice(s)  
 solidity of practices, 332, 528, 631–632, 686n55, 715n28  
 Solomon, W. David, 680n6  
 sortals, 304, 409–412, 420, 437–440, 445–446, 461, 471, 684nn36, 38, 690nn34, 39, 696n51. *See also* singular terms  
 Sosa, Ernest, 550–551, 566, 701n2  
 speaker's reference, 487–490, 511–513, 703n16  
 speech acts, xiii, xiv, 37, 82, 133, 142, 151, 157, 159, 172, 173, 182–186, 194, 230, 263, 343, 363, 367, 370, 531, 630, 636, 637, 649, 675n45. *See also* assertion(s); deontic scorekeeping; pragmatics  
 Spinoza, Benedictus de, 93  
 Sprachspielen. *See* language(s), games  
 SSR. *See* substitution-structural roles  
 Stalnaker, Robert, 148–149, 230, 337, 670n3, 671n12, 677n27

- stance(s), 55–62, 280, 287–289, 639, 644.  
*See also* intentional stance; interpretation; phenomenism  
discursive scorekeeping, 628, 630, 639–644, 648  
and intentionality, 55–62, 639–641  
states of affairs, 69–70, 76, 77, 85  
status(es). *See* deontic status(es); normative status(es)  
stimuli, 427–429. *See also* reliable differential responsive dispositions  
stipulation, 145–147, 671n8. *See also* conferring content; semantics  
STIT, 702n14  
Strawson, Peter F., 288, 369, 371  
strong *de re* ascriptions. *See* *de re* ascriptions, strong  
strong *de re* beliefs. *See* belief(s), strong *de re*  
structuralism, 80  
subordinate/superior, 241–243  
subsentential expressions, 136, 342, 360, 362–367, 375–378, 386, 392, 401, 402, 411, 414, 619–620, 640. *See also* singular terms; substitution  
substituted for, 301–303, 368–372, 376–386, 389, 392–398, 400–401, 403–405, 435, 436. *See also* singular terms; substitution; substitution-structural roles  
substituted in, 368–372, 378, 386, 404, 405, 409. *See also* substitution; substitution-structural roles  
substitution, xv, 65, 80, 112, 281–282, 344, 346–348, 367, 369, 372, 383, 391, 394, 403, 405, 430, 452, 509, 619, 623, 696n50. *See also* inference, substitution, and anaphora; semantics; singular terms  
and anaphora, 450–452, 465, 621–623  
and content, 347–349, 354–357  
and inference, 198, 283, 347–348, 352, 370–372, 414, 422, 425, 426  
intersubstitution, 308–310, 314, 324  
invariance under substitution, 281, 309–310, 317, 452  
and multivalued logic, 341–346, 352, 358, 359  
and subsentential expressions, 346–350, 359, 367, 399  
substitutional invariance, 281, 309–310, 317, 368  
substitutional variance, 367–369, 371, 379–381, 395, 405, 423, 451, 455  
substitutional commitments, xxii, 138, 282, 283, 319, 324, 325, 409, 419, 422, 426, 431, 433–440, 447, 451, 454, 459, 467, 472–474, 495, 499, 506, 640. *See also* explicit; inference, substitution, and anaphora; simple material substitution-inferential commitments; singular terms  
substitution frames, 368–369, 371, 376–381, 386, 394–396, 400–401, 403, 404, 406, 407, 692nn51, 54. *See also* predicates; substitution-structural roles  
substitution inferences, xv, 282, 319, 370–372, 374, 379–381, 450, 477, 619–622, 690n33. *See also* simple material substitution-inferential commitments  
substitution-inferential significance, 370–372, 376, 378, 400, 403  
asymmetric/symmetric, 377, 378, 384–385, 388, 391–395, 403  
substitution licenses. *See* identity, locutions  
substitution-structural roles, 367–371, 376, 378, 384–385, 400, 403, 404  
success, 13, 203, 236, 243, 280, 287, 290–291, 433, 521, 524, 526–528, 646  
supervenience, 47, 52, 292, 295–296, 628  
tactile Fregeanism, 579–583, 635. *See also* anaphora; content(s); senses  
taking-true, 5, 8, 11, 13, 46, 82, 202, 203, 228, 231, 233, 236, 240, 277, 278, 287–291, 297–299, 322, 521, 599, 625, 627. *See also* assertion(s); making true; phenomenism  
Tarski, Alfred, 84, 279, 317, 361, 364  
Tarskian contexts, 316, 318  
Tarski biconditionals, 302, 317  
terms. *See* singular terms  
testimony, 218, 234, 239, 531–533, 704n22. *See also* assertion(s); deontic scorekeeping; inheritance of entitlements  
authority of, 39–41, 175, 205, 234, 709n75  
‘that’, 68, 70, 329, 501–503, 505–506, 535, 538–539, 598, 608, 705n30  
theoretical concepts. *See* concepts, theoretical

- theoretical reasoning, 116, 520–522. *See also* practical reasoning  
 ‘thing’ as pseudosortal, 438  
 thoughts, 154, 567–568, 570. *See also* assertion(s); intentional states; propositional contents; rationality; sapience; singular thoughts  
 tokenings, 221, 303, 309–310, 432, 535, 537, 541, 566, 582, 664n10, 677n24, 685n40, 697n74, 705n26  
 token-recurrence structures, 432, 451–458, 469, 470, 488–489, 503, 539, 566, 573, 698n80. *See also* substitution  
 and anaphora, 432, 455, 456, 467, 472, 490, 499, 511, 580  
 asymmetric/symmetric, 455, 457, 490, 499, 564, 621–622, 706n33  
 tonk, 125  
 translation, 409–412, 535, 575  
 transparency, 128, 292–293, 570, 576, 583  
 triangulation, 426–432, 449, 467. *See also* reliable differential responsive dispositions  
 ‘true’, xvii, 112, 116, 137, 202, 278, 280, 283–285, 288, 297–300, 302, 304–307, 321–322, 327, 336, 499, 516–517, 527–528, 530. *See also* truth  
 expressive role of, 202, 278, 283–285, 324, 326–333, 568  
 truth, xii, xvii, 5, 8, 11–13, 17, 80–82, 99, 111, 112, 202, 236, 240, 277, 288, 289, 291, 298–303, 317, 318, 322–324, 344, 346, 349, 490, 498, 515, 516, 526, 528, 568, 594–596, 646, 649. *See also* making-true; representation; semantics; taking-true  
 conditions, 5, 6, 94, 129, 137, 278, 285, 326, 329, 356, 364, 513–517, 522, 599, 605, 669n89, 681n2, 710n95  
 correspondence theories of, 291, 317–318, 326, 330, 333  
 and *de re* ascriptions, 517, 595, 598, 646  
 explanatory role of, 286, 328–329  
 as expressive, 204, 232, 287, 299, 328–329, 333  
 and facts, 333, 622, 624, 625, 704n19  
 and inference, xvii, 5–6, 96–97, 104–105, 107–108, 277, 370, 519  
 performative account of, 287–288  
 as a property, 203, 303, 322, 323, 325–327  
 prosentential theory of, 301–306, 322  
 and redundancy of force, 291, 299–300  
 and representation, 70–72, 568  
 and success, 236, 286, 287, 291, 521, 527–528  
 and taking-true, 287, 292, 298, 322, 599, 627  
 talk, 17–18, 278, 279, 322, 515  
 truth-values, 80–81, 340, 344, 345, 352, 355–357, 617, 669–670n1. *See also* designatedness  
 trying, 294–295, 523–524. *See also* action(s)  
 twin-earth, 119–120, 331–332, 703n18  
 understanding, 4–5, 7, 13–14, 17, 31–32, 64–66, 74, 80, 90, 120, 215, 216, 294, 295, 478, 517, 526, 636. *See also* content(s); deontic scorekeeping; explicit; pragmatism  
 and interpretation, 508, 510, 658n39, 660n56  
 and knowledge, 209, 213–215  
 and meaning, 60, 62, 73, 88–89, 478  
 undertaking, 167, 194, 598, 627. *See also* acknowledgment; commitment(s); deontic attitudes; deontic scorekeeping  
 and acknowledgment, 194, 554  
 and attributing, 62, 161–163, 507, 554, 596–598, 677n11  
 unexplained explainers, 133, 360  
 unjustified justifiers, 221, 228. *See also* foundationalism  
 unrepeatability, 282, 432, 454, 466, 495, 592. *See also* anaphora; token-recurrence structures  
 validity, 10, 102–103, 117, 340, 345–346, 666n41, 687n8. *See also* formal inferences; logical vocabulary  
 assertional/inferential, 347–348  
 definition of, 346, 383  
 and material inference, 131, 133–134, 347–348, 359  
 and substitution, 104–105, 135, 348, 383, 619  
 verificationism, 121  
 vindication, 174, 176–177, 193, 204, 218, 226, 239, 265, 532–533. *See also* deference; deontic scorekeeping; justification  
 vocabulary. *See* logical vocabulary; normative vocabulary; semantic vocabulary

- volitions, 9, 258. *See also* practical commitments
- we, xi, xxi, 3–5, 7, 32, 39, 203, 275, 508, 587, 639, 641–645, 650, 716n36. *See also* demarcation; interpretation
- weak *de re* ascriptions. *See de re* ascriptions, weak
- Weber, Max, 48
- 'will', 245–246
- will, 32, 49, 233, 270, 271, 295, 661n69. *See also* action(s); akrasia; practical reasoning; trying
- Wittgenstein, Ludwig, 22, 52, 55, 74–76, 80–82, 111, 128, 178, 187, 226, 255–256, 333, 343, 461
- on interpretation, 61, 62, 508–510
- on norms, 13–16, 18, 46, 55, 200, 289, 603, 656n10
- pragmatism of, xii, 21–25, 29–30, 32, 42, 53, 55, 172, 199, 288
- on regress of rules, 20–22, 26, 45, 184, 206
- use of 'Regel', 64–66
- word-word relations, 306, 307, 325
- word-world relations, 306, 312–313, 324, 464
- Wright, Crispin, 37, 53–54, 656n10, 669n91